Revised September 21, 2006
(Page 178 replaced)

Defining and Validating Chronic Diseases: An Administrative Data Approach

July 2006

Manitoba Centre for Health Policy
Department of Community Health Sciences
Faculty of Medicine, University of Manitoba

Lisa Lix, PhD
Marina Yogendran, MSc
Charles Burchill, MSc
Colleen Metge, BSc (Pharm), PhD
Nancy McKeen, PhD, RN
David Moore, MD, PhD, DIC
Ruth Bond, MA

UNIVERSITY OF MANITOBA

This report is produced and published by the Manitoba Centre for Health Policy (MCHP). It is also available in PDF format on our website at **http://www.umanitoba.ca/centres/mchp/reports.htm**

Information concerning this report or any other report produced by MCHP can be obtained by contacting:

Manitoba Centre for Health Policy
Dept. of Community Health Sciences
Faculty of Medicine, University of Manitoba
4th Floor, Room 408
727 McDermot Avenue
Winnipeg, Manitoba, Canada R3E 3P5

Email: reports@cpe.umanitoba.ca
Phone: (204) 789 3819
Fax: (204) 789 3910

How to cite this report:

Lix L, Yogendran M, Burchill C, Metge C, McKeen N, Moore D, Bond R. Defining and Validating Chronic Diseases: An Administrative Data Approach. Winnipeg, Manitoba Centre for Health Policy, July 2006.

Legal Deposit:
Manitoba Legislative Library
National Library of Canada

ISBN 1-896489-25-7

1st Printing 07/25/2006

THE MANITOBA CENTRE FOR HEALTH POLICY

The Manitoba Centre for Health Policy (MCHP) is located within the Department of Community Health Sciences, Faculty of Medicine, University of Manitoba. The mission of MCHP is to provide accurate and timely information to health care decision-makers, analysts and providers, so they can offer services which are effective and efficient in maintaining and improving the health of Manitobans. Our researchers rely upon the unique Population Health Research Data Repository to describe and explain patterns of care and profiles of illness, and to explore other factors that influence health, including income, education, employment and social status. This Repository is unique in terms of its comprehensiveness, degree of integration, and orientation around an anonymized population registry.

Members of MCHP consult extensively with government officials, health care administrators, and clinicians to develop a research agenda that is topical and relevant. This strength along with its rigorous academic standards enable MCHP to contribute to the health policy process. MCHP undertakes several major research projects, such as this one, every year under contract to Manitoba Health. In addition, our researchers secure external funding by competing for other research grants. We are widely published and internationally recognized. Further, our researchers collaborate with a number of highly respected scientists from Canada, the U.S. and Europe.

We thank the University of Manitoba, Faculty of Medicine, Health Research Ethics Board for their review of this project. The Manitoba Centre for Health Policy complies with all legislative acts and regulations governing the protection and use of sensitive information. We implement strict policies and procedures to protect the privacy and security of anonymized data used to produce this report and we keep the provincial Health Information Privacy Committee informed of all work undertaken for Manitoba Health.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the contributions of many individuals whose efforts and expertise made it possible to produce this report. We appreciate the assistance of:

The project Working Group: Bev Cumming, Dr. Lawrence Elliott, Tannis Erickson, Dr. Sande Harlos, Dr. Dexter Harvey, Kelly McQuillen, Dr. Shahin Shooshtari, Roberta Vyse

External reviewers: Dr. Douglas Manuel, University of Toronto; Dr. Hude Quan, University of Calgary

Colleagues who provided expertise on the methods used in this research: Dr. Christine Peschken, Department of Internal Medicine, University of Manitoba; Dr. Alan Katz, Manitoba Centre for Health Policy; Dr. Anita Kozyrskyj, Manitoba Centre for Health Policy.

Programming support: Randy Walld.

Administrative support: Carole Ouelette, Shannon Lussier, Janine Harasymchuk, Jo-Anne Baribeau.

Literature searches and manuscript preparation: Nicole Fehr, Souradet Shaw, Sam Kovnats.

We acknowledge the financial support of the Department of Health of the Province of Manitoba. The results and conclusions are those of the authors and no official endorsement by Manitoba Health was intended or should be inferred. This report was prepared at the request of Manitoba Health, as part of the contract between the University of Manitoba and Manitoba Health.

TABLE OF CONTENTS

List of Tables

List of Appendix Tables

List of Figures

Executive Summary

Background

Chronic disease surveillance is critical for assessing the impact of disease on the population, identifying at-risk groups, and evaluating the effectiveness of population-based health promotion and disease prevention strategies. Multiple sources are used for chronic disease surveillance, including vital statistics records, disease-specific registries, population-based surveys, and administrative data. Evaluating the validity of each of these data sources is an important step in developing chronic disease surveillance systems.

Purpose and Objectives

This study examines the validity of administrative data, including hospital separations, physician billing claims, and prescription drug records, for monitoring the prevalence of selected chronic diseases in Manitoba. The specific objectives of this report are to:

1) Review the literature on the validity of administrative data for identifying chronic disease cases.
2) Evaluate the validity of multiple algorithms for identifying disease cases from Manitoba administrative data.
3) Test for differences in cross-sectional and longitudinal prevalence estimates for chronic disease algorithms.

Methods

The following diseases are the focus of the research (listed in alphabetical order): arthritis, asthma, coronary heart disease (CHD), diabetes, hypertension, and stroke. In addition, we reviewed the literature on the use of administrative data for identifying cases of congestive heart failure (CHF) and renal disease.

A chronic disease algorithm is a set of rules for identifying disease cases from administrative data. The elements of an algorithm include the type of data source, number of years of administrative data, diagnostic/medication code(s), and number of administrative data records (i.e., contacts) with a diagnostic/medication code(s).

Data sources for the research are hospital separations, physician billing claims, and prescription drug records in the Population Health Research Data Repository (PHRDR) housed at the Manitoba Centre for Health Policy (MCHP). Diagnostic codes in hospital and physician data are from the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) and medication codes in prescription drug data are from the Anatomic Therapeutic Chemical (ATC) coding system maintained by the World Health Organization (WHO) Collaborating Centre for Drug Statistics Methodology.

The algorithms selected for evaluation are based on the literature review results and consultations with clinicians and health services researchers. More than 200 algorithms are validated using self-report chronic disease data from cycle 1.1 of the Canadian Community Health Survey (CCHS). The survey data are linked to administrative data in the MCHP PHRDR. The CCHS validation cohort includes 5,589 adults 19 years of age and older and 833 youth between 12 and 18 years of age. Validation indices include the kappa (κ) statistic, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Youden's (1950) index. Youden's index is a summary measure of sensitivity and specificity, and is used to identify the optimal combination of these two validation indices for each chronic disease. CCHS data are used as the validation source because they are the only source of population-based data, besides administrative data, that can be used to identify individuals with multiple chronic diseases in Manitoba. Predictors of agreement between survey and administrative data, including age, sex, presence of comorbid conditions, region of residence, and income quintiles, are modeled using logistic regression.

Chronic disease prevalence estimates are generated for each algorithm. These prevalence estimates are computed for the Manitoba population from both cross-sectional and longitudinal administrative data for 1995/96 to 2003/04. Generalized linear regression models are used to test for differences in the population-based prevalence estimates from the algorithms. Specifically, the interactions between an algorithm effect and the sociodemographic effects of age, sex, region of residence, and income group are tested, to assess whether the relative rate of prevalence for different algorithms is constant across the sociodemographic characteristics of the population. Longitudinal trends in the prevalence estimates for different algorithms are also tested.

Key Findings

Administrative data exhibited very good to excellent validity for identifying cases of asthma, diabetes and hypertension. Administrative data exhibited fair to good validity for identifying cases of arthritis, osteoarthritis, non-fatal heart disease, and non-fatal stroke. Administrative data exhibited poor validity for identifying cases of rheumatoid arthritis. However, the latter result is likely due to bias in the validation data source.

Arthritis

Sixteen algorithms were validated for all forms of arthritis, rheumatoid arthritis, and osteoarthritis. For all forms of arthritis, agreement between survey and administrative data, as measured by κ, was highest (0.37) for the two-year algorithm based on one or more hospital separations or two or more physician billing claims, or one physician billing claim in combination with two or more prescription drug records. Youden's index was highest (0.40) for the algorithm based on one or more contacts in hospital separa-

tions, or two or more contacts in physician billing claims in five years. For rheumatoid arthritis, κ (0.17) and Youden's index (0.11) were highest for two algorithms, one of which was the five-year algorithm based on one or more physician claims. For osteoarthritis, κ (0.32) and Youden's index (0.39) were highest for the algorithm based on one or more physician billing claims in five years. However, for all forms of arthritis and rheumatoid arthritis, these two validation indices were almost equivalent for several other algorithms. The logistic regression analyses revealed that agreement between survey and administrative data was predicted by several sociodemographic variables, including age and income quintile.

Crude prevalence estimates for the algorithms with the maximum κ were 20.3% for all forms of arthritis, 1.6% for rheumatoid arthritis, and 13.2% for osteoarthritis. Crude prevalence estimates for the algorithms with the maximum value of Youden's index were 31.5% for all forms of arthritis, 1.0% for rheumatoid arthritis, and 13.2% for osteoarthritis for the Manitoba population 19 years of age and older. Analyses of the prevalence estimates revealed the relative rate for different algorithms was predicted by the sociodemographic variables of age, sex, region of residence and income quintile. There were no significant differences among the algorithms in the relative rate of change in prevalence over time.

Asthma

Twenty-eight algorithms were validated for each of the following age groups: 12 to 18 years, 19 to 49 years, and 50+ years, as well as for the combined age groups. For all ages, κ was highest (0.59) for the algorithm based on one or more hospital separations or two or more physician claims or two or more prescription drug records in five years. Youden's index was highest (0.73) for the algorithm based on one or more hospital separations or one or more physician billing claims or one or more prescription drug records in five years. However, it was equally high (0.72) for the algorithm based on one or more prescription drug records in five years. Agreement between survey and administrative data was predicted by age, presence of comorbid conditions, and income quintile.

The crude prevalence estimates for the Manitoba population 12 years of age and older were 11.6% for the algorithm with the maximum κ and 17.5% for the algorithm with the maximum value of Youden's index. Regression analyses revealed a number of interactions between the algorithm effect and the sociodemographic effects, indicating that the relative rate of asthma for different algorithms was not constant across the sociodemographic characteristics of the population. There were significant differences among the algorithms in the relative rate of change in asthma prevalence over time.

Coronary Heart Disease
A total of 20 algorithms were validated. The κ statistic was highest (0.55) for the algorithm based on one or more hospital separations or two or more physician billing claims, or one physician billing claim in combination with two or more prescription drug records in three years. Youden's index was highest (0.63) for the algorithm based on one or more hospital separations or one or more physician billing claims in five years. Logistic regression analysis revealed that agreement between survey and administrative data was predicted by age, sex, and the presence of comorbid conditions.

The algorithms with the highest values of κ and Youden's index resulted in crude prevalence estimates of 5.8% and 7.9%, respectively for the Manitoba population 19 years of age and older. Regression analyses revealed a number of statistically significant interactions between the algorithm effect and the sociodemographic variables, indicating that the relative rate of prevalence for different algorithms was not constant across the sociodemographic characteristics of the population. Almost all of the algorithms showed the same relative rate of change in heart disease prevalence over time.

Diabetes
Eighteen algorithms were validated. The κ statistic was highest (0.86) for the algorithm based on one or more hospital separations or one or more physician billing claims or two or more prescription drug records in two years. Youden's index was highest (0.88) for the algorithm based on one or more hospital separations or one or more physician billing claims or one or more prescription drug records in two years. However, several other two-year and three-year algorithms produced equally high values for this index (i.e., 0.83 – 0.87). Agreement between survey and administrative data was predicted by age and the presence of comorbid conditions.

Crude prevalence estimates for the algorithms with the highest κ and Youden's index were 6.3% and 7.5%, respectively for the Manitoba population 19 years of age and older. The regression analyses revealed a significant increase in the relative rate of diabetes over time, but the rate of change was not significantly different across the algorithms. There were no statistically significant interactions between the algorithm effect and the sociodemographic effects, suggesting that the relative rate of diabetes prevalence for different algorithms is constant across age, sex, income quintile, and region of residence of the population.

Hypertension
Eighteen algorithms were validated. The algorithm with the highest value of κ (0.70) was based on one or more hospital separations or one or more physician billing claims or two or more prescription drug records in one year. The algorithm with the highest value of Youden's index (0.79) was

based on one or more hospital separations or one or more physician billing claims or one or more prescription drug records in one year. Agreement between survey and administrative data was predicted by age and the presence of comorbid conditions.

The algorithms with the highest values of κ and Youden's index produced crude prevalence estimates of 21.6% and 22.5%, respectively for the Manitoba population 19 years of age and older. All algorithms showed the same relative rate of increase in hypertension prevalence over time. Regression analyses revealed interactions between the algorithm effect and the sociodemographic effects, indicating that the relative rate of prevalence for different algorithms is not constant across age, sex, income quintile, and region of residence of the population.

Stroke

A total of 24 algorithms were validated that varied in ICD-9-CM diagnostic codes as well as the number of years of administrative data and the type of data source. The algorithm with the highest κ (0.47) was based on one or more hospital separations or two or more physician billing claims, or one physician billing claim in combination with two or more prescription drug records in five years, and used a broad set of ICD-9-CM codes (430-438). The algorithm with the highest value of Youden's index (0.64) was based on one or more contacts in hospital separations or physician claims in five years, and the broad set of ICD-9-CM codes. Agreement between survey and administrative data was predicted by age, sex, the presence of comorbid conditions and income quintile.

The algorithms with the highest κ and Youden's index produced crude prevalence estimates of 2.9% and 3.8%, respectively for the Manitoba population 19 years of age and older. The relative rate of change in stroke prevalence over time was significantly different across the algorithms. Regression analyses revealed statistically significant interactions between the algorithm effect and several of the sociodemographic effects, indicating that the relative rate of prevalence for different algorithms was not constant across the sociodemographic characteristics of the population.

Conclusions and Recommendations

This research provides comparative information about the validity of a large number of algorithms for identifying cases of several chronic diseases from administrative data.

The research results illustrate the variations in validity that occur when the elements of a chronic disease algorithm are manipulated. The results demonstrate the substantial gains in sensitivity that can be achieved for some

chronic diseases, like asthma, when prescription drug data are used in addition to hospital separations and physician billing claims to identify chronic disease cases. These findings are relevant to researchers from other jurisdictions who seek to develop chronic disease algorithms that can be applied to their own administrative data.

Researchers and analysts can use the validation results contained in this report to select one or more algorithms to generate chronic disease prevalence estimates for the Manitoba population. Depending on the goals of future reports, chronic disease algorithms can be selected based on high agreement between survey and administrative data, high sensitivity to detect positive chronic disease cases, high specificity to avoid detecting false disease cases, or the maximum combination of sensitivity and specificity.

This research did not validate methods for estimating the incidence of chronic disease. Both empirical and model-based approaches have been proposed. These methods need to be validated before any recommendations can be made concerning the optimal methodology. Future research should also investigate opportunities to use statistical models, in combination with the expert advice of clinicians and researchers, to generate the optimal algorithm(s) for identifying disease cases from administrative data.

Chronic disease surveillance using administrative data, in addition to other validated data sources, has an essential role to play in public health initiatives at regional, provincial, and national levels. Future surveillance opportunities using the validated algorithms include comparative studies of chronic disease prevalence across socioeconomic groups and geographic regions of Manitoba, as well as cross-sectional and longitudinal studies of variations in health status and health service utilization for individuals with one or more chronic diseases.

CHAPTER 1: INTRODUCTION

A 2003 report estimates that almost 60 percent of Canadians 12 years of age or older have at least one chronic disease.

Chronic diseases are those conditions that are generally incurable, are often caused by a complex interaction of factors, and usually have a prolonged clinical course (Health Surveillance Coordination Division, 2003). Conditions like cardiovascular disease, respiratory illness, and diabetes have a large impact on the Canadian population. A 2003 report estimates that almost 60 percent of Canadians 12 years of age or older have at least one chronic disease (Schultz and Kopec, 2003). The total costs associated with caring for individuals with chronic disease in Canada are estimated to exceed $80 billion annually (Chronic Disease Prevention Alliance of Canada [CDPAC], 2006). In addition, the number of individuals with at least one chronic disease appears to be increasing. For example, it is estimated that there are more then 60,000 new cases of Type II diabetes in Canada annually (Health Canada, 2005).

Public health departments are being urged to "adjust to the epidemiological transition from communicable to chronic disease" (Frieden, 2004). Access to population-based chronic disease data is one critical factor in making this shift in orientation. These data are used to describe geographic and demographic variations in prevalence and incidence estimates, identify at-risk groups, and to examine trends over time in order to predict potential disease impact on the population in the future. This information is necessary to formulate public health policy around chronic disease treatment and prevention, and to evaluate the effectiveness of population-based health promotion and disease prevention strategies.

However, it is widely recognized that there is no single data source suitable for all aspects of chronic disease surveillance (Thacker et al., 1995). Data are compiled from multiple sources, including vitals statistics files, disease-specific registries, and population-based surveys. This report focusses on the use of administrative data, including hospital, physician, and prescription drug records, for chronic disease surveillance, and their validity for that purpose.

Administrative data have been used in numerous studies of chronic disease incidence and prevalence. For example, hospital discharge data and/or physician billing claims have been used both nationally and internationally to generate diabetes prevalence estimates (Hux et al., 2002; Maskarinec, 1997; Saydah et al., 2004), and are the basis for diabetes surveillance in Manitoba (Young et al., 1991; Blanchard et al., 1996; Blanchard, et al., 1997).

Administrative data are a potentially valuable tool for chronic disease surveillance because they are relatively easy to access and process, can be used to monitor a variety of diseases, and can provide both cross-sectional and longitudinal information about disease prevalence and incidence for entire popu-

Administrative data overcome several of the limitations associated with other sources of chronic disease surveillance data. However, because administrative data are collected for purposes of health system management and provider payment and not for chronic disease surveillance, it is important to assess their validity for the latter purpose.

lations. Administrative data overcome several of the limitations associated with other sources of surveillance data. For example, while vital statistics data are an accessible source in many jurisdictions, they cannot be used to monitor diseases with low case-mortality rates like arthritis. As well, cause of death is rarely attributed to a chronic disease itself, but rather to complications that arise from having the disease. Vital statistics data are not always a timely source of information on disease prevalence because of potentially long lag times between changes in the population prevalence of a disease and its detection using mortality data. While disease-specific clinical registries can usually provide accurate estimates of disease prevalence, they are expensive and time-consuming to establish and maintain, and are being subjected to increased scrutiny under current health privacy legislation. As well, registries which are specific to particular geographic areas, clinical groups, or facilities do not provide estimates of incidence or prevalence that can be generalized to larger populations. Longitudinal population-based health surveys, another well-established source of chronic disease data, suffer from respondent attrition, which can result in inaccurate estimates of incidence and prevalence.

However, because administrative data are collected for purposes of health system management and provider payment and not for chronic disease surveillance, it is important to assess their validity for the latter purpose. Fowles et al. (1998) compared U.S. Medicare claims data to data abstracted from medical charts to assess the sensitivity and specificity of the former for identifying individuals with 17 different chronic diseases. Sensitivity refers to how well administrative data detect the presence of a disease for individuals who actually have it, while specificity refers to how well the administrative data avoids the problem of falsely detecting individuals who do not have the disease. The sensitivity of administrative data varied substantially, from a low of 20% for alcohol and drug abuse to a high of 100% for diabetes. Besides diabetes, chronic diseases for which administrative data had the highest sensitivity were hypertension (90%), asthma (82%), mental health conditions (71%), and joint problems (68%). Specificity was very high (i.e., 95% or above) for all diseases with the exception of joint problems (88%).

Robinson et al. (1997) compared Manitoba administrative data (hospital and physician records) to survey data from the Manitoba Heart Health Project for diabetes, hypertension, elevated cholesterol, stroke, acute myocardial infarction, and non-specific forms of heart disease. Agreement between the two sources, as measured by the kappa (κ) statistic, was highest for diabetes (i.e., $\kappa > .70$) and hypertension ($\kappa > .50$), and lowest for non-specific heart disease ($\kappa = .38$). However, for any given disease, kappa varied with the algorithm, the set of rules used to identify disease cases. Specifically, Robinson et al. (1997) considered the effect of both the number of years of administrative data required to establish disease diagnosis, and the number

of times a diagnosis code was required to appear in the administrative data to confirm a disease case (i.e., the number of required contacts). As expected, there was a positive relationship between κ and the number of years of data and a negative relationship between κ and the number of required contacts.

Hospital data are the sole source of prevalence and incidence estimates in some studies (Brameld et al., 2003; Huff et al., 1996). The diagnostic and medical procedure data contained in hospital records is coded by a medical archivist, which increases the likelihood of accurate documentation. In contrast, the diagnostic information contained in the billing claims of physicians is often not confirmed at the source because diagnosis is not linked to physician remuneration. Studies have been undertaken to assess the validity of different sources of administrative data for identifying chronic disease cases. Wilchesky et al. (2004) compared diagnostic information in Quebec physician records to medical chart data abstracted from the primary care physician for more than 14,000 individuals, for 14 chronic diseases associated with drug-disease contraindications. Among these were hypertension, renal failure, diabetes, asthma, chronic obstructive pulmonary disease (COPD), coronary heart disease (CHD), and dementia. Specificity was very high for all investigated conditions; the lowest value was for hypertension (82%). Sensitivity was much more variable. It was highest for glaucoma (76%), hypertension (69%), and diabetes (64%). It was low for renal disease (19%) and dementia (19%), and moderate for asthma (43%), COPD (46%), and congestive heart failure ([CHF], 42%). The authors recommend that future studies examine the improvements in sensitivity and specificity that can be achieved by using multiple administrative data sources to identify disease cases.

The importance of this recommendation is emphasized in two studies which examined the percentage of chronic disease cases that can be validly identified from different administrative data sources. Wigertz and Westerling (2001) found that only 17% of individuals with a diagnosis for diabetes in administrative data (i.e., hospital and medical records), were identified solely from inpatient hospital records, while 78% were identified solely from primary care physician records. For asthma, only 8% of individuals were identified exclusively from inpatient hospital records, while 85% were identified solely from primary care physician data. In contrast, for cerebrovascular disease, 53% of individuals were identified from inpatient hospital records alone, and 67% were identified only from primary care physician data. Similarly, Robinson et al. (1997) found that 5% of individuals with a diagnosis of diabetes in administrative data (i.e., combined medical and hospital records) were identified from hospital records alone, but 68% were detected solely from physician records. For hypertension, only 2% were identified from hospital records but 86% were found only in medical claims. In con-

trast to Wigertz and Westerling, Robinson et al. found that only 7% of stroke cases were identified solely in the hospital data, but 54% were detected solely from physician records.

While many studies are limited to hospital and physician data for identifying individuals with chronic diseases, prescription drug records have been recognized as another potentially valuable source of data (Sartor and Walckiers, 1995). For example, Cricelli et al. (2003) used the Health Search Database in Italy to identify individuals with several different chronic diseases. This database combines linked administrative records on prescriptions, clinical events and diagnoses, hospital admission, and cause of death. Prevalence estimates obtained from self-reported survey data and administrative data were compared for the following diseases: diabetes, COPD, gastroduodenal ulcer, and hypertension. Similar prevalence estimates were obtained from the two data sources for diabetes and hypertension, although no quantitative measure of the degree of concordance between the measures was calculated. The estimates were less similar for gastroduodenal ulcer and COPD. The authors suggest that the consistency of self-reported and administrative data for hypertension and diabetes occurs because these are diseases that have a "clear-cut diagnosis" (p. 255) and require continuous medical treatment. Rector et al. (2004) used a combination of physician, facility (i.e., hospital), and pharmacy claims in the U.S. to conduct an extensive validation of algorithms for six chronic diseases: hypertension, CHF, chronic lung disease, arthritis, glaucoma, and diabetes. The use of multiple sources resulted in high specificity for all diseases, and high sensitivity for all but arthritis, lung disease, and CHF. Maio et al. (2005) used prescription drug records for one region of Italy to estimate the prevalence of disease in each of 31 chronic condition groups, including cardiovascular diseases, rheumatologic conditions, respiratory illness, gastrointestinal disease, and psychiatric disease.

1.1 Purpose and Objectives

The purpose of this report is to examine the validity of administrative data for monitoring the prevalence of chronic disease in Manitoba. The specific objectives are to:

1) Review the literature on the validity of administrative data for identifying chronic disease cases.
2) Evaluate the validity of multiple algorithms for identifying disease cases from Manitoba administrative data.
3) Test for differences in cross-sectional and longitudinal prevalence estimates for chronic disease algorithms.

In this study, we conduct a systematic evaluation of the improvement in validity that can be achieved when prescription drug data, as well as hospital and physician data, are used to identify disease cases.

This study extends previous research on methods for identifying chronic disease cases from administrative data in several important ways. First, only a small number of studies have systematically compared the sensitivity and specificity of multiple algorithms for identifying chronic disease cases (Robinson et al., 1997; Rector et al., 2004). In this study, we examine the changes in validity that result when the elements of a chronic disease algorithm, including the data source, number of years of data, and number of required occurrences of a diagnostic or medication code are varied. Second, the number of studies that have examined the gains in agreement, sensitivity, specificity, and predictive value that are achieved when prescription drug data are used in addition to hospital and physician data is limited. In this study, we conduct a systematic evaluation of the improvement in validity that can be achieved when prescription drug data, as well as hospital and physician data, are used to identify disease cases. Finally, this study, unlike previous research, tests the potential confounding effects of sociodemographic variables like age, sex, and income group in estimating disease prevalence from different algorithms. It also tests for differences in the longitudinal trends in prevalence estimates from different chronic disease algorithms.

1.2 Report Organization

The report is organized as follows. Chapter 2 describes the methods adopted in this research. It begins with a description of the group-based consensus process used to select the chronic diseases that were the focus of the research. The techniques used to define and validate chronic disease algorithms are described in this chapter, as are the methods to calculate and test cross-sectional and longitudinal prevalence estimates for the Manitoba population.

Chapters 3 through 8 are each devoted to one of the chronic diseases that was the focus of in-depth research. These are discussed in alphabetical order: arthritis, asthma, CHD, diabetes, hypertension, and stroke. Each chapter begins with a review of published research. The validity of multiple algorithms to identify disease cases is compared. Provincial prevalence estimates are reported and tested.

Chapter 9 concludes the report, with a discussion of recommended algorithms for the chronic diseases that were investigated in this research. As well, we review the literature on methods for identifying cases for two diseases that were not investigated in this study, but that were of interest to the Working Group. These are renal disease and CHF. Opportunities for further research on the methods for identifying chronic disease cases from administrative data are also discussed. The chapter ends with a discussion of the role of population-based data in provincial chronic disease surveillance initiatives.

CHAPTER 2: METHODS

This chapter describes the methods used to identify chronic disease cases from administrative data.

This chapter describes the methods used to identify chronic disease cases from administrative data. It begins with a description of the group-based consensus process to select the chronic diseases that were the focus of the research. The administrative data sources in the Population Health Research Data Repository housed at Manitoba Centre for Health Policy (MCHP) that were used to conduct the research, are enumerated. The diagnostic and medication codes to identify disease cases (i.e., individuals with a chronic disease) are listed. Next, the validation of the chronic disease algorithms is described. The methods used to obtain provincial cross-sectional and longitudinal prevalence estimates from administrative data are also described. A discussion of the descriptive and inferential analyses applied to test provincial prevalence estimates from chronic disease algorithms concludes the chapter.

2.1 Group-Based Consensus Process for Chronic Disease Selection

Members of the Working Group for this research were initially tasked with identifying the chronic diseases for which algorithms would be developed and validated. Five criteria were presented to the Working Group to facilitate the process of establishing a priority list of chronic diseases:

1. **Magnitude of prevalence estimates:** Prevalence estimates were obtained through a search of the research literature. For each chronic disease, this criterion provides an indication of the burden of illness in the population.
2. **Availability of a validation data source:** Manitoba and national data sources to validate the administrative data algorithms were identified through literature searches and discussions with researchers and Manitoba Health representatives. This criterion provides an indication of the feasibility of undertaking algorithm validation for each disease.
3. **Health system utilization:** Information on rates of physician visits and/or hospitalizations associated with a disease was collected from the published literature. This criterion provides an indication of the importance or significance of a disease to the health system.
4. **Health system costs:** Information on the direct costs of each disease to the health system was collected from the published literature. This criterion provides another indication of the importance or significance of a disease to the health system.
5. **Body system:** This criterion provides an indication of the body systems covered by different chronic diseases.

A preliminary list of chronic diseases was formulated, and each Working Group member rank ordered the diseases according to their priority for the

research. The criteria described above were considered in this ranking process. The following 10 diseases were given the highest research priority by the Working Group:

1. Hypertension
2. Coronary Heart Disease (CHD)
3. Renal Disease
4. Asthma
5. Arthritis (Both Rheumatoid Arthritis and Osteoarthritis)
6. Stroke
7. Congestive Heart Failure (CHF)
8. Depression
9. Dementia
10. Diabetes

In subsequent discussions, members of the Working Group noted that research conducted by MCHP and other University of Manitoba researchers had produced validated algorithms for depression and dementia. Accordingly, these chronic diseases were dropped from further consideration.

Subsequent to the Working Group consensus meeting, we determined that a validation data source for renal disease was not available in the PHRDR. Validation data existed for CHF, but the number of validation cases was too small to produce reliable results. Therefore, these two chronic diseases were dropped from further consideration. However, opportunities for further research on methods to identify renal disease cases and CHF cases from administrative data are discussed in Chapter 9.

2.2 Methods for Review of Literature

MEDLINE was searched for studies published between 1990 and 2005 inclusive that reported on methods for identifying chronic disease cases from administrative data. The following terms were used: administrative data, databases, population surveillance, chronic disease, diabetes, heart disease, coronary/ischemic heart disease, arthritis, rheumatoid arthritis, osteoarthritis, hypertension, high blood pressure, asthma, respiratory disease, and stroke. The reference lists of identified papers were checked for additional citations.

The websites of relevant research groups that conduct population-based research using administrative data (i.e., Institute of Clinical Evaluative Sciences [ICES]) were reviewed. In addition, the Statistics Canada and Health Canada websites were reviewed for relevant documents.

2.3 Sources of Administrative Data to Define Chronic Disease Algorithms

Administrative data to define chronic disease algorithms were obtained from the Population Health Research Data Repository (PHRDR) housed at MCHP. The sources of the data included hospital separations, physician billing claims, and prescription drug records.

Hospital abstracts are completed at the point of discharge for all separations from acute care facilities in Manitoba. They include up to 16 diagnosis codes based on the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). Only inpatient separations were used to define the algorithms.

Manitoba physicians who are paid on a fee-for-service basis submit billing claims to Manitoba Health. These claims contain a single ICD-9 diagnostic code. A small proportion of physicians are salaried, but most submit parallel billing claims for administrative purposes. Accordingly, almost all contacts with Manitoba physicians are captured in the Repository.

Table 1 summarizes the three-digit ICD-9-CM codes that were selected to define the chronic diseases from hospital separations and physician billing claims. Justifications for the selection of these codes are provided in subsequent chapters of the report.

Table 1 : Diagnosis codes used to define chronic diseases with administrative data

Disease	ICD-9-CM Diagnosis Codes
Arthritis	714: rheumatoid arthritis 715: osteoarthritis 446, 710: connective tissue disorders (446 = Polyarteritis nodosa and allied conditions; 710 = Diffuse diseases of connective tissue) 720: ankylosing spondylitis 274: gout 711-713, 716, 717, 718, 719, 721, 725-729, 739: other arthritis and related conditions
Asthma	493: asthma
Coronary Heart Disease	410: acute myocardial infarction 411: other acute and subacute forms of ischemic heart disease 412: history of myocardial infarction 413: angina pectoris 414: All other forms of chronic ischemic heart disease
Diabetes	250: diabetes mellitus
Hypertension	401: essential hypertension
Stroke	430-438: cerebrovascular disease

Source: Manitoba Centre for Health Policy, 2006

The third source of data for defining chronic disease algorithms, prescription drug records, are maintained in the pharmaceutical database, which is a subset of the Drug Programs Information Network (DPIN), an electronic, on-line, point-of-sale prescription drug database connecting all retail pharmacies in Manitoba to a central database. The DPIN system was initiated in 1995. It captures information about prescription drug dispensations for all Manitoba residents, regardless of insurance coverage or final payer. DPIN contains a variety of information about each pharmaceutical dispensation, including the drug identification number (DIN). The DINs are linked to Anatomic Therapeutic Chemical (ATC) codes in the Drug Product Directory maintained by Health Canada. A list of all DINs for the drugs covered by the Manitoba Pharmacare program is maintained in the MCHP Master Formulary. The Formulary contains classification codes for drugs, generic product name, and brand name, and is derived from the Manitoba Formulary.

The ATC system is maintained by the World Health Organization (WHO) Collaborating Centre for Drug Statistics Methodology, and was first published in 1976. Under this system, drugs are divided into groups at each of five levels according to the organ or system on which they act and/or their therapeutic and chemical characteristics. The levels are: (1) anatomical group, (2) therapeutic main group, (3) therapeutic/pharmacological subgroup, (4) chemical/therapeutic/pharmacological subgroup, and (5) subgroup for chemical substance. The methodology that was used to select ATC codes and DINs for this research is described in subsequent chapters of the report.

2.4 Validating Chronic Disease Algorithms

Validation Data Source

Data from the Canadian Community Health Survey (CCHS), cycle 1.1, collected between September 2000 and November 2001 were used to evaluate the agreement, sensitivity, specificity, positive predictive value, and negative predictive value of each chronic disease algorithm selected for investigation. The CCHS was conducted by Statistics Canada to provide regular and timely cross-sectional estimates of health determinants, health status and health system utilization for a total of 136 health regions in Canada, including the territories. Survey respondents were sampled from 11 regions in Manitoba. Respondents were 12 years of age and older; the sampling methodology was designed to ensure over-representation of youth under 19 years of age and seniors 65 years of age and older. Sample sizes were chosen to produce reliable estimates at the health region level. The sample size allocated to each province was based on total population, and then each province's sample size was allocated among its health regions proportionately to the square root of the estimated population of a region.

Manitoba CCHS cycle 1.1 data were linked to the administrative data in the Data Repository housed at MCHP using the unique encrypted personal health identification number (PHIN) for those individuals who consented to the linkage. The linkage was successful for 6,812 Manitoba residents 12 years of age and older. From this sub-sample, the cohort of survey respondents with at least five years of continuous coverage under the Manitoba Health Services Insurance Plan prior to the date of their CCHS interview was created. Information on health insurance coverage was obtained from the population registry. The registry contains coverage information for individuals registered with the Plan since 1970. For all diseases, the validation cohort was limited to individuals with five years of coverage because some chronic disease algorithms were based on as many as five years of administrative data.

The validation cohort included 5,589 adult survey respondents 19 years of age and older, and 833 youth survey respondents between 12 and 18 years of age. Slightly less than half (46.1%) of the respondents in the adult validation cohort were 50 years of age or older and more than half (55.3%) were female. Almost one-quarter (23.4%) of the respondents in the adult validation cohort were from the Winnipeg Regional Health Authority (RHA), 12.2% were from the northern RHAs, and the remaining 64.4% were from the southern rural RHAs. Among members of the youth validation cohort, 50.2% were female, 20.7% were from the Winnipeg RHA, 15.4% were from the northern RHAs, and 63.9% were from southern rural RHAs.

Only the adult cohort was used to validate chronic disease algorithms for arthritis, CHD, diabetes, hypertension, and stroke. For asthma, the algorithms were validated using both the adult and youth cohorts, and separate validations were conducted for adults between 19 and 49 years of age, and 50 years of age or older.

Chronic disease algorithms were defined using one, two, three, or five years of administrative data. The years of administrative data that were searched to determine whether a survey respondent could be classified as a disease case in the administrative data were based on the date of the interview. For example, if an individual was interviewed on October 31, 2001 then a one-year algorithm was defined using data from November 1, 2000 to October 31, 2001.

Validation Questions

The CCHS questions used to identify survey respondents with each of the investigated chronic diseases are listed in Table 2. Respondents were asked to

report chronic diseases according to the following directions:
Now I'd like to ask about certain chronic health conditions which you may have. We are interested in 'long-term conditions' that have lasted or are expected to last 6 months or more and that have been diagnosed by a health professional.

These directions were repeated to survey respondents throughout their completion of the set of questions about chronic diseases.

Table 2: CCHS questions used to identify survey respondents with chronic diseases

Disease	Relevant CCHS Question(s)
Arthritis	Do you have arthritis or rheumatism, excluding fibromyalgia? What kind of arthritis do you have?
Asthma	Do you have asthma?
Diabetes	Do you have diabetes?
Hypertension	Do you have high blood pressure?
Heart Disease	Do you have heart disease? Do you have congestive heart failure? (Used as an exclusion variable)
Stroke	Do you suffer from the effects of a stroke?

Source: Manitoba Centre for Health Policy, 2006

The number and percent of individuals in the adult and youth validation cohorts who reported having the selected chronic diseases is found in Table 3. As well, this table reports the crude provincial prevalence of each chronic disease using the CCHS self-report data. These prevalence estimates and their associated 95% confidence intervals (CIs) are calculated using the survey sample weights for the Manitoba population. The CIs were computed using the bootstrap methodology provided for the CCHS.

The results in Table 3 show that more than 1,000 individuals reported having either arthritis or hypertension, but smaller numbers of respondents had self-reported asthma, diabetes, non-fatal CHD, and non-fatal stroke. For asthma, the adult validation cohort was divided into the younger (i.e., 19 to 49 years) and older (i.e., 50+ years) age groups. There were $N = 228$ (7.6%) survey respondents who reported having asthma in the younger adult age group and $N = 190$ (7.4%) survey respondents who reported having asthma in the older adult age group.

Table 3: Number (percent) of respondents in the CCHS validation cohort with each chronic disease and crude provincial prevalence using CCHS sample weights

Disease	Adult Validation Cohort *N* (%)	Youth Validation Cohort *N* (%)	Crude Prevalence (%) (95% CI)
Arthritis (All Forms)	1,344 (24.1)	--	18.6 (17.4 – 19.8)
Rheumatoid Arthritis	459 (8.2)	--	7.4 (6.5 – 8.3)
Osteoarthritis	601 (10.8)	--	7.7 (6.9 – 8.5)
Asthma	418 (7.5)	111(13.5)	8.6 (7.6 – 9.6)
Diabetes	337 (6.0)	--	4.5 (3.9 – 5.2)
Heart Disease	371 (6.6)	--	5.1 (4.4 – 5.7)
Hypertension	1,033 (18.5)	--	15.3 (14.2 – 16.5)
Stroke	108 (1.9)	--	1.5 (1.1 – 1.8)

Note: Crude prevalence estimates were obtained using CCHS sample weights. The bootstrap was used to obtain the 95% confidence intervals.

Source: Manitoba Centre for Health Policy, 2006

For asthma, one set of validation analyses were conducted in which individuals with COPD were excluded. There were a total of 23 survey respondents in the validation cohort who indicated that they had both asthma and COPD; all but one of these respondents was 50 years of age or older. The reason for conducting the validation analysis with this exclusion is that COPD is often misdiagnosed or treated as asthma even though these are distinct conditions.

Survey data have been used in several previous studies to validate algorithms for the chronic diseases investigated in this report. The validity of survey data for identifying disease cases has also been evaluated using medical chart data. For example, Hux et al. (2002) validated two diabetes algorithms using both National Population Health Survey (NPHS) data and medical chart data. Estimates of sensitivity and specificity were similar for both validation sources. For example, sensitivity of their first diabetes algorithm, one or more hospital separations or one or more physician service claims in Ontario administrative data in a one-year period, was 90% in survey data and 91% in medical chart data. However, estimates of positive predictive value (PPV) were lower for survey data (44%) than for medical chart review (64%). O'Connor et al. (1998) found that compared to medical charts, survey data resulted in very high sensitivity (95%), specificity (99%), and positive predictive value (81%) for identifying previous cases of heart attack. For overall heart disease, survey data were less sensitive (58%), but maintained very high specificity (99%) and PPV (81%). Okura et al. (2004) compared self-report data and medical chart data for heart failure, diabetes, hypertension, myocardial infarction, and stroke. Sensitivity of self-report measures was highest for "medical conditions that are well defined and relatively easily diagnosed" (p. 1101), including myocardial infarction (90%), hypertension (82%), and stroke (78%). Sensitivity was lower for heart failure (69%) and diabetes (66%). Specificity was above 90% for all conditions. Martin et al. (2000) compared survey data and medical chart data for three chronic health conditions: hypertension, hypercholesterolemia, and diabetes.

Sensitivity of the survey data was highest for hypertension (83%), moderate for diabetes (73%), and lowest for hypercholesterolemia (59%). Specificity was highest for diabetes (99.3%), moderate for hypercholesterolemia, and lowest for hypertension (81.4%). Rector et al. (2004) used survey data to validate chronic disease algorithms defined from Medicare data; the following conditions were investigated: hypertension, heart failure, chronic lung disease, arthritis, glaucoma, and diabetes.

While medical chart data is considered by some to be an ideal gold standard, previous research recommends a cautious approach when conducting population-based validation studies using medical charts.

The CCHS was selected as the validation source in this study because next to administrative data, it is the only source for obtaining population-based chronic disease prevalence estimates in Manitoba. As well, the sample size for Manitoba in cycle 1.1 of the CCHS was large enough to ensure sufficient numbers of positive disease cases even for relatively rare conditions such as rheumatoid arthritis and stroke. While medical chart data is considered by some to be an ideal gold standard, previous research recommends a cautious approach when conducting population-based validation studies using medical charts. Hux et al. (2002) observed that for diabetes, validation of population-based algorithms with medical chart data was difficult. "Migration between providers and lack of efficient vertical integration of care may contribute to under detection if data are abstracted from the office chart of a single practitioner because that practitioner may not represent the patient's regular source of care" (p. 515). Wilchesky et al. (2004) made a similar observation in their medical chart-based validation of chronic disease algorithms from administrative data. Sensitivity was lower when only a single physician's chart (i.e., the regular provider) was reviewed than when the charts of all physicians that an individual had visited were reviewed. There are other problems that may arise from using medical chart data for validation studies. For example, privacy or security concerns of health care providers may result in restricted access to patient charts, even with patient consent. A large number of charts must be sampled in order to obtain a sufficient number of positive disease cases for rare chronic diseases. Finally, some chart entries may be difficult to decipher or interpret (Ritter et al., 2001), which can bias validation results.

However, self-report data may not be an unbiased gold standard. Under-reporting or over-reporting of some chronic diseases in surveys may occur because respondents are not aware of all the diagnoses reported in a patient chart, or because of the lack of correspondence between the lay language used in surveys and the clinical terminology used to record diagnoses in the medical chart. Accurate reporting is more likely to occur for conditions that result in frequent contacts with a health professional.

Validation Methods

Six indices were used to evaluate the validity of chronic disease algorithms. The first was the kappa statistic (κ), a measure of agreement between two

sources, each of which is measured on a binary scale (i.e., disease present/absent). The interpretation of κ used in this report is (Altman, 1991):

- Poor agreement: $\kappa < 0.20$
- Fair agreement: $\kappa = 0.20$ to 0.39
- Moderate agreement: $\kappa = 0.40$ to 0.59
- Good agreement: $\kappa = 0.60$ to 0.79
- Very good agreement: $\kappa = 0.80$ to 1.00

Ninety-five percent CIs were calculated for κ. These intervals are calculated using the square-root of the asymptotic variance and a critical value from the standard normal distribution.

Sensitivity and specificity were calculated for each chronic disease algorithm. Sensitivity was defined as the percentage of true positives an algorithm detects among all positive disease cases. Positive disease cases are survey respondents in the CCHS validation cohort who reported having the specified disease. Specificity was defined as the percentage of true negatives an algorithm detects among all the negative disease cases. Negative disease cases are survey respondents in the CCHS validation cohort who did not report having the specified disease. For both sensitivity and specificity, 95% CIs were calculated. These confidence intervals are based on the asymptotic standard error and a critical value from the standard normal distribution.

Positive and negative predictive values (NPV) are also reported for each chronic disease algorithm. PPV refers to the percentage of individuals with a positive result for an algorithm among those who reported having the disease. NPV refers to the percentage of individuals with a negative result for an algorithm who did not report having the disease. Ninety-five percent CIs were also calculated for PPV and NPV, and were based on the asymptotic standard error and a critical value from the standard normal distribution. Figure 1 illustrates the calculations of sensitivity, specificity, PPV, and NPV.

Figure 1: Calculation of Validation Indices for Chronic Disease Algorithms

CCHS Data

Repository Data	Has Disease	Does Not Have Disease
Has Disease	A	B
Does Not Have Disease	C	D

Sensitivity = A/(A + C)* 100
Specificity = D/(B + D)* 100
PPV = A/(A +B)* 100
NPV = D/(C + D)* 100

Source: Manitoba Centre for Health Policy, 2006

Youden's (1950) index, which combines information on sensitivity and specificity, was computed for each algorithm. The index is defined as sensitivity + specificity – 1, where sensitivity and specificity are calculated as proportions. Youden's index has minimum and maximum values of –1 and +1, respectively, with a value of +1 representing the optimal value for an algorithm.

Finally, logistic regression analysis was used to test the sociodemographic variables associated with agreement between survey and administrative data. Only the algorithm which resulted in the maximum value of the κ statistic was selected for the regression analysis for each chronic disease. A binary outcome variable that defined agreement (yes/no) between survey and administrative data agreement was created. Agreement between the two sources of data is represented by cells A and D in Figure 1 and disagreement between the two sources of data is represented by cells B and C.

The explanatory variables in the logistic regression models were age group (i.e., 10-year groupings were used in most models), sex, region of residence (i.e., rural north RHAs, rural south RHAs, Winnipeg RHA), income adequacy quintile, and the presence of comorbid conditions. Income adequacy is a variable developed by CCHS methodologists. Each survey respondent was assigned to a quintile (i.e., approximately one-fifth of all survey respondents) using an algorithm based on total household income and number of persons living in the household. Using these two variables to assign income adequacy quintile means that individuals living in households with the same income but different numbers of household residents could be assigned to different quintiles. A comorbidity variable was included in the models, because previous research has demonstrated that the agreement between survey data and medical chart data is influenced by the presence of comorbid conditions (Okura et al., 2004). In the models for diabetes and stroke, the presence (yes/no) of either of the comorbid conditions of heart disease and hypertension was included as an explanatory variable. In the model for heart disease, the presence of either of the comorbid conditions of diabetes and hypertension was included as an explanatory variable. In the model for hypertension, the presence of either of the comorbid conditions of diabetes and heart disease was included. In the model for asthma, the presence of allergies as well as emphysema or COPD was included as an explanatory variable. There were no comorbid conditions included in the model for all forms of arthritis, rheumatoid arthritis, or osteoarthritis. Age group, sex, and region of residence were defined from the population registry, income adequacy quintiles were defined from the CCHS data, and the presence of comorbid conditions was also defined from the CCHS data.

All variables were treated as categorical in the regression analyses. Odds ratios (ORs) and their 95% CIs are reported. The fit of each logistic regression model was assessed using the Hosmer-Lemeshow test.

2.5 Calculating Provincial Prevalence Estimates

The population registry in the PHRDR housed at MCHP was used to define population cohorts to derive numerator and denominator data for calculating crude provincial prevalence estimates for each algorithm. The registry includes health insurance program coverage information in addition to information on demographic characteristics (e.g. age and sex), and location of residence for each resident of Manitoba eligible to receive health services. For all diseases with the exception of asthma, provincial prevalence estimates were calculated for the population 19 years of age and older. Asthma prevalence estimates were calculated for the population 12 years of age and older.

Cross-sectional provincial prevalence estimates were calculated to facilitate comparisons among the chronic disease algorithms at a single point or period in time. Table 4 lists the years that were used to calculate cross-sectional estimates based on algorithms defined for one, two, three, and five years of administrative data. For example, all estimates based on one year of data were defined for the Manitoba population continuously registered with the Manitoba Health Services Insurance Plan for the period April 1, 2002 to March 31, 2003. Similarly, all estimates based on two years of data were defined for the Manitoba population continuously registered for the period April 1, 2001 to March 31, 2003.

Table 4: Time periods used to define cross-sectional provincial chronic disease prevalence estimates

# Years	Time Period
1	April 1, 2002 to March 31, 2003
2	April 1, 2001 to March 31, 2003
3	April 1, 2000 to March 31, 2003
5	April 1, 1998 to March 31, 2003

Source: Manitoba Centre for Health Policy, 2006

Next, crude provincial prevalence estimates were calculated over time, to facilitate longitudinal comparisons among the chronic disease algorithms. Longitudinal estimates were calculated for algorithms based on one, two, or three years of administrative data as follows:

- One-year estimates were calculated for populations defined for each of the following five years:
  - April 1st 1999 to March 31st 2000
  - April 1st 2000 to March 31st 2001
  - April 1st 2001 to March 31st 2002
  - April 1st 2002 to March 31st 2003
  - April 1st 2003 to March 31st 2004

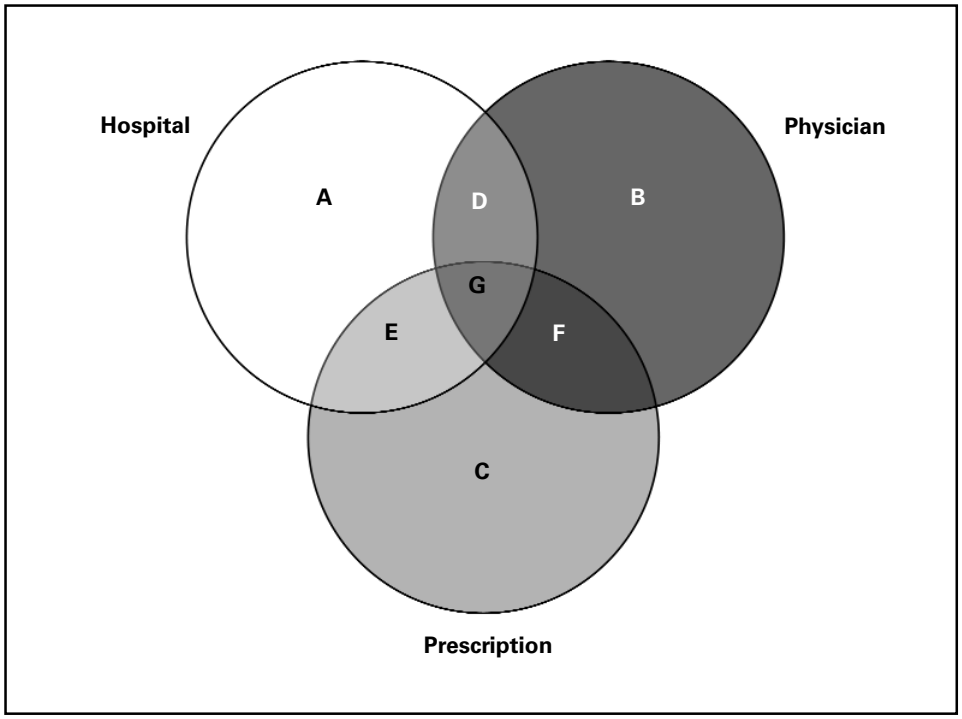
- Two-year estimates were calculated for populations defined for each of the following four time periods:
  - April 1st 1996 to March 31st 1998
  - April 1st 1998 to March 31st 2000
  - April 1st 2000 to March 31st 2002
  - April 1st 2002 to March 31st 2004
- Three-year estimates were calculated for populations defined for each of the following three time periods:
  - April 1st 1995 to March 31st 1998
  - April 1st 1998 to March 31st 2001
  - April 1st 2001 to March 31st 2004

Longitudinal prevalence estimates for algorithms based on five years of administrative data were not examined because there were an insufficient number of years of prescription drug data available to compute estimates for more than a single time period.

Analyses of Provincial Prevalence Estimates

For the cross-sectional data, Venn diagrams were used to describe chronic disease case counts (i.e., the numerator data for provincial prevalence estimates) for each algorithm. These diagrams compare the number and percent of disease cases obtained from each of the three sources of administrative data: hospital separations, physician billing claims, and prescription drug records. This information is important for assessing potential biases in chronic disease prevalence estimates if one or more administrative data sources is not available for defining an algorithm. A Venn diagram, as depicted in Figure 2 depicts both the unique and shared number of disease cases from each source. For example, section A of the Venn diagram represents the number of disease cases identified only from hospital data. Similarly, sections B and C represent the number of disease cases identified only from physician and prescription drug data, respectively. Sections D, E, and F represent the number of cases identified from each combination of two data sources, and section G represents the number of cases identified in all three administrative data sources.

Figure 2: Venn Diagram for Describing Counts of Chronic Disease Cases



Source: Manitoba Centre for Health Policy, 2006

For the longitudinal provincial data, graphical techniques were used to depict the change in prevalence estimates over successive time periods. The data are presented separately for algorithms based on one, two, and three years of administrative data.

Regression analyses were used to test for differences in provincial prevalence estimates for the chronic disease algorithms. For all regression models, the dependent variable was the number of chronic disease cases identified for an algorithm. The population count (i.e., denominator) was used as an offset variable in the model. The explanatory variables were algorithm, age, sex, region of residence, and income quintile. Age, sex, and region of residence were defined from the population registry. Income quintiles were defined using the methodology described by Roos and Mustard (1997), which links Statistics Canada Census data to population registry data. Income quintile is an area-level measure based on total household income for dissemination areas (DAs), the smallest geographic unit for which Census data are reported. Each quintile (i.e., Q1 to Q5) represents approximately 20 percent of the total Manitoba population.

The case counts and population counts for each algorithm were stratified by age group (i.e., 10-year groupings were used in most models), sex, region of

residence (i.e., rural north RHAs, rural south RHAs, Winnipeg RHA), and income quintile (i.e., Q1 [lowest] to Q5 [highest]). The quintiles were defined separately for rural and urban residents. In the case of the longitudinal analyses, the data were also stratified by year or time period. Age, sex, region of residence, income quintile, and algorithm were all treated as categorical variables in the regression models.

Generalized linear models (McCulloch and Searle, 2001) with a negative binomial distribution were used for all regression analyses because preliminary descriptive results revealed that disease prevalence estimates, which might be expected to follow a Poisson distribution (particularly for rare diseases) were overdispersed (i.e., the variance was greater than the mean). For all models specified with a negative binomial distribution, the ratio of the model deviance to its degrees of freedom, which should equal 1 for a well-fitting model, was always close to this value. For the longitudinal data, generalized estimating equations (GEEs) were adopted to account for the correlation among the estimates across successive time periods (Fitzmaurice et al., 2004).

For the cross-sectional prevalence estimates, two models were fit to the data. In the first model, the full model, the explanatory variables were the main effects of algorithm, age, sex, region of residence, and income quintile, and the two-way interactions of algorithm x age, algorithm x sex, algorithm x region, and algorithm x quintile. In the second model, the reduced model, the explanatory variables were the main effects of algorithm, age, sex, region, and quintile. A likelihood ratio test (LRT), which is computed as -2 times the natural logarithm of the difference in the likelihood for the full and reduced models, was used to test the null hypothesis that there was no difference in model fit between the full and reduced models (McCulloch and Searle, 2001). A statistically significant LRT indicates that the two-way interactions account for a significant proportion of total model variation. If one or more interaction terms is significant, this means that the relative rate (RR) of prevalence for different algorithms is not constant across one or more of the sociodemographic variables of age, sex, region of residence, and income quintile in the population.

For the longitudinal prevalence estimates three models were fit to the data. In the full model, the explanatory variables were the main effects of algorithm, age, sex, region of residence, income quintile and year/time, the two-way interactions of algorithm x age, algorithm x sex, algorithm x region, algorithm x quintile, algorithm x time, and the three-way interactions of algorithm x age x time, algorithm x sex x time, algorithm x region x time, algorithm x quintile x time. In the second model, the partial model, the explanatory variables were the main effects of algorithm, age, sex, region, quintile, and time and the two-way interaction of algorithm x year/time. In

the third model, the reduced model, the explanatory variables were the main effects of algorithm, age, sex, region, quintile, and year/time. LRTs could only be performed for the model comparisons when independence of observations was assumed because GEEs are based on quasi-maximum likelihood estimation (Fitzmaurice et al., 2004). The LRT for the partial and reduced models is used to test the null hypothesis that there is no difference in model fit between the partial and reduced models, or, correspondingly, that the algorithm x year/time does not account for a significant proportion of total model variation. A statistically significant LRT indicates that the RR for different algorithms is not constant over time. The LRT for the full and reduced models is used to test the null hypothesis that there is no difference in model fit between the full and reduced models, or, correspondingly, that the two-way and three-way interactions do not account for a significant proportion of total model variation. If one or more of the interaction terms is significant, this indicates that the RR for different algorithms over time is not constant across one or more of the sociodemographic variables of age, sex, region of residence, and income quintile.

Tests of individual model effects were conducted if a LRT was statistically significant. In addition, the regression coefficients were exponentiated to produce the RR estimates. The RR describes the prevalence rate for one level of a categorical variable relative to another level of a categorical variable, which is called the reference category. The reference categories for age, sex, region of residence, and income quintile were the oldest age group, females, Winnipeg, and the wealthiest income quintile, respectively. For the algorithm variable, the reference category was different for each chronic disease, but was typically the algorithm based on the fewest number of years of data and/or data sources. For the time variable which was a continuous variable in all models, the exponentiated regression coefficient describes the RR of increase/decrease in prevalence across individuals years or time periods.

CHAPTER 3: ARTHRITIS

3.1 Introduction and Review of Literature

This chapter focusses on administrative data algorithms for all forms of arthritis, as well as algorithms for rheumatoid arthritis (RA) and osteoarthritis (OA). Arthritis is a growing public health concern because it is associated with high health care utilization and costs.

This chapter focusses on administrative data algorithms for all forms of arthritis, as well as algorithms for rheumatoid arthritis (RA) and osteoarthritis (OA). Arthritis is a growing public health concern because it is associated with high health care utilization and costs. Hootman et al. (2002) found, for example, that arthritis and other rheumatic conditions account for as many physician visits as cardiovascular disease, and more physician visits than respiratory diseases like asthma, COPD, and chronic bronchitis.

National and international estimates of arthritis prevalence show substantial variability. For example, a recent U.S. study estimated that 21% of adults 18 years of age and older had doctor-diagnosed arthritis, and another 11% had possible arthritis (Bolen et al., 2005). Badley and DesMeules (2003) estimated the prevalence of OA in Canada at about 10% and the prevalence of RA at approximately 1%. Wang et al. (2000) estimated the prevalence of arthritis and rheumatism in Canada using 1994 NPHS data at 14.2% for the population 20 years of age and older. Badley and Wang (1998) used Canadian survey data from the early 1990s to project an estimated increase in the prevalence of arthritis from 10.7% to 15.7% through to 2031.

Table A.1 in Appendix A summarizes seven studies that used administrative data to identify cases with all forms of arthritis, RA or OA. The studies primarily relied on physician billing claims to identify cases of arthritis, although the Powell et al. (2003), Rector et al. (2004), and Singh et al. (2004) studies also identified cases of RA using prescriptions for disease-modifying anti-rheumatic drugs (DMARDS). A wide variety of other administrative sources were used to identify disease cases, including hospital separations, emergency department records, and laboratory results. The study by Powell et al. (2003) suggests that more than a single year of data is required to obtain a valid prevalence estimate from administrative data. results.

Sensitivity as high as 90% and specificity greater than 95% were observed in the published studies. However, in the Losina et al. (2003) study the maximum observed sensitivity was only 65% for RA. As well, Fowles et al. (1998) observed sensitivity of 68% and specificity of 88% when they used administrative data to identify individuals with joint problems. Thus, while some studies suggest that administrative data can validly be used to identify individuals with arthritis, there is a lack of consistency in this observation.

The majority of the studies retrieved through the literature review used ICD-9-CM codes 714 to identify RA cases and 715 to define OA cases,

although a few used more specific sets of four- or five-digit codes for each of these forms of arthritis. For all forms of arthritis, a wide variety of ICD-9-CM codes were adopted in previous research and procedure codes were also used in some studies (i.e., Katz et al., 1997).

3.2 Description of Arthritis Algorithms

In this study, a single ICD-9-CM code, 714, was used to identify RA cases and a single code, 715, was used to identify OA cases. For all forms of arthritis, these two diagnostic codes were selected, in addition to several diagnostic codes for connective tissue disorders, gout, anklyosing spondylitis, and other forms of arthritis. The codes selected for all forms of arthritis are the same as those adopted in a national report on arthritis prepared by Health Canada (Badley and DesMeules, 2003). As noted previously, the selected ICD-9-CM codes are listed in Table 1 in Chapter 2.

Pharmacological treatment of RA is primarily by: (1) DMARDS, which include xenobiotic agents and biologic agents, (2) anti-inflammatory agents including glucocorticoids and non-steroidal anti-inflammatory agents (NSAIDs), and (3) analgesics such as acetaminophen, opiates, and topical agents. Pharmacological treatment of OA is primarily by anti-inflammatory agents and analgesics. The following process was used to select the prescription drugs for inclusion in this research. A set of relevant ATC codes was identified through the literature search and consultations with experts (i.e., rheumatologist and pharmaco-epidemiological researchers). Then, all of the DINs associated with these ATC codes were identified from the MCHP Master Formulary. Since the list of drugs was very extensive, it was reviewed again with the experts to ensure that no relevant drugs had been missed. Appendix B contains a complete list of the fourth- or fifth-level ATC codes that were selected for the research, and the generic drug names associated with these codes.

The 16 algorithms that were investigated for all forms of arthritis, RA, and OA are listed in Table 5. These algorithms are based on up to five years of administrative data because Powell et al. (2003) observed that multiple years of data are required to obtain a valid algorithm. Algorithm #1 was based on one or more physician claims with an arthritis diagnostic code in one year of data, while algorithm #2 was based on at least two physician claims in one year of data. For algorithm #4, individuals were classified as disease cases if they had one or more hospital separations with an arthritis diagnostic code, or two or more physician billing claims with an arthritis diagnostic code, or if they had a single physician billing claim with an arthritis diagnostic code and at least two prescription drug records with an arthritis medication code in one year of administrative data. Two-, three-, and five-year algorithms were similarly defined. For example, algorithm #5 required at least one

physician claim with an arthritis diagnosis code in two years for an individual to be classified as a disease case.

Table 5: Arthritis algorithms selected for validation

# Years	Algorithm	Hospital Separations or	Physician Claims or	Physician Claims and Prescription Drug Records
1	1		1 or more	
	2		2 or more	
	3	1 or more	2 or more	
	4	1 or more	2 or more	1 and 2 or more
2	5		1 or more	
	6		2 or more	
	7	1 or more	2 or more	
	8	1 or more	2 or more	1 and 2 or more
3	9		1 or more	
	10		2 or more	
	11	1 or more	2 or more	
	12	1 or more	2 or more	1 and 2 or more
5	13		1 or more	
	14		2 or more	
	15	1 or more	2 or more	
	16	1 or more	2 or more	1 and 2 or more

Source: Manitoba Centre for Health Policy, 2006

None of the arthritis algorithms relied solely on the prescription drug data for the identification of arthritis cases. That is, at least one contact in hospital separations or physician billing claims had to occur in combination with two or more prescription drug records for an individual to be classified as a disease case. This requirement was implemented because there are no unique marker drugs for arthritis. Many of the prescription drugs, such as NSAIDs are used to treat several other diseases. A possible exception, as noted above, is DMARDs.

An arthritis algorithm that relied only on prescription drug data for identification of arthritis cases was expected to have very low specificity. To verify this, we examined several algorithms for all forms of arthritis, RA, and OA which were based on one or more contacts in prescription drug records without requiring a diagnosis in hospital separations or physician billing claims. A summary of the validation results for these algorithms is reported in Appendix C. It is important to note that when we conducted the validation analyses, we limited our attention to DMARDs to identify cases of RA,

and to NSAIDs and analgesics to identify cases of OA. However, we included all of the prescription drugs listed in Appendix B to identify cases of all forms of arthritis.

The summary shows that for all forms of arthritis, specificity was as low as 36.8% for an algorithm based on five years of data and one or more contacts in hospital separations, or physician billing claims, or prescription drug records. Specificity was high for RA, but sensitivity was no higher for these supplemental algorithms than for the other algorithms we investigated. For OA, specificity was as low as 49.7% using five years of data. Moreover, prevalence estimates for the entire Manitoba population were substantially higher than expected for all forms of arthritis, RA, and OA, when an individual did not require an arthritis diagnostic code in addition to a medication code in order to be classified as a disease case. Accordingly, we did not include these algorithms in subsequent phases of the research.

3.3 Validation Results

Validation Indices

Table 6 contains the point estimates for the six validation indices for the 16 algorithms for all forms of arthritis. The 95% CIs for each of these estimates are reported in Appendix Table D.1. In this table, and in the tables contained in the remainder of the report, the abbreviation H is used to denote hospital separations, P is used to denote physician billing claims, and Rx is used to denote prescription drug records.

There was fair agreement between administrative and survey data for all forms of arthritis, with values of κ ranging from 0.24 to 0.37. The highest value of κ was for the algorithm based on one or more hospital separations or two or more physician billing claims, or one physician billing claim and two or more prescription drug records in five years.

Using one year of data, the algorithm based on a single physician claim (i.e., algorithm #1) had the highest sensitivity (43.4%); the corresponding algorithm based on two years of data had a sensitivity of 60.0%, and the corresponding algorithm based on three years of data had a sensitivity of 69.0%. Thus, the largest improvement in sensitivity was obtained when moving from the one-year to the two-year algorithms. However, the highest overall sensitivity was observed for algorithm #13, which was based on one or more physician billing claims in five years of data.

The highest specificity was observed for algorithm #2, based on two or more physician claims in one year of administrative data (93.8%). Specificity estimates for algorithms #3 and #4 were almost the same (93.6% and 91.9%, respectively). There was a decrease in specificity of approximately 5 percent-

age points between algorithm #2 and the corresponding algorithm based on two years of data (88.4%).

For all forms of arthritis, Youden's index ranged from 0.19 to 0.41; the latter value was obtained for the five-year algorithm based on a combination of hospital, physician, and prescription drug data (i.e., algorithm #16). However, several other two-year and three-year algorithms produced almost equivalent estimates of Youden's index, including algorithm #15, which was based on a simpler definition of one or more hospital separations or two or more physician billing claims in five years.

The PPV of an arthritis diagnosis was highest for the one-year algorithm based on a combination of all three data sources (57.4 %). There was little variation in the PPV for the one-year and two-year algorithms; the lowest value was 44.1%, which was for the two-year algorithm based on one or more physician billing claims. NPV also showed little variation, with values ranging from 79.8% to 89.4%; it was highest for the algorithm based on one or more physician billing claims in five years.

Table 6: Estimates of agreement, sensitivity, specificity, and predictive values for arthritis algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.28	43.4	84.3	0.28	46.6	82.4
	2	2+ P	0.24	25.3	93.8	0.19	56.4	79.8
	3	1+ H or 2+ P	0.25	26.3	93.6	0.20	56.4	80.0
	4	1+ H or 2+ P or (1 P & 2+ Rx)	0.30	34.5	91.9	0.26	57.4	81.6
2	5	1+ P	0.32	60.0	75.9	0.36	44.1	85.7
	6	2+ P	0.32	41.7	88.4	0.30	53.2	82.7
	7	1+ H or 2+ P	0.33	42.6	88.1	0.31	53.2	82.9
	8	1+ H or 2+ P or (1 P & 2+ Rx)	0.37	51.7	84.9	0.37	52.0	84.7
3	9	1+ P	0.31	69.0	68.7	0.38	41.1	87.5
	10	2+ P	0.34	50.9	83.7	0.35	49.7	84.3
	11	1+ H or 2+ P	0.35	51.8	83.4	0.35	49.7	84.5
	12	1+ H or 2+ P or (1 P & 2+ Rx)	0.36	60.3	78.9	0.39	47.6	86.3
5	13	1+ P	0.27	78.1	58.6	0.37	37.4	89.4
	14	2+ P	0.35	63.1	76.2	0.39	45.7	86.7
	15	1+ H or 2+ P	0.35	63.7	75.9	0.40	45.6	86.8
	16	1+ H or 2+ P or (1 P & 2+ Rx)	0.34	71.1	70.1	0.41	42.9	88.4

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

For RA, there was poor agreement between administrative and survey data; values of κ never exceeded 0.20 (see Table 7). The highest value of κ (0.17) was for the five-year algorithm based on one or more physician billing claims and also for the five-year algorithm based on one or more hospital separations, or two or more physician billing claims, or one physician billing claim and two or more prescription drug records.

Sensitivity ranged from 5.0% to 11.3%. The highest sensitivity was observed for the five-year algorithm based on one or more physician billing claims. However, sensitivity was almost equivalent for the corresponding three-year algorithm (10.7%). Specificity was near 100% for all RA algorithms. Youden's index ranged from 0.05 to 0.11. It was highest for two of the five-year algorithms: (a) one or more physician billing claims, and (b) one or more hospital separations or two or more physician billing claims, or one physician billing claim and two or more prescription drug records.

The PPV of an RA diagnosis ranged from 55.9% to 80.6% and was, as expected, highest for the algorithms based on a single year of data. There was almost no variation in NPV; it was approximately 92% for all RA algorithms.

Table 7: Estimates of agreement, sensitivity, specificity, and predictive values for rheumatoid arthritis algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.12	7.4	99.8	0.07	73.9	92.3
	2	2+ P	0.08	5.0	99.9	0.05	76.7	92.2
	3	1+ H or 2+ P	0.09	5.4	99.9	0.05	78.1	92.2
	4	1+ H or 2+ P or (1 P & 2+ Rx)	0.12	6.3	99.9	0.07	80.6	92.3
2	5	1+ P	0.14	8.9	99.6	0.09	65.1	92.4
	6	2+ P	0.11	6.5	99.8	0.06	71.4	92.3
	7	1+ H or 2+ P	0.11	7.0	99.8	0.07	72.7	92.3
	8	1+ H or 2+ P or (1 P & 2+ Rx)	0.14	7.6	99.7	0.08	72.9	92.3
3	9	1+ P	0.16	10.7	99.4	0.10	62.8	92.6
	10	2+ P	0.13	7.8	99.7	0.08	70.6	92.4
	11	1+ H or 2+ P	0.14	8.5	99.7	0.08	72.2	92.4
	12	1+ H or 2+ P or (1 P & 2+ Rx)	0.14	8.9	99.7	0.10	70.7	92.4
5	13	1+ P	0.17	11.3	99.2	0.11	55.9	92.6
	14	2+ P	0.13	8.3	99.7	0.08	69.1	92.4
	15	1+ H or 2+ P	0.14	8.9	99.7	0.09	70.7	92.4
	16	1+ H or 2+ P or (1 P & 2+ Rx)	0.17	9.4	99.6	0.11	68.3	92.5

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

For OA, there was poor to fair agreement between survey and administrative data with values of κ ranging from 0.16 to 0.32 (see Table 8). This statistic was highest for the algorithm based on one or more physician billing claims in five years. Sensitivity was highest for this same algorithm (49.9%); it was more than double the sensitivity estimate for the corresponding algorithm based on a single year of data. However, the largest improvement in the estimate of sensitivity was observed when moving from the one-year algorithm based on one or more physician billing claims, to the two-year algorithm based on one or more physician billing claims (difference = 11.6%).

Specificity was greater than 90% for all of the algorithms with the exception of the five-year algorithm based on one or more physician billing claims (88.7%). Youden's index was highest for the five-year algorithm based on one or more physician billing claims (0.39).

The PPV of an OA diagnosis ranged from 34.8% to 49.3%, and was highest for the one-year algorithm based on two or more physician billing claims. NPV showed little variation, from 90.4% to 93.6%.

Table 8: Estimates of agreement, sensitivity, specificity, and predictive values for osteoarthritis algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.24	23.5	95.9	0.19	40.8	91.2
	2	2+ P	0.16	12.3	98.5	0.11	49.3	90.3
	3	1+ H or 2+ P	0.17	13.1	98.3	0.11	48.5	90.4
	4	1+ H or 2+ P or (1 P & 2+ Rx)	0.22	19.0	97.2	0.16	45.2	90.9
2	5	1+ P	0.30	35.1	93.3	0.28	38.9	92.3
	6	2+ P	0.23	19.8	97.2	0.17	45.8	90.9
	7	1+ H or 2+ P	0.23	20.8	97.0	0.18	45.3	91.0
	8	1+ H or 2+ P or (1 P & 2+ Rx)	0.28	29.8	94.7	0.25	40.5	91.8
3	9	1+ P	0.31	41.4	91.2	0.33	36.3	92.8
	10	2+ P	0.25	24.5	95.9	0.21	41.8	91.3
	11	1+ H or 2+ P	0.25	25.3	95.7	0.21	41.3	91.4
	12	1+ H or 2+ P or (1 P & 2+ Rx)	0.25	34.9	92.9	0.26	37.2	92.2
5	13	1+ P	0.32	49.9	88.7	0.39	34.8	93.6
	14	2+ P	0.29	31.6	94.3	0.26	40.2	92.0
	15	1+ H or 2+ P	0.29	32.8	94.0	0.27	39.8	92.1
	16	1+ H or 2+ P or (1 P & 2+ Rx)	0.31	43.1	90.7	0.34	35.8	93.0

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

Agreement Between Survey And Administrative Data

Logistic regression analysis was used to test the sociodemographic variables associated with agreement between survey and administrative data. For all forms of arthritis, RA, and OA the models contained the main effects of age, sex, region of residence, and income adequacy quintile. Two-way interactions were tested but were not statistically significant and were therefore excluded from the final models. The Hosmer-Lemeshow test indicated that all of the main-effect models provided a good fit to the data.

For all forms of arthritis, the logistic regression model was applied to the data for algorithm #8, for RA it was applied to the data for algorithm #13, and for OA it was also applied to the data for algorithm #13. For all forms of arthritis, the following variables were statistically significant predictors of agreement between the two data sources: age (χ^2 = 204.8, degrees of freedom [df] = 6, p < .0001), sex (χ^2 = 8.4, df = 1, p = .0038), and income adequacy quintiles (χ^2 = 14.0, df = 5, p = .0156). For RA, the following variables were statistically significant: age (χ^2 = 158.7, df = 6, p < .0001), region of residence (χ^2 = 11.4, df = 2, p = .0033), and income adequacy quintiles (χ^2 = 17.0, df = 5, p = .0045). For OA, the following variables were statistically significant: age (χ^2 = 371.1, df = 6, p < .0001), sex (χ^2 = 13.5, df = 1, p = .0002), region of residence (χ^2 = 7.5, df = 2, p = .0238), and income adequacy quintiles (χ^2 = 12.9, df = 5, p = .0243). Table 9 contains the ORs and 95% CIs for the explanatory variables in each of the three models.

Table 9: Odds Ratio (OR) estimates and 95% CIs for predictors of agreement between administrative and survey data for arthritis, rheumatoid arthritis, and osteoarthritis

	Arthritis		RA		OA	
Predictors	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Age						
19 – 34 years	4.1	(2.9, 5.8)	23.1	(14.5, 36.9)	12.7	(6.9, 23.1)
35 – 44 years	2.4	(1.7, 3.4)	7.7	(5.2, 11.4)	4.1	(2.5, 6.7)
45 – 54 years	1.5	(1.1, 2.1)	3.3	(2.3, 4.7)	2.3	(1.4, 3.7)
55 – 64 years	1.2	(0.9, 1.7)	1.9	(1.3, 2.7)	1.2	(0.8, 2.0)
65 – 74 years	1.1	(0.8, 1.6)	1.3	(0.9, 1.9)	1.0	(0.7, 1.6)
75 – 84 years	1.0	(0.7, 1.4)	1.3	(0.9, 1.8)	1.0	(0.7, 1.6)
85+ years	Ref	--	--	--	--	--
Sex						
Males	1.2	(1.1, 1.4)	1.4	(1.2, 1.6)	1.0	(0.8, 1.2)
Females	Ref	--	--	--	--	--
Region of Residence						
North Rural RHAs	1.1	(0.8, 1.4)	0.7	(0.5, 0.9)	0.8	(0.6, 1.2)
South Rural RHAs	1.0	(0.9, 1.2)	1.0	(0.8, 1.2)	1.3	(1.0, 1.7)
Winnipeg RHA	Ref	--	--	--	--	--
Income Quintile						
Lowest	0.6	(0.4, 0.8)	0.4	(0.2, 0.7)	0.5	(0.4, 0.8)
Low Middle	1.1	(0.9, 1.5)	0.6	(0.4, 0.9)	0.9	(0.7, 1.3)
Middle	1.0	(0.8, 1.2)	0.6	(0.4, 0.8)	0.9	(0.7, 0.1)
Upper Middle	1.0	(0.9, 1.3)	0.8	(0.6, 1.1)	1.0	(0.8, 1.3)
Not Stated	0.8	(0.6, 1.1)	0.7	(0.5, 1.2)	0.9	(0.6, 1.2)
Highest	Ref	--	--	--	--	--

Source: Manitoba Centre for Health Policy, 2006

For all models, the results indicate that the odds of agreement between the two data sources were higher for individuals in younger age groups than in the 85+ year age group. They were also higher for males than for females. The odds of agreement were lower for residents of the Northern RHAs than for residents of Winnipeg RHA for both RA and OA. Finally, for all three chronic diseases, the odds of agreement between the two data sources were lower for individuals in the lower income groups than for individuals in the highest income group.

3.4 Provincial Prevalence Estimates

Cross-Sectional Prevalence Estimates

Prevalence estimates for all 16 algorithms for all forms of arthritis, RA, and OA are summarized in Table 10. As noted previously, estimates for one-year algorithms were calculated using 2002/03 data; estimates for two-year algorithms were calculated using 2001/02–2002/03 data; estimates for three-year algorithms were calculated using 2000/01–2002/03 data; estimates for five-year algorithms were calculated using 1998/99–2002/03 data.

Prevalence estimates varied substantially across the algorithms. For all forms of arthritis, estimates ranged from 9.2% to 47.3%. Prevalence estimates ranged from 0.4% to 1.6% for RA, and from 2.2% to 13.2% for OA. For all forms of arthritis, the prevalence estimate for the most sensitive algorithm (i.e., algorithm #13) was 47.3%. For the two algorithms with the highest overall values of Youden's index—algorithms #15 and #16—the prevalence estimates were 31.5% and 37.4%, respectively. For the algorithm with the highest κ the estimate was 20.3%. For RA, the algorithm with the highest sensitivity and Youden's index (i.e., algorithm #16) resulted in a prevalence estimate of 1.0%. Youden's index was equally high for the algorithm based on one or more physician billing claims in five years of data; this algorithm resulted in a prevalence estimate of 1.6% for the Manitoba population 19 years of age and older. Finally, for OA, the algorithm with the highest κ, sensitivity, and Youden's index (i.e., algorithm #13) resulted in a prevalence estimate of 13.2%.

Table 10: Crude provincial prevalence estimates for arthritis algorithms, 1998/99 – 2002/03

# Years		Algorithm	Arthritis (%)	RA (%)	OA (%)
1	1	1+ P	18.8	0.7	5.0
	2	2+ P	9.2	0.4	2.2
	3	1+ H or 2+ P	9.4	0.5	2.3
	4	1+ H or 2+ P or (1 P & 2+ Rx)	12.0	0.5	3.5
2	5	1+ P	28.8	1.0	7.8
	6	2+ P	16.1	0.6	3.9
	7	1+ H or 2+ P	16.4	0.6	4.1
	8	1+ H or 2+ P or (1 P & 2+ Rx)	20.3	0.7	6.0
3	9	1+ P	36.5	1.2	10.0
	10	2+ P	21.9	0.7	5.3
	11	1+ H or 2+ P	22.3	0.8	5.6
	12	1+ H or 2+ P or (1 P & 2+ Rx)	27.1	0.8	8.1
5	13	1+ P	47.3	1.6	13.2
	14	2+ P	31.1	0.9	7.4
	15	1+ H or 2+ P	31.5	1.0	7.8
	16	1+ H or 2+ P or (1 P & 2+ Rx)	37.4	1.0	11.1

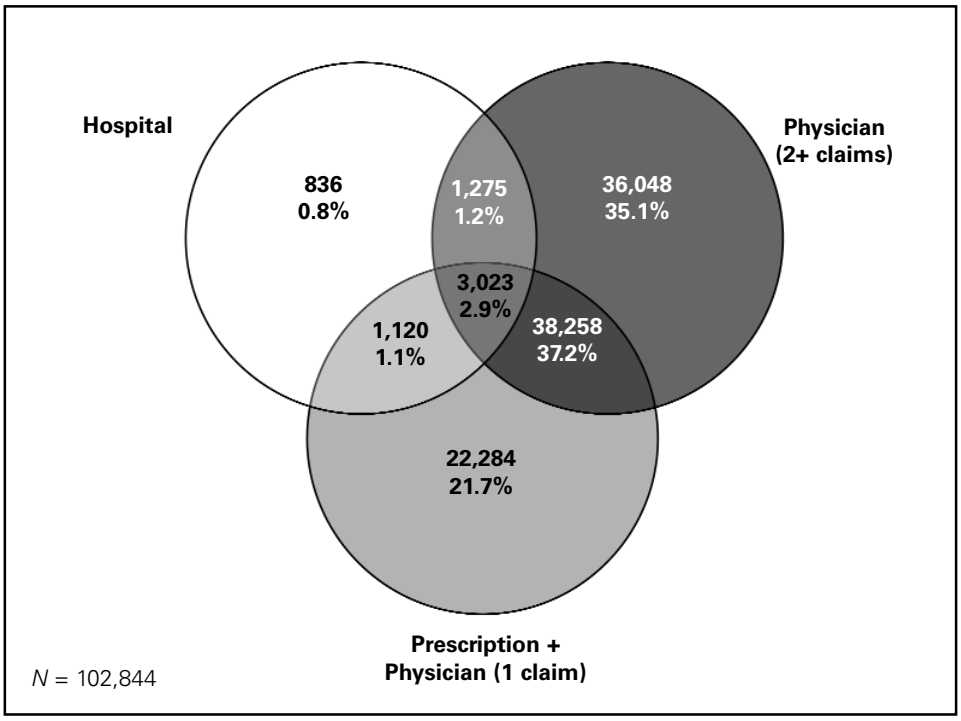
Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 1-year estimates are for 2002/03, 2-year estimates are for 2001/02 – 2002/03, 3-year estimates are for 2000/01 – 2002/03, 5-year estimates are for 1998/99 – 2002/03.

Source: Manitoba Centre for Health Policy, 2006

Venn Diagrams

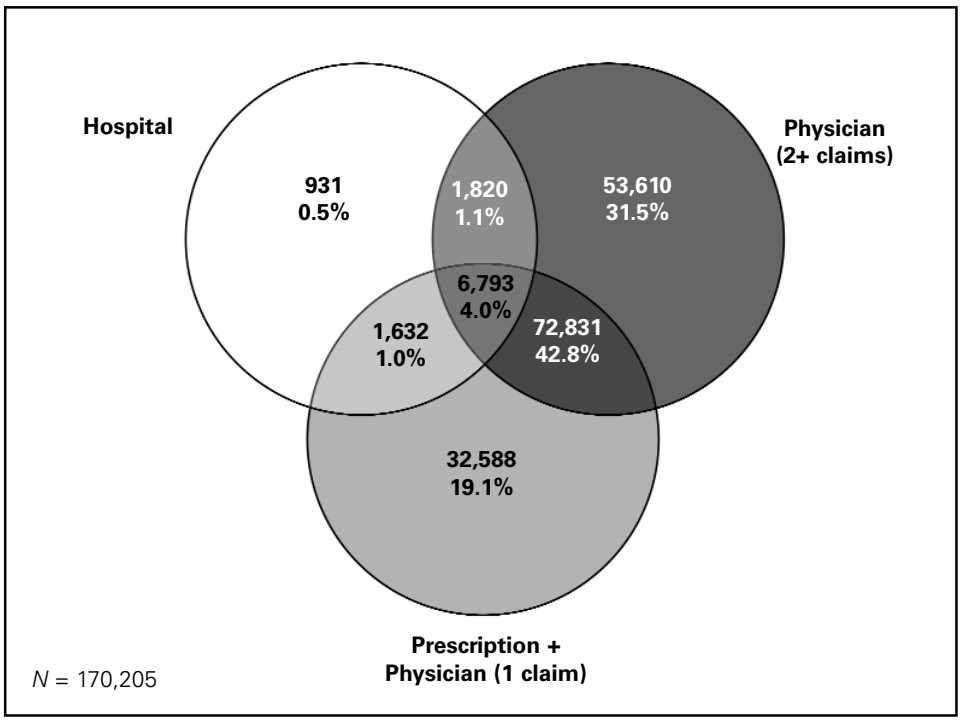
Figures 3 to 6 contain Venn diagrams of the arthritis case counts for a select number of algorithms (i.e., algorithms #4, #8, #12, and #16) for all forms of arthritis. Figure 3 reveals that of the 102,844 individuals who were identified as arthritis cases by applying the one-year algorithm to 2002/03 data, more than one-third of these cases (35.1%) were identified as having two or more physician billing claims with an arthritis diagnostic code. Only 0.8% were identified as having one or more hospital separations with an arthritis diagnostic code. As well, 21.7% of individuals were identified from the combination of one physician claim and two or more prescription drug records in a one-year period. Only a very small percentage of individuals (2.9%) were identified in all three administrative data sources.

Figure 3: Arthritis Algorithm #4: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 1 Year



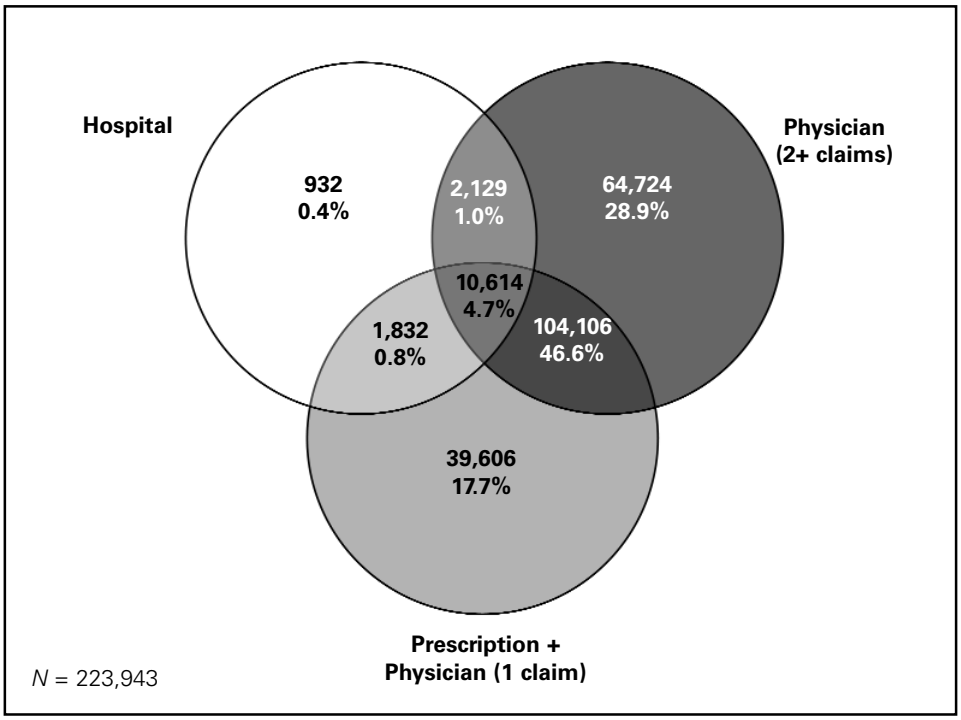
Source: Manitoba Centre for Health Policy, 2006

Figure 4: Arthritis Algorithm #8: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 2 Year



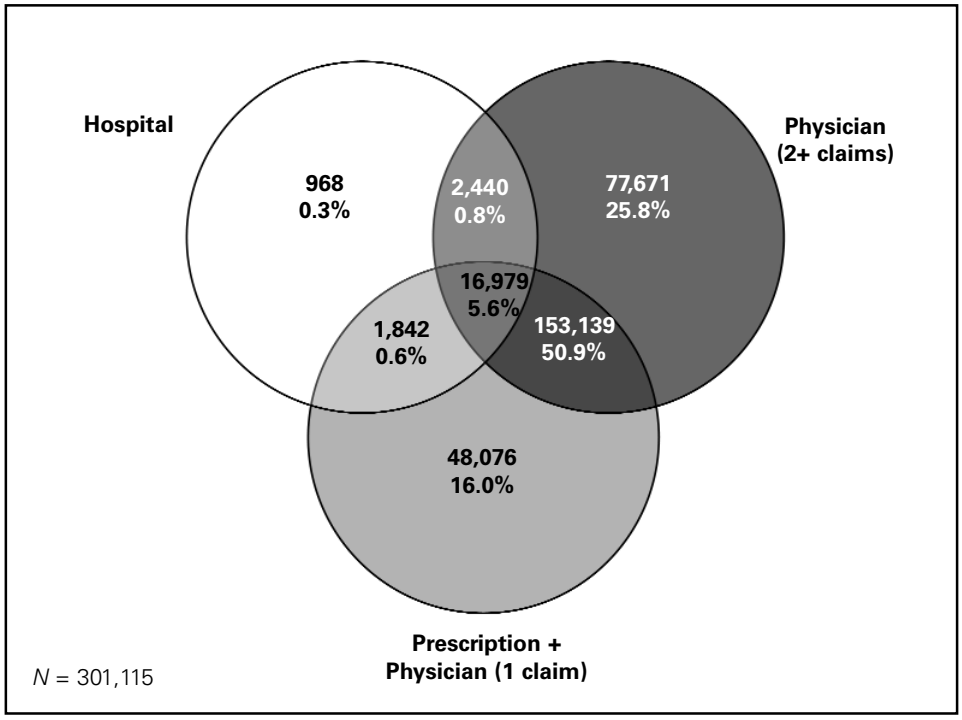
Source: Manitoba Centre for Health Policy, 2006

Figure 5: Arthritis Algorithm #12: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 3 Year



Source: Manitoba Centre for Health Policy, 2006

Figure 6: Arthritis Algorithm #16: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 5 Year

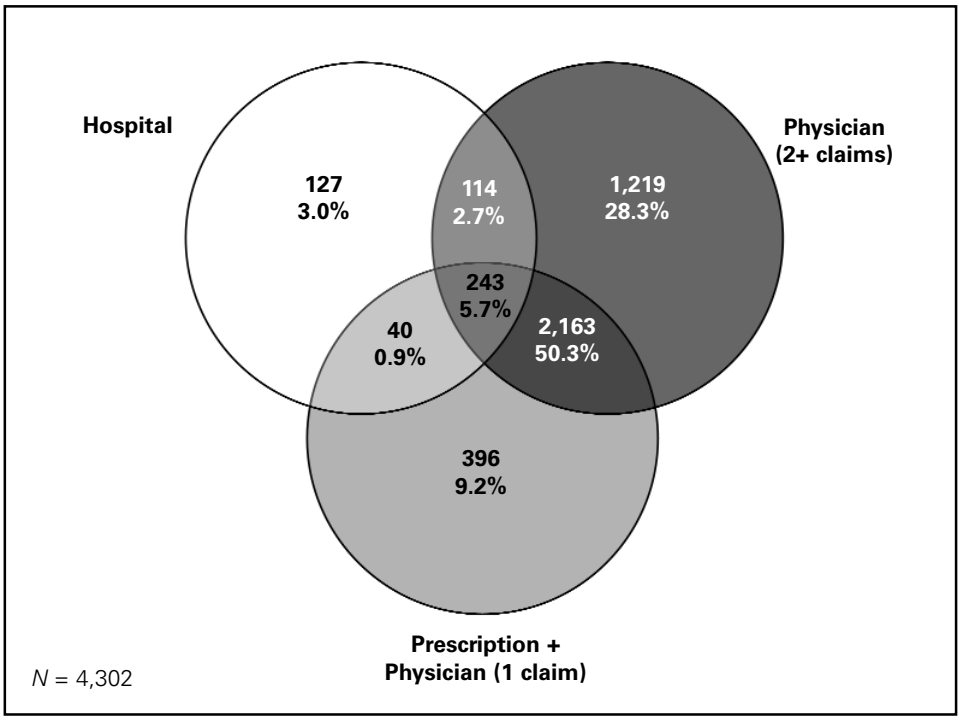


Source: Manitoba Centre for Health Policy, 2006

As revealed in Figures 4, 5 and 6, the percentage of individuals identified as arthritis cases using a combination of the physician and prescription data was lower for the two-year algorithm than for the one-year algorithm (19.1%). This was also true for the three-year algorithm (17.7%) and the five-year algorithm (16.0%).

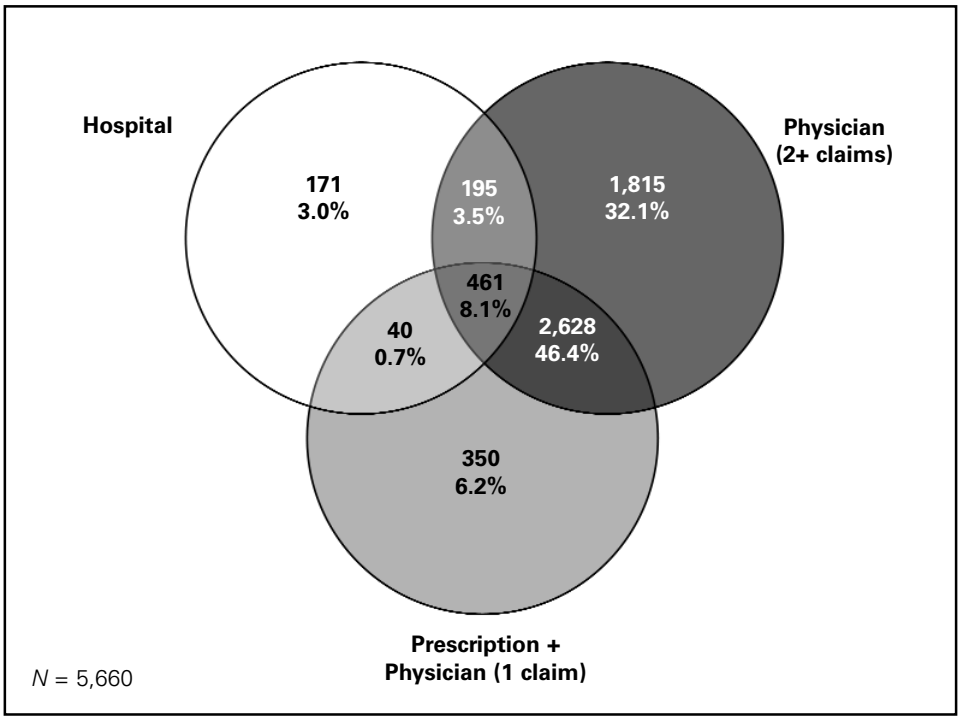
Figures 7 to 10 contain the Venn diagrams for algorithms #4, #8, #12, and #16 for RA. There were 4,302 RA cases identified in one year (i.e., 2002/03) using algorithm #4 (Figure 7). More than one-quarter (28.3%) of these individuals were identified as having two or more physician claims with a RA diagnostic code. Less than 10% had a single physician claim and two or more prescription drug records. The two-year algorithm resulted in the identification of 5,660 RA cases, and only 6.2% had a single physician claim and two or more prescription drug records while 32.1% had two or more physician claims in this time period. The number of individuals identified solely from the physician data increased to 35.1% for the three-year to algorithm and to 37.6% for the five-year algorithm.

Figure 7: Rheumatoid Arthritis Algorithm #4: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 1 Year



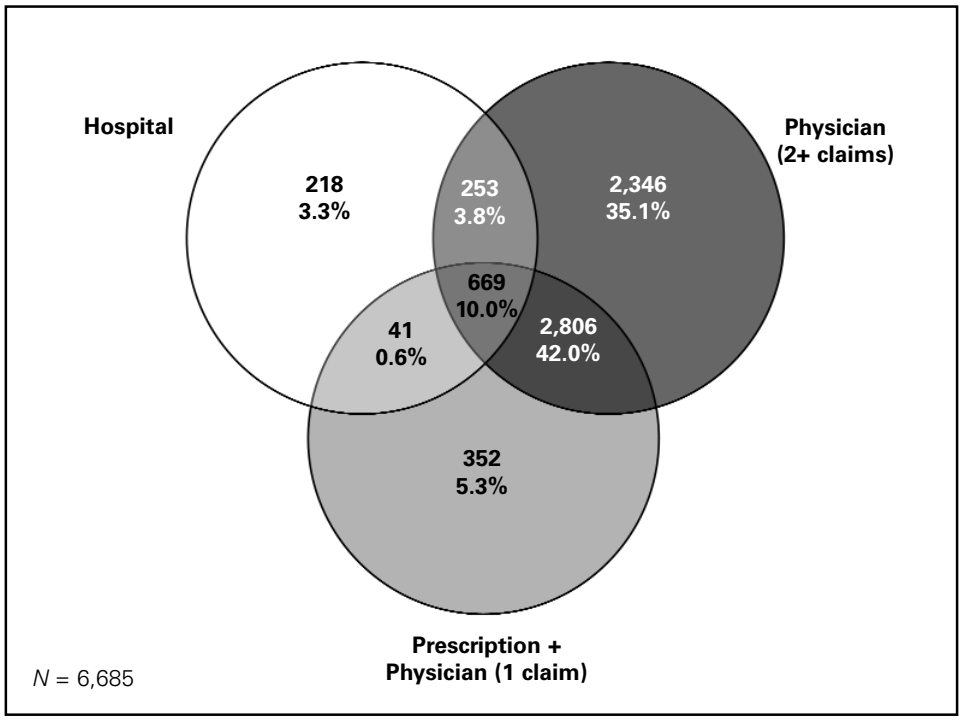
Source: Manitoba Centre for Health Policy, 2006

Figure 8: Rheumatoid Arthritis Algorithm #8: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 2 Years



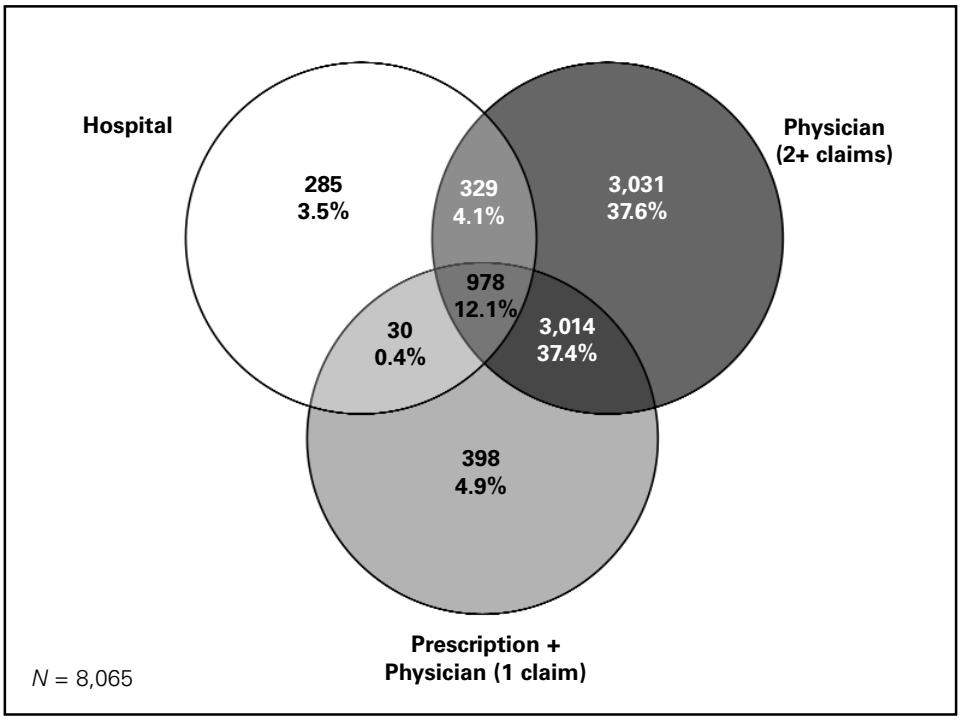
Source: Manitoba Centre for Health Policy, 2006

Figure 9: Rheumatoid Arthritis Algorithm #12: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 3 Years



Source: Manitoba Centre for Health Policy, 2006

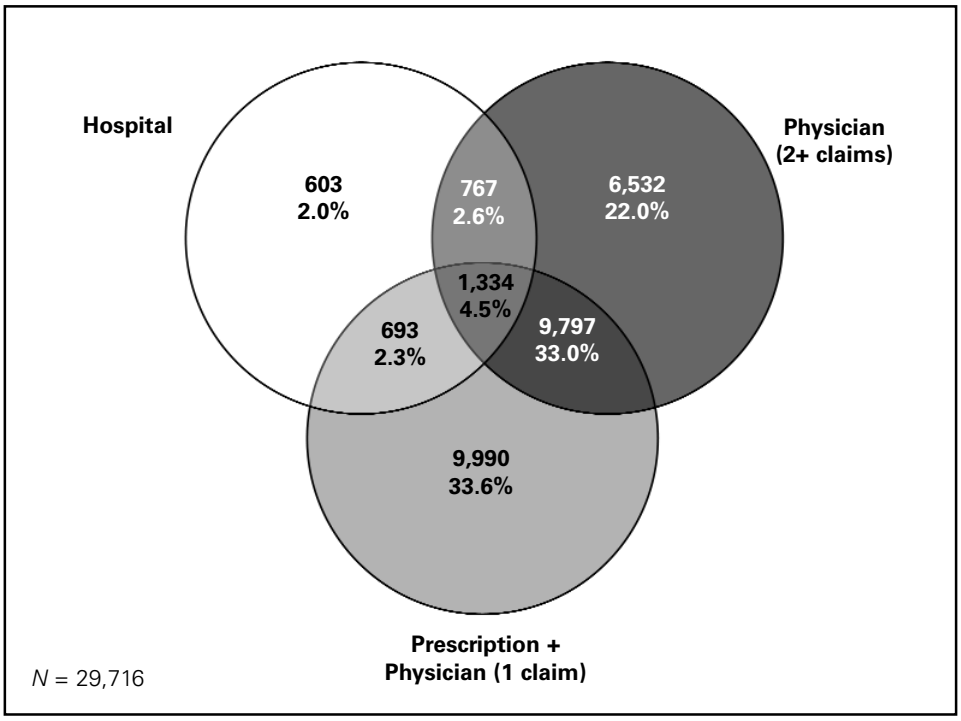
Figure 10: Rheumatoid Arthritis Algorithm #16: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 5 Years



Source: Manitoba Centre for Health Policy, 2006

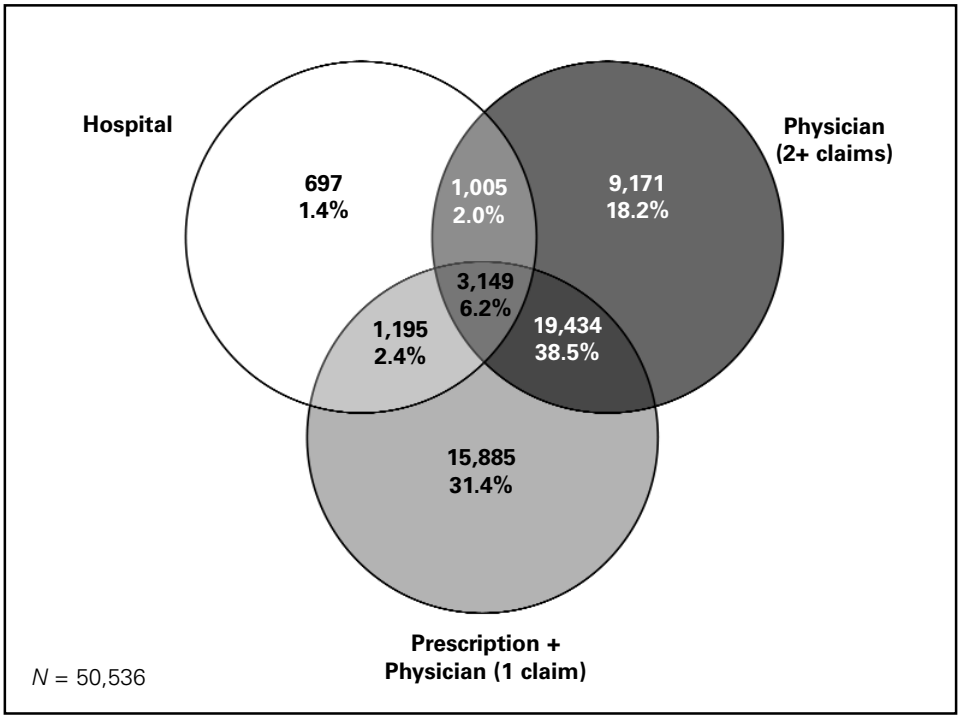
In Figures 11 to 14, Venn diagrams for OA algorithms #4, #8, #12, and #16 applied to the Manitoba data are presented. The one-year algorithm resulted in the identification of close to 30,000 osteoarthritis cases from the three sources of administrative data in 2002/03. One-third (33.6%) of these cases had one physician billing claim with an OA diagnostic code and two or more prescription drug records with a relevant medication code. Less than one-quarter (22.0%) had two or more physician billing claims. Only 2.0% were identified solely from hospital separations. The total number of OA cases increased to more than 50,000 when two years of administrative data were used, and the percentage of cases identified using a combination of the physician and prescription data decreased slightly (31.4%). The results for the three-year algorithm revealed that 30.9% of cases were identified from both physician and prescription drug data, and a relatively small number (15.2%) were identified as having two or more physician billing claims in this time period. For the five-year year algorithm, 12.4% of cases had two or more physician billing claims with an OA diagnostic code.

Figure 11: Osteoarthritis Algorithm #4: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 1 Year



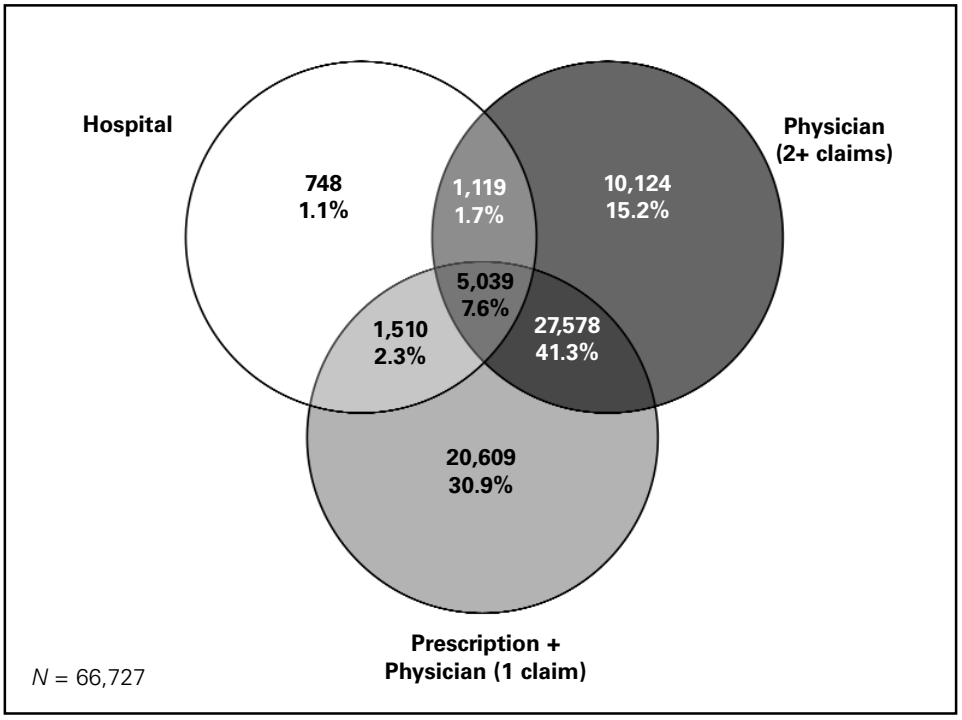
Source: Manitoba Centre for Health Policy, 2006

Figure 12: Osteoarthritis Algorithm #8: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 2 Years



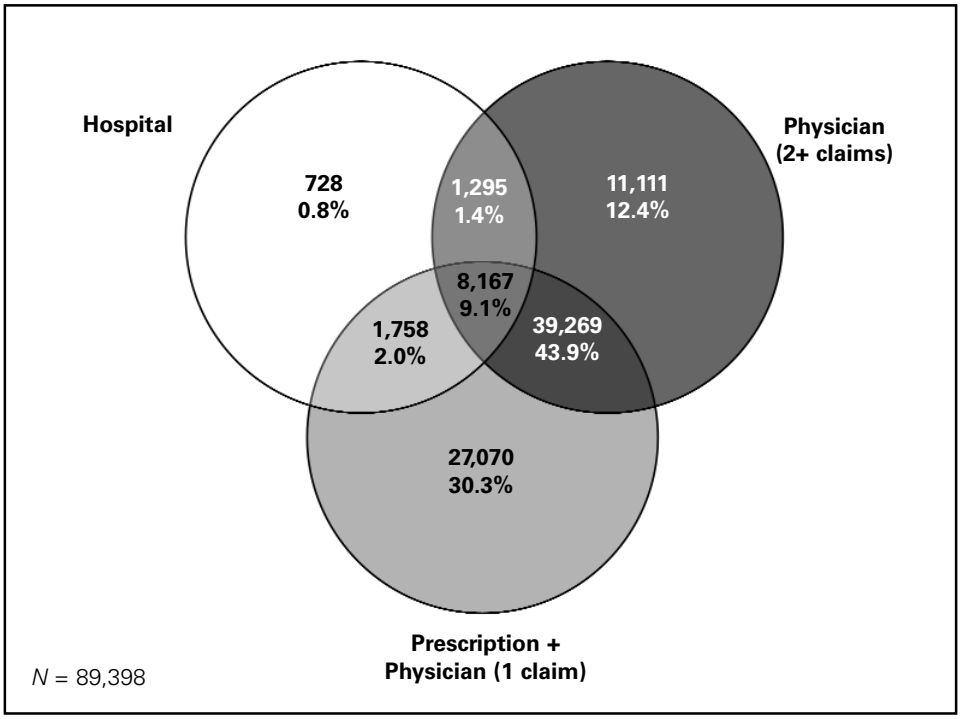
Source: Manitoba Centre for Health Policy, 2006

Figure 13: Osteoarthritis Algorithm #12: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 3 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 14: Osteoarthritis Algorithm #16: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 5 Years



Source: Manitoba Centre for Health Policy, 2006

Regression Analyses for Cross-Sectional Prevalence Estimates

In the regression models for the provincial prevalence estimates, we tested for differences among the algorithms with high κ, sensitivity and/or high values of Youden's index. Prevalence estimates for algorithms #5, #8, #9, #12, #13, and #16 were tested for all forms of arthritis, RA, and OA.

For all forms of arthritis, the LRT for the difference between the full model, which contained the main effects of algorithm, age, sex, region, and quintile in addition to selected two-way interactions, and the reduced model, which contained main effects only, was statistically significant (χ^2 = 726.3, df = 65, p < .0001). Further analysis revealed that all four of the two-way interaction effects specified in the full model were statistically significant: algorithm x age (p < .0001), algorithm x sex (p = .0026), algorithm x quintile (p = .0029), and algorithm x region (p < .0001). These results indicate that for all forms of arthritis, the relative rate (RR) of arthritis for different algorithms varied with the sociodemographic characteristics of the population.

For RA, the LRT for the full and reduced models was not statistically significant (χ^2 = 81.0, df = 65, p =.0870). This result indicates that the RR of RA prevalence for different algorithms did not vary with the sociodemographic characteristics of the population. However, the main effect of algorithm was statistically significant (χ^2 = 951.2; df = 5, p < .0001). Wald tests revealed that all of the algorithms had prevalence estimates that were significantly

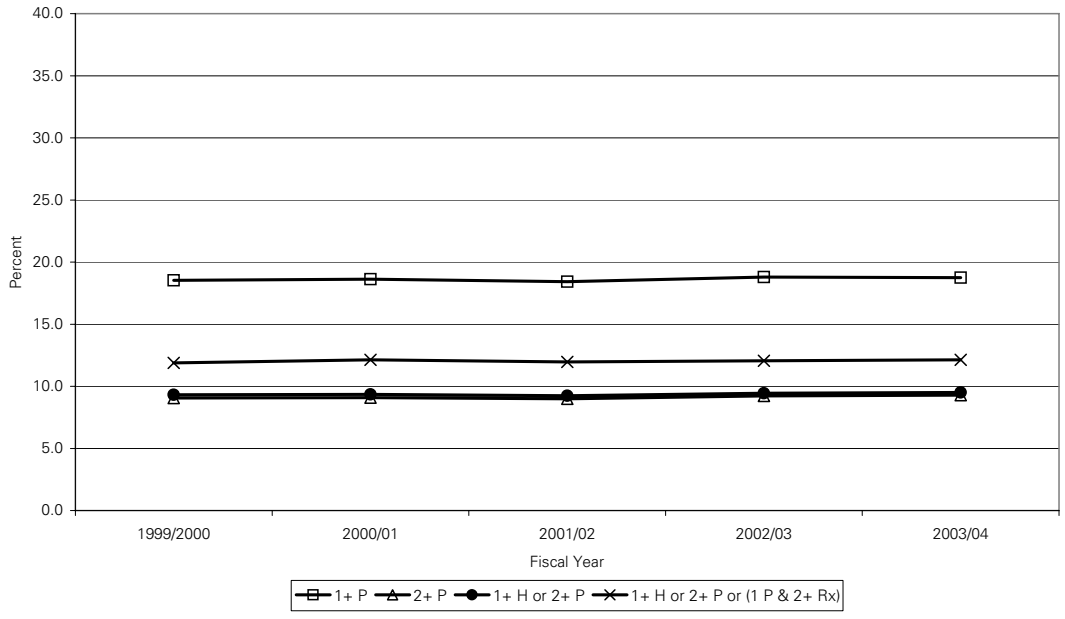
different from the prevalence estimate for algorithm #5 (reference) with the exception of algorithm #12 (1+ H or 2+ P or [1 P & 2+ Rx] in three years).

For OA, the LRT for the full and reduced models was statistically significant (χ^2 = 180.4, df = 65, p <.0001). The algorithm x age (p < .0001) and algorithm x region (p = .0004) interactions were statistically significant. This finding means that the RR of OA prevalence for different algorithms varied with both the age and region of residence of the Manitoba population.

Longitudinal Prevalence Estimates

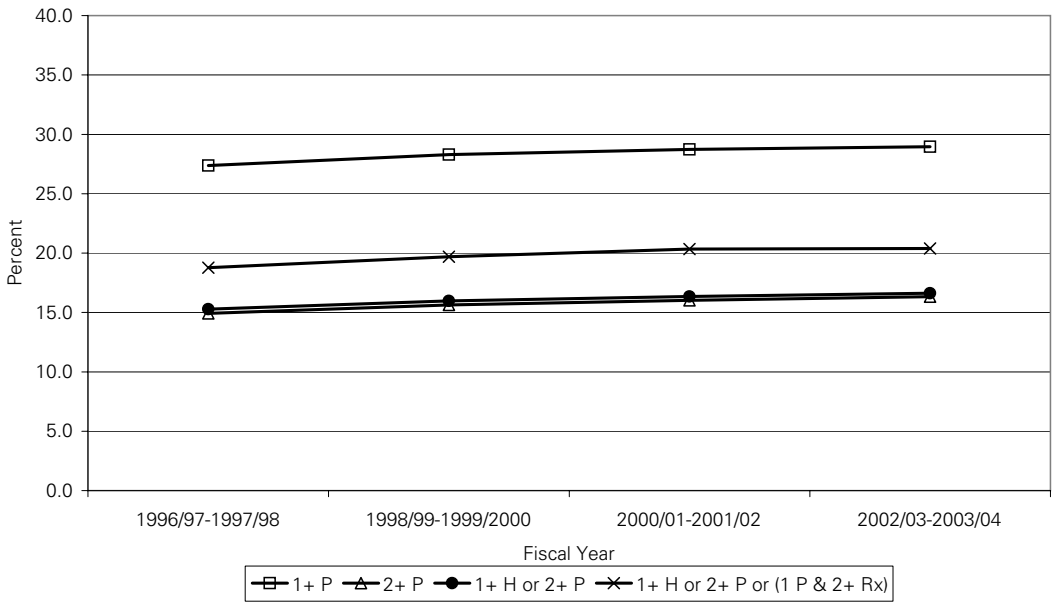
Figures 15, 16, and 17 depict the change in crude prevalence estimates of arthritis for the one-, two-, and three-year, algorithms, respectively. Figure 15 shows relatively little change in the prevalence of arthritis over a five-year period. As well, the trend lines for all four algorithms were roughly parallel, indicating that each provided a similar picture of the change in prevalence of arthritis over time. Both the two-year and three-year algorithms showed a slight increase in prevalence over time, but again the trend lines were approximately parallel, indicating that each provided a similar picture of the rate of change in arthritis prevalence over time.

Figure 15: Provincial Trends in Arthritis Prevalence for One-Year Algorithms, 1999/2000 – 2003/04



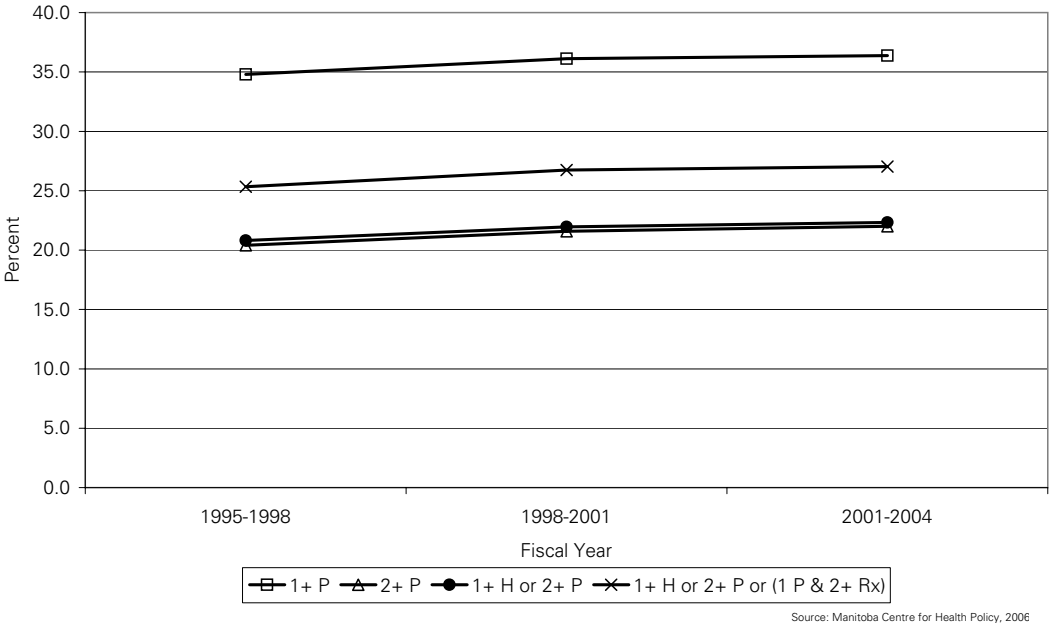
Source: Manitoba Centre for Health Policy, 2006

Figure 16: Provincial Trends in Arthritis Prevalence for Two-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

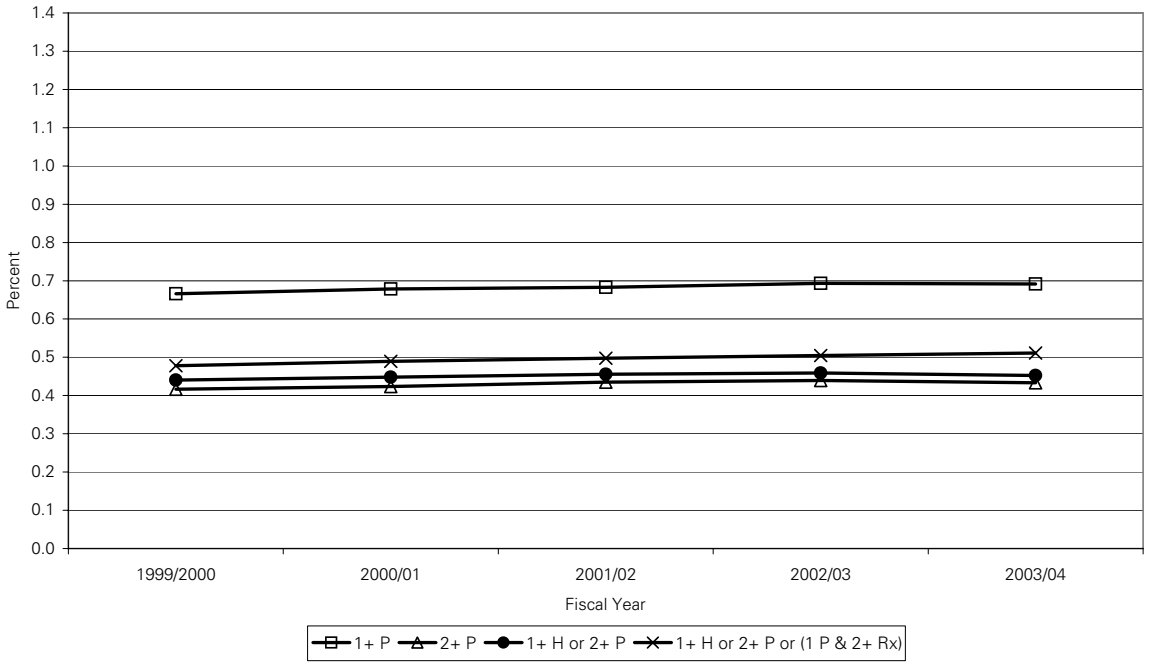
Figure 17: Provincial Trends in Arthritis Prevalence for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

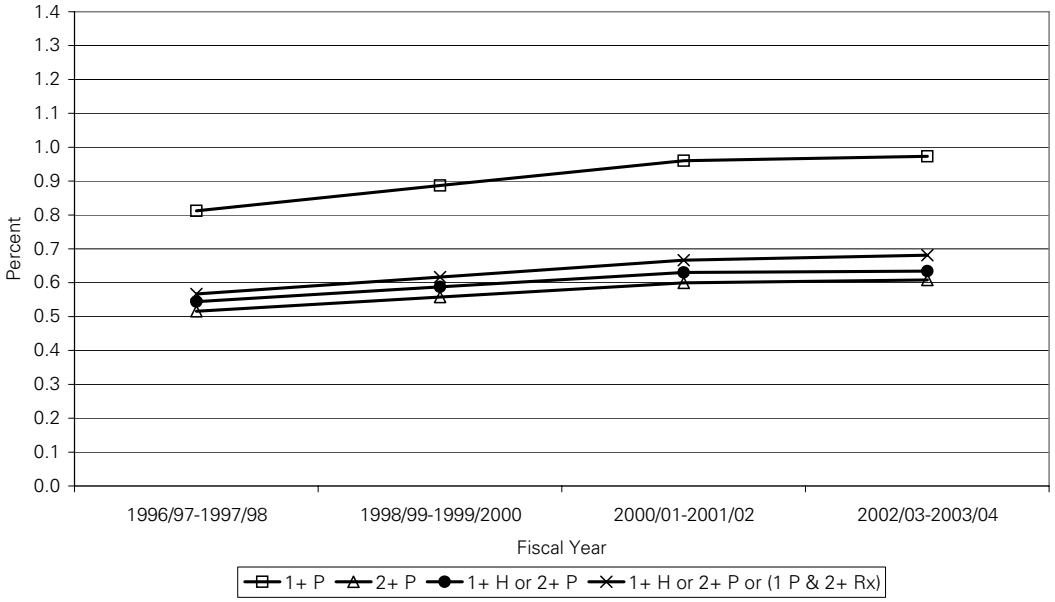
The trends in prevalence estimates for RA for the one-, two-, and three-year algorithms are shown in Figures 18, 19, and 20, respectively. In Figure 18, the trend lines for the four algorithms were approximately parallel, indicating that all one-year algorithms showed similar rates of change in RA prevalence. Moreover, the trends for these one-year algorithms showed a relatively constant prevalence of RA in the most recent five years of the study period. In contrast, both the two-year and three-year algorithms showed a significant increase. There was, however, little difference in the rate of change across the two-year and three-year algorithms.

Figure 18: Provincial Trends in Rheumatoid Arthritis Prevalence for One-Year Algorithms, 1999/2000 – 2003/04



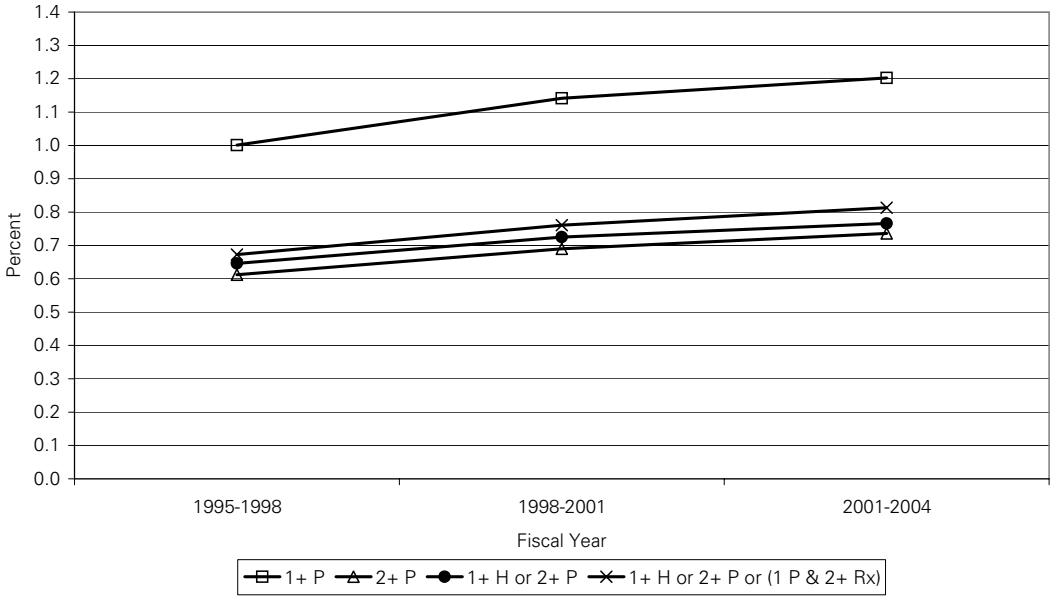
Source: Manitoba Centre for Health Policy, 2006

Figure 19: Provincial Trends in Rheumatoid Arthritis Prevalence for Two-Year Algorithms, 1996/97 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

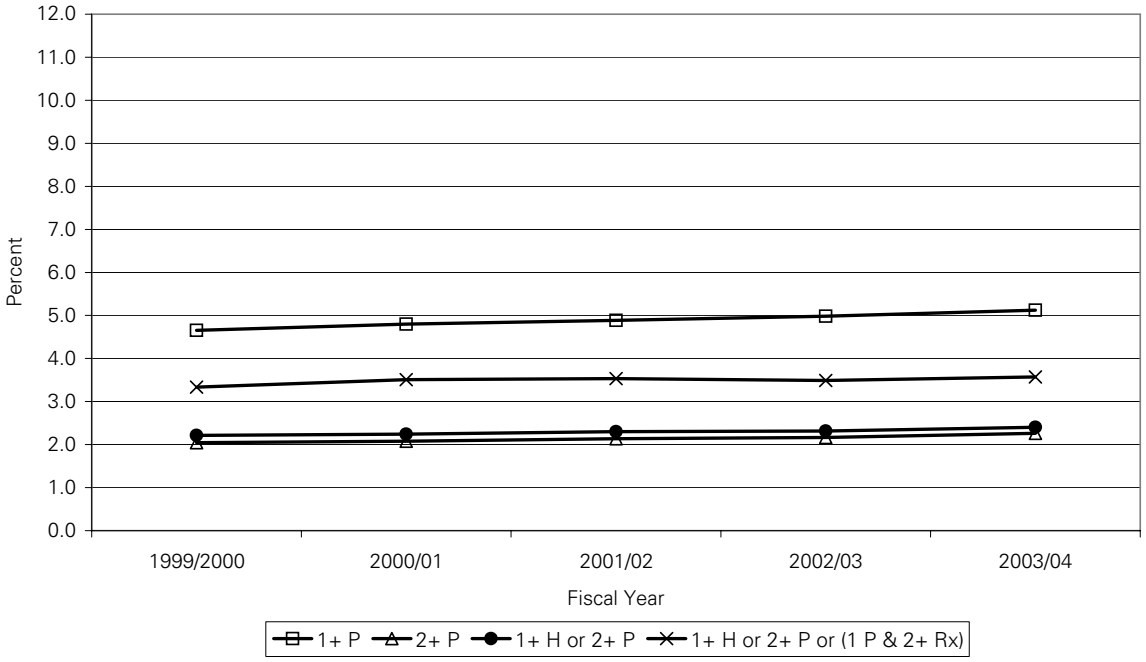
Figure 20: Provincial Trends in Rheumatoid Arthritis Prevalence for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

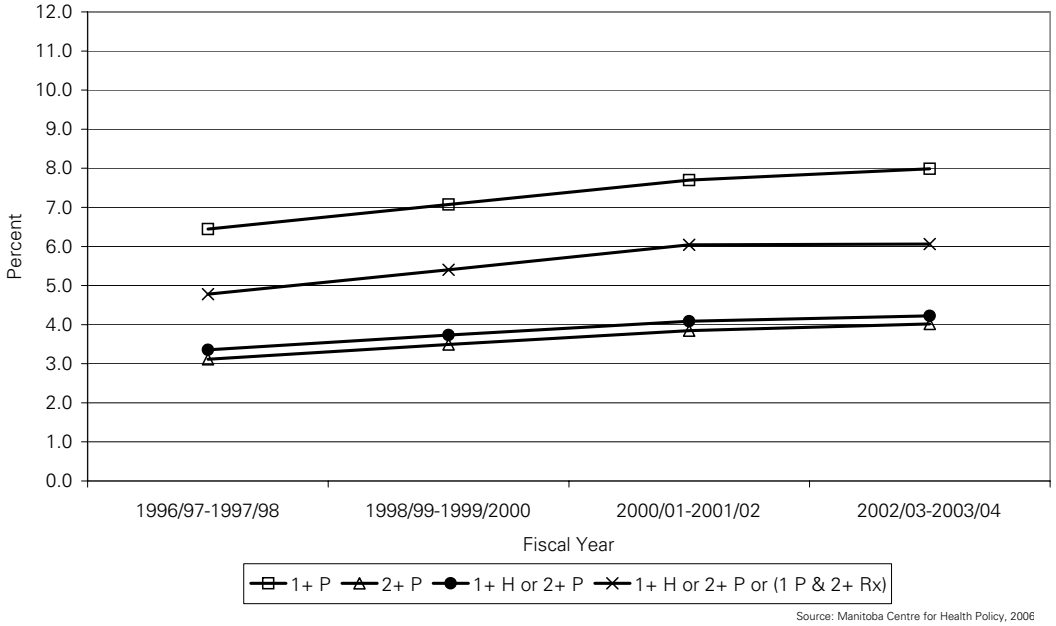
Figures 21, 22, and 23, depict the trends in prevalence estimates for OA for one-, two-, and three-year algorithms, respectively. The trend estimates for the one-year algorithms showed that the algorithm based on one or more contacts in physician billing claims (i.e., algorithm #1) and the algorithm based on a combination of all three data sources (i.e., algorithm #4) showed a slightly greater increase over time than the remaining two algorithms. The same pattern was evident for the two-year results, where algorithms #5 and #8 showed the greatest increase over time, and for the three-year results, where algorithms #9 and #12 showed the greatest increase over time.

Figure 21: Provincial Trends in Osteoarthritis Prevalence for One-Year Algorithms, 1999/2000 – 2003/04



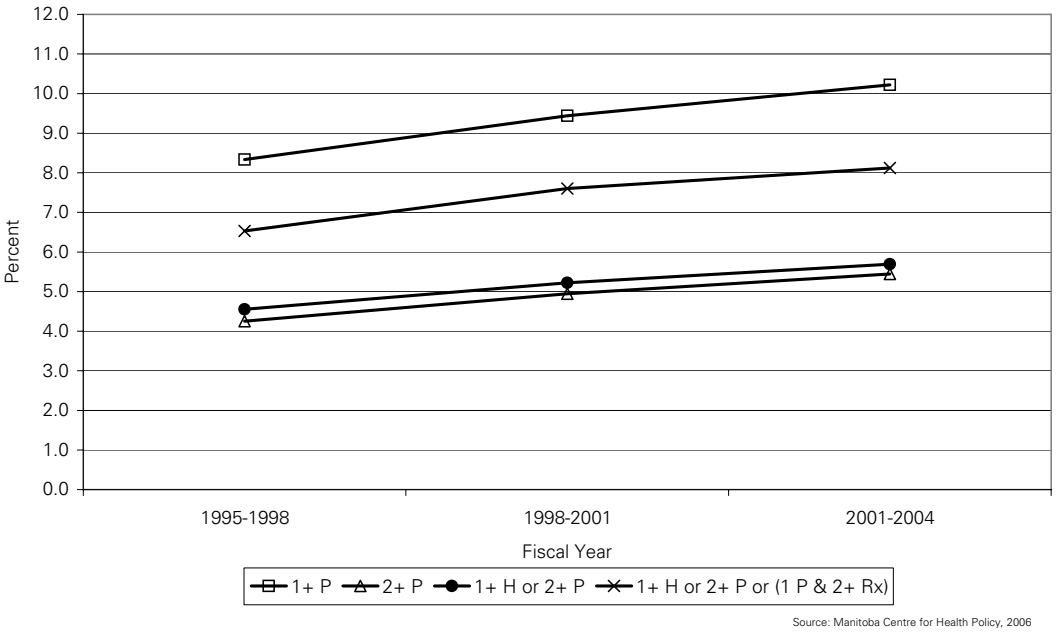
Source: Manitoba Centre for Health Policy, 2006

Figure 22: Provincial Trends in Osteoarthritis Prevalence for Two-Year Algorithms, 1996/97 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Figure 23: Provincial Trends in Osteoarthritis Prevalence for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Regression Analyses for Longitudinal Prevalence Estimates

Regression analyses were conducted for these longitudinal prevalence estimates, to test for differences among the algorithms in the RR over time, and also to test whether the trends for the algorithms varied across the sociodemographic variables of age, sex, region of residence, and income quintile. Three sets of regression analyses were conducted, which focused on the one-, two-, and three-year algorithms for all forms of arthritis, RA, and OA. The model testing results are summarized in Table 11.

Table 11: Summary of likelihood ratio test (LRT) results for longitudinal arthritis prevalence estimates

# Years	Model Comparison	Arthritis	RA	OA
1	Full & Reduced Model	χ^2 = 1799.1 df = 94 p < .0001	χ^2 = 185.4 df = 94 p < .0001	χ^2 = 1184.1 df = 94 p < .0001
	Partial & Reduced Model	χ^2 =0.1 df = 3 p = .9873	χ^2 = 1.4 df = 3 p = .7080	χ^2 = 1.5 df = 3 p = .6948
2	Full & Reduced Model	χ^2 = 1607.8 df = 94 p < .0001	χ^2 = 172.8 df = 94 p < .0001	χ^2 = 1199.4 df = 94 p < .0001
	Partial & Reduced Model	χ^2 = 2.26 df = 3 p = .5205	χ^2 = 1.8 df = 3 p = .7610	χ^2 = 1.9 df = 3 p = 0.5859
3	Full & Reduced Model	χ^2 = 1258.4 df = 94 p < .0001	χ^2 =153.9 df = 94 p < .0001	χ^2 = 1104.2 df = 94 p < .0001
	Partial & Reduced Model	χ^2 =2.03 df = 3 p = .5661	χ^2 =0.3 df = 3 p = .9537	χ^2 = 1.7 df = 3 p = .6335

Note: Full model contains main effects of algorithm, age, sex, region, and income quintiles and time period, two-way interactions of algorithm x age, algorithm x sex, algorithm x region, algorithm x time period, and three-way interactions of algorithm x age x time, algorithm x sex x time, algorithm x region x time, and algorithm x quintile x time; Partial model contains all main effects and algorithm x time; Reduced model contains main effects only. Results which are statistically significant indicate an improvement in model fit with the addition of one or more interaction terms.

Source: Manitoba Centre for Health Policy, 2006

The results are consistent for all forms of arthritis, RA, and OA. The LRTs for the partial and reduced models were never significant. However, the LRTs for the full and reduced models were all statistically significant. The GEE model results for all of the full models revealed that there were statistically significant interaction effects (p < .05) for algorithm x age, algorithm x region, algorithm x age x time, and algorithm x region x time. These results indicate that the RR of change in prevalence was not statistically different across the algorithms. At the same time, the RR of change in prevalence for different algorithms was not constant across the age and region of residence of the Manitoba population.

3.5 Chapter Summary

The validation results reveal that administrative data exhibited fair to moderate agreement with survey data for all forms of arthritis and osteoarthritis and poor agreement for rheumatoid arthritis.

The validation results reveal that administrative data exhibited fair to moderate agreement with survey data for all forms of arthritis and osteoarthritis and poor agreement for rheumatoid arthritis. Furthermore, when survey data were adopted as the gold standard, the validity of administrative data for identifying cases of arthritis and osteoarthritis was moderate, but poor for identifying cases of rheumatoid arthritis. However, the very low estimates of sensitivity and Youden's index for rheumatoid arthritis are likely attributable to bias in the validation source. The estimate of prevalence for the validation cohort was 8.2%, which is much higher than the prevalence that has been reported in other national and international studies. The provincial prevalence estimates for rheumatoid arthritis produced using Manitoba's administrative data are similar to the estimates produced in other Canadian jurisdictions from administrative data (e.g., Lacaille et al., 2005).

The algorithm that resulted in the highest agreement between survey and administrative data is not the same for all forms of arthritis, rheumatoid arthritis, and osteoarthritis. For all forms of arthritis, the algorithm that exhibited the highest agreement between the two sources was based on just two years of data, but relied on a combination of all three sources of administrative data. For rheumatoid arthritis, the algorithm that exhibited the highest agreement between the two sources required only one or more physician billing claims in five years. For osteoarthritis, the algorithm with the highest agreement between the two data sources was also based on one or more physician claims in five years of data. However, it is important to note that agreement between the survey and administrative data was predicted by several sociodemographic characteristics. This means that if the data were stratified by age group or other variables, agreement levels may not be consistently high for the same algorithms as identified in these aggregate results.

If maximum sensitivity and specificity is of primary interest, then for all forms of arthritis the algorithm based on one or more hospital separations, or two or more physician billing claims in five years should be adopted. For rheumatoid arthritis, the algorithm based on one or more physician billing claims resulted in the maximum value of Youden's index. For osteoarthritis, it was also the case that the five-year algorithm based on one or more physician billing claims resulted in the maximum estimate of this index. However, it should be noted that other algorithms based on five years of data produced similar results.

Regression analyses applied to the cross-sectional data revealed that the relative rate of arthritis prevalence for different algorithms varied across age groups. For rheumatoid arthritis no interactions between algorithm and the

sociodemographic characteristics of the population were observed. The relative rate of osteoarthritis prevalence for the investigated algorithms varied across age groups and regions of the province. Regression analyses applied to the longitudinal data revealed that the relative rate of change in prevalence of arthritis, rheumatoid arthritis, and osteoarthritis for different algorithms varied with the sociodemographic characteristics of the population, but did not vary over time. These results, combined with those from the cross-sectional analyses, indicate that the algorithms do not provide the same picture of the relative difference in the prevalence of arthritis and osteoarthritis across sociodemographic groups within the population.

CHAPTER 4: ASTHMA

4.1 Introduction and Review of Literature

Asthma affects a large percentage of both children and adult populations.

From the 1996/97 NPHS, physician-diagnosed asthma was estimated to occur in 12.8% of youth 10 to 14 years of age, and 14.1% of youth 15 to 19 years of age.

Asthma affects a large percentage of both children and adult populations. National estimates of asthma prevalence have been derived primarily from survey data. From the 1996/97 NPHS, physician-diagnosed asthma was estimated to occur in 12.8% of youth 10 to 14 years of age, and 14.1% of youth 15 to 19 years of age. In the adult population this estimate was placed at 6.3% (The National Asthma Control Task Force, 2000). Chen et al. (2005) used CCHS data from 2003 and estimated that 8.4% of the population 12 years of age and older have been diagnosed with asthma. For Manitoba, the estimated prevalence using CCHS data was 8.9%. Using National Health Interview Survey (NHIS) data from the U.S., Dey and Bloom (2003) estimated that 12.5% of children 18 years of age or younger have been diagnosed with asthma. Rhodes et al. (2003) distinguished between lifetime asthma prevalence and current asthma prevalence; using data from the U.S. Behavioral Risk Factor Surveillance Survey (BRFSS) they estimated current asthma prevalence at 7.2% in individuals 18 years of age and older, and lifetime asthma prevalence at 11.0% in this same age group.

Table A.2 in Appendix A summarizes eight published studies that used administrative data to identify asthma cases. Three of the studies used Manitoba data; among these studies, two used only physician billing claims to identify asthma cases, while the third (Kozyrskyj et al., 2004) used a combination of hospital separations, physician claims, and prescription drug records to identify asthma cases. For the studies that reported validation results, estimates of sensitivity, specificity, and predictive values varied substantially. For example, in the Huzel et al. (2002) study, the maximum sensitivity for the adult population was 70.1%, but specificity was very high (99.8%). Kozyrskyj et al. (2004) achieved a much higher sensitivity of 93.9% in a youth population, but specificity was slightly lower (91.4%). Borzecki et al. (2004) examined administrative data algorithms for identifying both COPD and asthma cases (combined), and achieved a sensitivity of 81.0% and a specificity of 92.0%. Wilchesky et al. (2004) concluded that physician data alone are not a valid source for asthma case identification; the authors observed a maximum sensitivity of only 43.0% in an adult population using Quebec's physician claims data to identify asthma cases.

4.2 Description of Asthma Algorithms

A single ICD-9-CM code, 493, was used to identify asthma cases in hospital and physician data. This is consistent with all of the previous studies reported in the summary table. The methodology to extract prescription drug records from the pharmaceutical database was based on previous

research (Kozyrskyj et al., 2004) and consultations with MCHP researchers with expertise in using administrative data to identify cases of asthma. DINs with the following third-level ATC codes were initially selected: R03A (adrenergics, inhalants), R03B (other drugs for obstructive airway diseases, inhalants), R03C (other drugs for obstructive airway diseases, inhalants), and R03D (other systemic drugs for obstructive airway diseases). A very small number of DINs within these four classes were excluded after discussions with experts. As well, one drug outside of these four classes (R06AX17 Ketotifen) was included. Appendix E contains a list of the drugs selected for this research, and identifies the exclusions that were made.

Table 12 lists the 28 algorithms that were investigated in this study. Like the arthritis algorithms, the asthma algorithms were based on as many as five years of administrative data. Two of the algorithms in each time period were based on only physician billing claims, one algorithm was based on only prescription drug data, and the remaining algorithms were based on a combination of two or more of these data sources. For example, algorithm #1 identifies individuals as asthma cases if they had one or more physician claims with an asthma diagnosis code in a one-year period. Algorithm #7 identifies individuals as asthma cases if they had one or more hospital separations, or two or more physician claims, or two or more prescription drug records with relevant diagnostic or medication codes in one year. The algorithms for two, three, and five years of administrative data are interpreted in a similar way.

Table 12: Asthma algorithms selected for validation

# Years	Algorithm	Hospital Separations or	Physician Claims or	Prescription Drug Records
1	1		1 or more	
	2		2 or more	
	3			1 or more
	4	1 or more	1 or more	
	5	1 or more	2 or more	
	6	1 or more	1 or more	1 or more
	7	1 or more	2 or more	2 or more
2	8		1 or more	
	9		2 or more	
	10			1 or more
	11	1 or more	1 or more	
	12	1 or more	2 or more	
	13	1 or more	1 or more	1 or more
	14	1 or more	2 or more	2 or more
3	15		1 or more	
	16		2 or more	
	17			1 or more
	18	1 or more	1 or more	
	19	1 or more	2 or more	
	20	1 or more	1 or more	1 or more
	21	1 or more	2 or more	2 or more
5	22		1 or more	
	23		2 or more	
	24			1 or more
	25	1 or more	1 or more	
	26	1 or more	2 or more	
	27	1 or more	1 or more	1 or more
	28	1 or more	2 or more	2 or more

Source: Manitoba Centre for Health Policy, 2006

It is important to note that unlike the algorithms for arthritis, the algorithms for asthma did not require that the prescription drug records with the relevant medication codes appear in combination with a diagnostic code in physician billing claims. This is because the drugs identified for inclusion in the study are specific to the treatment asthma, and would be used only infrequently for the treatment of other chronic diseases.

The validation analyses were conducted for three age groups: 12 to 18 years, 19 to 49 years, and 50+ years, as well as for the combined age groups (i.e., 12+ years of age). We selected these three age groups because only a limited number of studies have validated asthma algorithms in younger populations. As well, confounding of the diagnoses of asthma and COPD may result in lower validity of asthma algorithms in older age groups.

4.3 Validation Results

Validation Indices

Table 13 contains the point estimates for the six validation indices for each of the 28 algorithms for the combined age groups (i.e., 12 years of age or older). The corresponding 95% CIs are reported in Table D.1 in Appendix D. Appendix F contains the results when cases of self-reported COPD were excluded from the validation cohort. These supplementary results are very similar to those reported in this section and therefore will not be discussed in great detail.

There was fair to moderate agreement between the administrative and survey data, with values of κ ranging from 0.24 to 0.59. The highest value was for the five-year algorithm based on one or more hospital separations or two or more physician claims or two or more prescription drug records. However, this estimate was almost identical to the estimate for the corresponding three-year algorithm (0.58).

Sensitivity was highly variable, and ranged from 18.1% to 84.3%. It was consistently the case that for the one-, two-, three-, and five-year results, the two algorithms that were the most sensitive were based on one or more prescription drug records, or one or more contacts in hospital separations or physician billing claims or prescription drug records. The highest specificity was observed for the five-year algorithms.

Specificity was consistently high. It ranged from 88.6% to 99.5% across the 28 algorithms. Overall, the most specific algorithms were those based on two or more physician claims in one, two, three, or five years of data (i.e., algorithms #2, #9, #16, and #23). However, there was very little change in specificity when two or more physician billing claims were required instead of only one physician billing claim, and the decrease in sensitivity was sub-

stantial. For example, the one-year algorithm based on one or more physician billing claims (i.e., algorithm #1) had a sensitivity of 30.8% while the algorithm based on two or more physician billing claims (i.e., algorithm #2) had a sensitivity of 18.1%.

Youden's index ranged from 0.18 to 0.73. The highest value was observed for the five-year algorithm based on one or more hospital separations or one or more physician billing claims or one or more prescription drug records in five years. However, Youden's index was very similar for the five-year algorithm based on only one or more prescription drug records (0.72), as well as for the algorithm based on one or more hospital separations or two or more physician billing claims or two or more prescription drug records in five years (0.70).

Table 13: Estimates of agreement, sensitivity, specificity, and predictive values for asthma algorithms, all ages

# Years	Algorithm		κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.40	30.8	98.8	0.30	70.3	94.1
	2	2+ P	0.27	18.1	99.5	0.18	77.4	93.1
	3	1+ Rx	0.51	55.4	95.9	0.51	54.7	96.0
	4	1+ H or 1+ P	0.40	31.4	98.8	0.30	70.3	94.1
	5	1+ H or 2+ P	0.28	18.9	99.5	0.18	77.5	93.2
	6	1+ H or 1+ P or 1+ Rx	0.52	58.4	95.5	0.54	53.7	96.2
	7	1+ H or 2+ P or 2+ Rx	0.51	46.9	97.7	0.45	64.4	95.3
2	8	1+ P	0.37	43.9	98.0	0.42	66.1	95.1
	9	2+ P	0.26	30.1	99.1	0.29	74.3	94.0
	10	1+ Rx	0.54	69.8	93.7	0.64	49.9	97.2
	11	1+ H or 1+ P	0.38	44.8	97.9	0.43	66.2	95.2
	12	1+ H or 2+ P	0.28	31.2	99.0	0.30	74.3	94.1
	13	1+ H or 1+ P or 1+ Rx	0.54	72.6	93.1	0.66	48.4	97.4
	14	1+ H or 2+ P or 2+ Rx	0.55	57.5	96.7	0.54	60.9	96.2
3	15	1+ P	0.36	52.9	96.9	0.50	60.7	95.8
	16	2+ P	0.24	37.6	98.6	0.36	71.3	94.6
	17	1+ Rx	0.52	74.7	92.3	0.67	46.5	97.6
	18	1+ H or 1+ P	0.37	53.7	96.9	0.51	60.8	95.9
	19	1+ H or 2+ P	0.28	38.6	98.6	0.37	71.1	94.7
	20	1+ H or 1+ P or 1+ Rx	0.52	77.7	91.2	0.69	44.2	97.9
	21	1+ H or 2+ P or 2+ Rx	0.58	66.2	95.7	0.62	57.8	96.9
5	22	1+ P	0.45	63.5	95.3	0.59	55.0	96.7
	23	2+ P	0.35	50.3	97.7	0.48	66.2	95.6
	24	1+ Rx	0.51	81.5	90.2	0.72	42.7	98.2
	25	1+ H or 1+ P	0.48	63.7	95.3	0.59	54.9	96.7
	26	1+ H or 2+ P	0.37	50.7	97.6	0.48	65.7	95.7
	27	1+ H or 1+ P or 1+ Rx	0.50	84.3	88.6	0.73	39.9	98.4
	28	1+ H or 2+ P or 2+ Rx	0.59	75.4	94.2	0.70	53.7	97.7

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

Within each of the one-, two-, three-, and five-year sets of algorithms, the PPV of an asthma diagnosis was highest for the algorithms based on: (1) two or more physician billing claims or one or more hospital separations, or (2) two or more physician billing claims. The NPV of an asthma diagnosis was consistently above 90% for all of the algorithms, and was greater than 95% for more than half of the algorithms.

The point estimates for the validation indices for the youth cohort 12 to 18 years of age are reported in Table 14. The estimates are similar to those for the entire cohort (see Appendix D, Table D.5 for the 95% CIs). However, there was stronger agreement between survey and administrative data for this age group, with κ ranging from 0.24 (fair agreement) to 0.70 (good agreement). The highest was for the algorithm based on one or more hospital separations or two or more physician billing claims or two or more prescription drug records in five years.

Sensitivity was highly variable and ranged from 16.2% to 87.4%. The five-year algorithm based on one or more contacts in hospital separations or physician billing claims or prescription drug records had the highest sensitivity. The improvement in sensitivity over the corresponding three-year algorithm was substantial (i.e., 9.0%).

Specificity was very high (i.e., above 95%) for almost all of the algorithms. Youden's index was highest for the five-year algorithm based on one or more prescriptions (0.77). It was almost the same for the algorithm based on one or more hospital separations or one or more physician billing claims or one or more prescription drug records in five years, as well as for the algorithm based on one or more hospital separations or two or more physician billing claims or two or more prescriptions in five years.

Overall, the PPV of an asthma diagnosis in this youngest age group was highest for algorithms #9 and #12, which were based on (1) two or more physician claims in a two-year period (88.5%), or (2) one or more hospital separations or two or more physician billing claims (88.9%). The NPV of an asthma diagnosis was consistently high, only falling below 90% for 6 of the 28 investigated algorithms.

Table 14: Estimates of agreement, sensitivity, specificity, and predictive values for asthma algorithms, 12-18 years

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.37	29.7	98.1	0.29	70.2	90.1
	2	2+ P	0.24	16.2	99.4	0.16	81.8	88.5
	3	1+ Rx	0.48	43.2	97.2	0.40	70.6	91.8
	4	1+ H or 1+ P	0.37	29.7	98.1	0.28	70.2	90.1
	5	1+ H or 2+ P	0.24	16.2	99.4	0.16	81.8	88.5
	6	1+ H or 1+ P or 1+ Rx	0.52	48.6	96.8	0.45	70.1	92.5
	7	1+ H or 2+ P or 2+ Rx	0.42	33.3	98.6	0.32	78.7	90.6
2	8	1+ P	0.43	45.9	96.5	0.42	67.1	92.1
	9	2+ P	0.30	33.3	99.0	0.32	84.1	90.6
	10	1+ Rx	0.64	67.6	95.4	0.63	69.4	95.0
	11	1+ H or 1+ P	0.44	45.9	96.5	0.42	67.1	92.1
	12	1+ H or 2+ P	0.31	33.3	99.0	0.32	84.1	90.6
	13	1+ H or 1+ P or 1+ Rx	0.63	69.4	94.5	0.64	65.8	95.3
	14	1+ H or 2+ P or 2+ Rx	0.55	48.6	98.2	0.47	80.6	92.6
3	15	1+ P	0.46	62.2	95.2	0.57	66.3	94.2
	16	2+ P	0.30	47.7	98.3	0.46	81.5	92.4
	17	1+ Rx	0.66	76.6	93.9	0.70	65.9	96.3
	18	1+ H or 1+ P	0.48	62.2	95.2	0.57	66.3	94.2
	19	1+ H or 2+ P	0.34	47.7	98.3	0.46	81.5	92.4
	20	1+ H or 1+ P or 1+ Rx	0.66	78.4	92.4	0.71	61.3	96.5
	21	1+ H or 2+ P or 2+ Rx	0.68	69.4	96.8	0.66	77.0	95.4
5	22	1+ P	0.49	74.8	92.0	0.67	58.9	96.0
	23	2+ P	0.36	60.4	96.7	0.57	73.6	94.1
	24	1+ Rx	0.64	86.5	90.7	0.77	58.9	97.8
	25	1+ H or 1+ P	0.51	74.8	92.0	0.67	58.9	96.0
	26	1+ H or 2+ P	0.38	60.4	96.7	0.57	73.6	94.1
	27	1+ H or 1+ P or 1+ Rx	0.63	87.4	88.5	0.76	53.9	97.9
	28	1+ H or 2+ P or 2+ Rx	0.70	80.2	94.6	0.75	69.5	96.9

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

Table 15 contains the point estimates for the six validation indices for the young adult cohort (i.e., 19 to 49 years of age). The corresponding 95% CIs are reported in Appendix D in Table D.6.

Table 15: Estimates of agreement, sensitivity, specificity, and predictive values for asthma algorithms, 19-49 years

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+P	0.39	29.8	99.0	0.29	71.6	94.5
	2	2+ P	0.26	17.1	99.5	0.17	73.6	93.6
	3	1+ Rx	0.52	50.0	97.5	0.48	62.3	96.0
	4	1+H or 1+P	0.40	30.3	99.0	0.29	71.9	94.5
	5	1+ H or 2+ P	0.27	18.0	99.5	0.18	74.5	93.7
	6	1+H or 1+P or 1+Rx	0.53	53.5	97.1	0.51	60.1	96.2
	7	1+ H or 2+ P or 2+ Rx	0.52	41.7	99.0	0.41	77.9	95.4
2	8	1+P	0.37	42.5	98.5	0.41	69.3	95.4
	9	2+ P	0.24	28.9	99.1	0.28	73.3	94.4
	10	1+ Rx	0.56	65.4	95.7	0.61	55.6	97.1
	11	1+H or 1+P	0.38	43.0	98.4	0.41	69.0	95.5
	12	1+ H or 2+ P	0.26	29.8	99.1	0.29	73.1	94.5
	13	1+H or 1+P or 1+Rx	0.57	70.2	95.1	0.65	53.9	97.5
	14	1+ H or 2+ P or 2+ Rx	0.59	53.5	98.4	0.52	73.1	96.3
3	15	1+P	0.35	48.7	97.6	0.46	62.0	95.9
	16	2+ P	0.21	34.2	98.8	0.33	70.9	94.8
	17	1+ Rx	0.55	70.2	94.4	0.65	50.6	97.5
	18	1+H or 1+P	0.35	49.1	97.5	0.47	61.9	95.9
	19	1+ H or 2+ P	0.24	35.1	98.8	0.34	70.8	94.9
	20	1+H or 1+P or 1+Rx	0.54	74.6	93.3	0.68	47.8	97.8
	21	1+ H or 2+ P or 2+ Rx	0.59	60.1	97.6	0.58	67.5	96.8
5	22	1+P	0.47	60.5	95.9	0.56	54.8	96.7
	23	2+ P	0.35	50.4	97.8	0.48	65.0	96.0
	24	1+ Rx	0.53	78.1	92.2	0.70	45.1	98.1
	25	1+H or 1+P	0.47	60.5	95.9	0.56	54.5	96.7
	26	1+ H or 2+ P	0.36	50.4	97.7	0.48	64.6	96.0
	27	1+H or 1+P or 1+Rx	0.51	82.0	90.4	0.72	41.3	98.4
	28	1+ H or 2+ P or 2+ Rx	0.61	71.1	96.0	0.67	59.1	97.6

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

The κ values were slightly lower for this age group than for the 12 to 19 years age group. They ranged from 0.21 (fair agreement) to 0.61 (good agreement). Consistent with the results for the 12 to 18 years age group, the highest κ was observed for the algorithm based on one or more hospital separations or two or more physician billing claims or two or more prescription drug records in five years.

The maximum observed sensitivity was 82.0% for the five-year algorithm based on one or more contacts in hospital separations or physician billing claims or prescription drug records. Sensitivity was slightly lower for the five-year algorithm based on one or more prescription drug records (78.1%). Specificity was above 90% for all of the algorithms.

Youden's index was almost equivalent for the two algorithms with the highest sensitivity (i.e., approximately 0.70). PPV was highest for all of the algorithms that required at least two physician claims to identify an asthma case. There was a substantial decrease in PPV, of between 10% and 20%, for algorithms based on one or more hospital separations, or one or more physician billing claims, or one or more prescription drug records, and algorithms based on one or more hospital separations, two or more physician claims, or two or more prescription drug records. NPV was consistently high (i.e., above 90%) for all algorithms.

For the 50+ years age group, the validation estimates were similar to those for the other two age groups (see Appendix D, Table D.7 for the 95% CIs). Appendix F contains the corresponding estimates when individuals with self-reported COPD or emphysema were removed from the validation cohort. The estimates for all validation indices are slightly lower when these exclusions were made.

For the oldest age group, values of κ (see Table 16) ranged from 0.24 (fair agreement) to 0.56 (moderate agreement). Consistent with the results for other age groups, the highest κ was for the algorithm based on one or more hospital separations or two or more physician billing claims or two or more prescription drug records in five years. The highest sensitivity was estimated for algorithm #27, which was based on one or more contacts in hospital separations, or physician billing claims or prescription drug records in five years (85.3%). However, the sensitivity for the five-year algorithm based on one or more prescription drug records was similar (82.6%). Specificity was very high for all of the algorithms. Regardless of the number of years of administrative data, Youden's index was highest for the algorithms based on one or more contacts in hospital separations, or physician billing claims, or prescription drug records (0.72), one or more prescription drug records (0.70), and one or more hospital separations or two or more physician billing claims or two or more prescription drug records.

In this oldest age group, the PPV of an asthma diagnosis was highest (79.6%) for the one-year algorithm based on two or more physician billing claims (i.e., algorithm #2). The NPV of an asthma diagnosis was very high (i.e., above 90%) for all of the algorithms.

Table 16: Estimates of agreement, sensitivity, specificity, and predictive values for asthma algorithms, 50+ years

# Years	Algorithm		κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.42	32.6	98.8	0.31	68.9	94.9
	2	2+ P	0.30	20.5	99.6	0.20	79.6	94.0
	3	1+ Rx	0.51	68.9	93.6	0.63	46.0	97.4
	4	1+ H or 1+ P	0.42	33.7	98.8	0.33	68.8	94.9
	5	1+ H or 2+ P	0.32	21.6	99.5	0.21	78.8	94.1
	6	1+ H or 1+ P or 1+ Rx	0.50	70.0	93.2	0.63	45.1	97.5
	7	1+ H or 2+ P or 2+ Rx	0.54	61.1	95.8	0.57	53.7	96.9
2	8	1+ P	0.33	44.2	97.9	0.42	62.2	95.7
	9	2+ P	0.25	29.5	99.0	0.28	70.0	94.6
	10	1+ Rx	0.47	76.3	90.9	0.67	39.9	98.0
	11	1+ H or 1+ P	0.36	46.3	97.8	0.44	62.9	95.8
	12	1+ H or 2+ P	0.28	31.6	99.0	0.31	70.6	94.8
	13	1+ H or 1+ P or 1+ Rx	0.47	77.4	90.3	0.67	38.8	98.0
	14	1+ H or 2+ P or 2+ Rx	0.52	67.4	94.3	0.62	48.3	97.3
3	15	1+ P	0.31	52.6	96.7	0.49	56.2	96.3
	16	2+ P	0.24	35.8	98.5	0.34	65.4	95.1
	17	1+ Rx	0.45	78.9	89.4	0.68	37.1	98.2
	18	1+ H or 1+ P	0.33	54.2	96.7	0.51	56.6	96.4
	19	1+ H or 2+ P	0.28	37.4	98.4	0.36	65.1	95.2
	20	1+ H or 1+ P or 1+ Rx	0.44	81.1	88.4	0.69	35.7	98.3
	21	1+ H or 2+ P or 2+ Rx	0.52	71.6	93.0	0.65	44.9	97.6
5	22	1+ P	0.39	60.5	95.7	0.56	52.8	96.8
	23	2+ P	0.34	44.2	97.9	0.42	62.7	95.7
	24	1+ Rx	0.43	82.6	87.7	0.70	34.7	98.4
	25	1+ H or 1+ P	0.41	61.1	95.6	0.57	52.7	96.9
	26	1+ H or 2+ P	0.37	45.3	97.8	0.43	61.9	95.7
	27	1+ H or 1+ P or 1+ Rx	0.43	85.3	86.4	0.72	33.3	98.7
	28	1+ H or 2+ P or 2+ Rx	0.56	77.9	91.9	0.70	43.4	98.1

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

Agreement Between Survey and Administrative Data

The logistic regression models to test the sociodemographic variables associated with agreement between survey and administrative data contained the main effects of age, sex, region of residence, income adequacy quintile, and comorbidity (i.e., presence of allergies, emphysema, or COPD). Only the data for algorithm #28 for all age groups were modeled. Two-way interactions were tested. None of the interactions were statistically significant, and they were therefore excluded from the final models. The Hosmer-Lemeshow test indicated that the model containing the main effects fit the data well.

The following variables were statistically significant predictors of agreement between the two data sources: age (χ^2 = 38.8, df = 6, p < .0001), comorbidity (χ^2 = 106.0, df = 1, p < .0001), and income adequacy quintile (χ^2 = 21.4, df = 5, p < .0001). Table 17 contains the ORs and 95% CIs for the explanatory variables in the model.

The odds of agreement between the two sources were higher for younger age groups than for the oldest age group. The odds were also higher for individuals who indicated they did not have any of the comorbid conditions than for individuals who did report having one or more comorbid conditions. The odds were also lower for individuals in lower income groups than for individuals in the highest income group.

Table 17: Odds Ratio (OR) estimates and 95% CIs for predictors of agreement between administrative and survey data for asthma, all ages

Predictors	OR	95% CI
Age		
12 – 18 years	1.7	(1.2, 2.4)
19 – 24 years	1.9	(1.3, 3.0)
25 – 34 years	2.0	(1.4, 2.9)
35 – 49 years	2.0	(1.4, 2.9)
50 – 64 years	1.2	(1.6, 3.1)
65 – 74 years	1.8	(1.3, 2.5)
75+ years	Ref	--
Sex		
Males	0.9	(0.8, 1.1)
Females	Ref	--
Region of Residence		
North Rural RHAs	1.2	(0.8, 1.7)
South Rural RHAs	1.0	(0.8, 1.2)
Winnipeg RHA	Ref	--
Comorbidity		
Absent	2.4	(2.0, 3.0)
Present	Ref	--
Income Quintile		
Lowest	0.4	(0.2, 0.6)
Low Middle	0.7	(0.5, 1.1)
Middle	0.6	(0.4, 0.9)
Upper Middle	0.6	(0.5, 0.9)
Not Stated	0.8	(0.5, 1.3)
Highest	Ref	--

Source: Manitoba Centre for Health Policy, 2006

4.4 Provincial Prevalence Estimates

Cross-Sectional Prevalence Estimates

Cross-sectional prevalence estimates for each of the asthma algorithms are reported in Table 18. These estimates are provided for the entire population 12 years of age and older, as well as for each of the age groups.

The prevalence estimates varied substantially with the source of data and number of years of data. The algorithm that resulted in the highest overall estimate of κ (i.e., algorithm #28) produced a prevalence estimate of 11.6%

for all ages and age-specific estimates of 15.3%, 9.8%, and 12.8% for the 12 to 18 years, 19 to 49 years, and 50+ years age groups, respectively. The algorithm that resulted in the highest estimate of sensitivity for all ages, which was based on one or more hospital separations, or one or more physician claims, or one or more prescription drug records in five years (i.e., algorithm #27), produced a prevalence estimate of 17.5% in the Manitoba population 12 years of age and older. The corresponding age-group specific estimates for this algorithm were 22.5% for the population 12 to 18 years of age, 15.6% for the population 19 to 49 years of age, and 18.4% for the population 50 + years of age. The algorithm with similar sensitivity, which was based on one or more contacts in the prescription drug records (i.e., algorithm #24), resulted in a prevalence estimate of 15.5% in the Manitoba population, and age-specific estimates of 19.5% for youth (12–18 years), 13.6% for younger adults (19–49 years), and 16.7% for older adults (50+ years).

Table 18: Crude provincial prevalence estimates for asthma algorithms, 1998/99 – 2002/03

# Years		Algorithm	All Ages (%)	12 – 18 Years (%)	19 – 49 Years (%)	50+ Years (%)
1	1	1+P	3.6	5.3	3.4	3.2
	2	2+ P	1.9	2.7	1.8	1.9
	3	1+ Rx	7.1	7.7	5.7	9.0
	4	1+H or 1+P	3.6	5.4	3.4	3.3
	5	1+ H or 2+ P	2.0	2.7	1.8	2.0
	6	1+H or 1+P or 1+Rx	8.0	9.0	6.6	9.7
	7	1+ H or 2+ P or 2+ Rx	5.3	5.3	4.0	7.1
2	8	1+P	5.6	8.8	5.4	5.0
	9	2+ P	3.4	5.1	3.1	3.2
	10	1+ Rx	9.8	11.3	8.3	11.6
	11	1+H or 1+P	5.7	8.8	5.4	5.1
	12	1+ H or 2+ P	3.5	5.1	3.2	3.4
	13	1+H or 1+P or 1+Rx	11.0	13.2	9.4	12.6
	14	1+ H or 2+ P or 2+ Rx	7.4	8.5	5.8	9.1
3	15	1+ P	7.3	11.6	6.9	6.4
	16	2+ P	4.7	7.1	4.3	4.3
	17	1+ Rx	11.9	14.2	10.3	13.4
	18	1+ H or 1+ P	7.4	11.6	6.9	6.6
	19	1+ H or 2+ P	4.8	7.2	4.4	4.6
	20	1+H or 1+ P or 1+ Rx	13.4	16.5	11.7	14.7
	21	1+ H or 2+ P or 2+ Rx	9.0	11.0	7.3	10.6
5	22	1+ P	10.0	16.5	9.5	8.6
	23	2+ P	6.8	10.9	6.3	6.1
	24	1+ Rx	15.5	19.5	13.6	16.7
	25	1+ H or 1+P	10.1	16.6	9.6	8.8
	26	1+ H or 2+ P	6.9	11.0	6.4	6.3
	27	1+ H or 1+ P or 1+ Rx	17.5	22.5	15.6	18.4
	28	1+ H or 2+ P or 2+ Rx	11.6	15.3	9.8	12.8

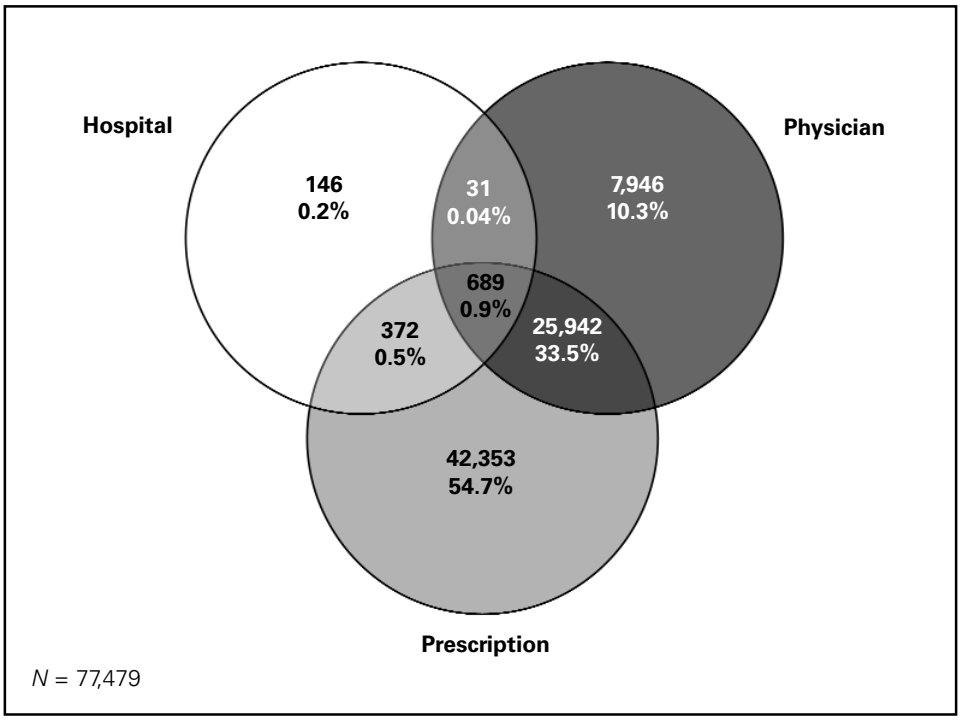
Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 1-year estimates are for 2002/03, 2-year estimates are for 2001/02 – 2002/03, 3-year estimates are for 2000/01 – 2002/03, 5-year estimates are for 1998/99 – 2002/03.

Source: Manitoba Centre for Health Policy, 2006

Venn Diagrams

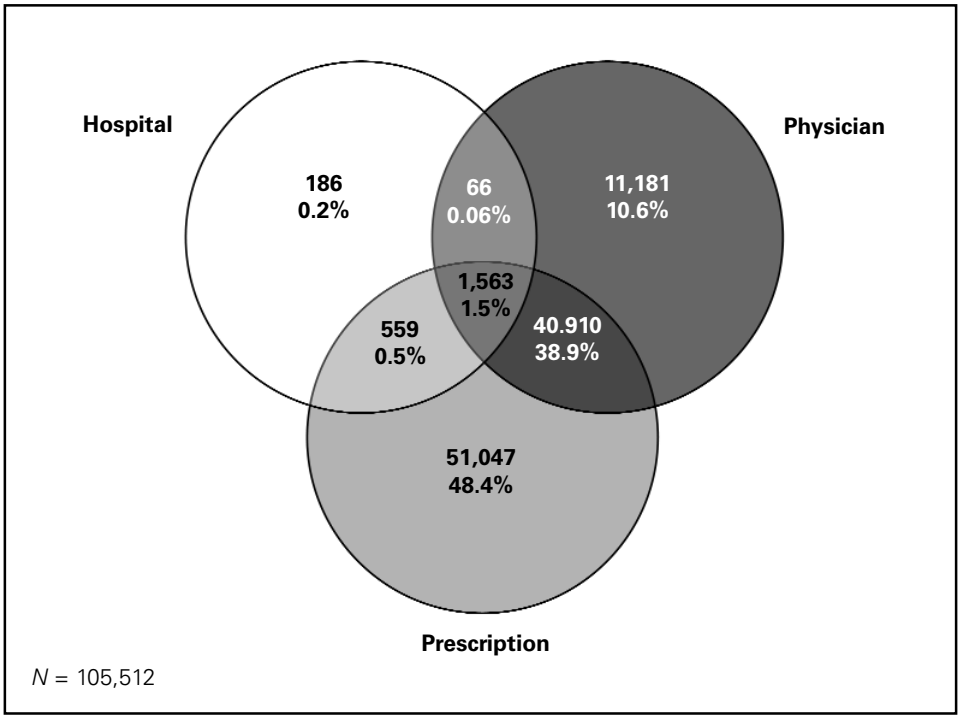
Venn diagrams are presented for algorithms #6, #13, #20, and #27. These are the algorithms based on one or more hospital separations, or one or more physician billing claims, or one or more prescription drug records in one, two, three, or five years, respectively. The Venn diagrams describe the number and percent of asthma cases identified in each of the three data sources for the Manitoba population 12 years of age and older. The Venn diagrams for age-specific groups are not reported because they produced similar results.

Figure 24: Asthma Algorithm #6: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 1 Year



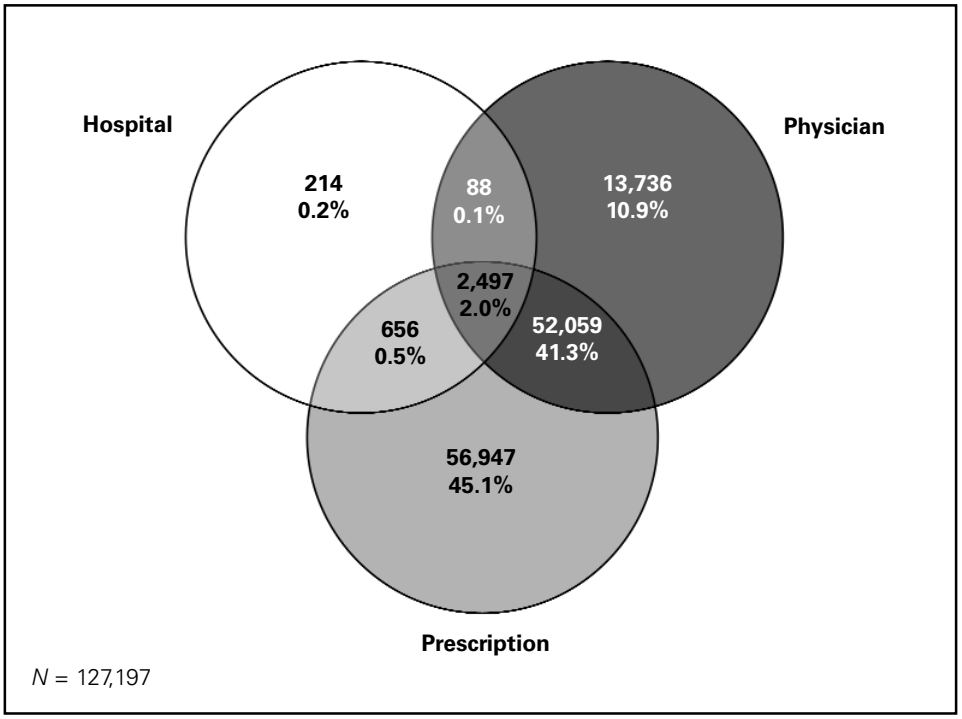
Source: Manitoba Centre for Health Policy, 2006

Figure 25: Asthma Algorithm #13: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 2 Years



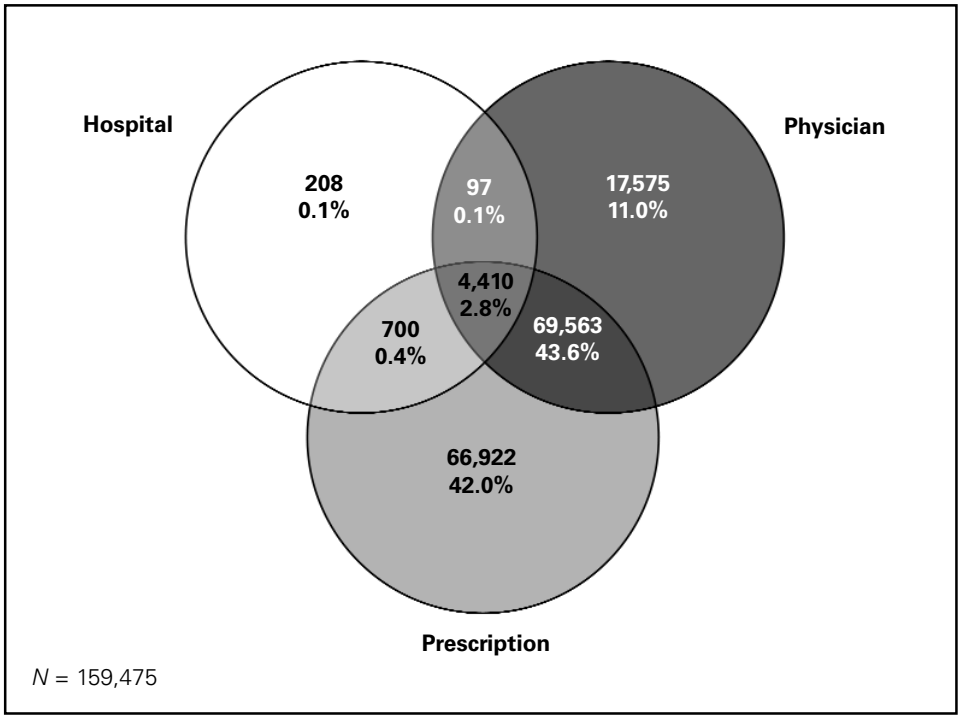
Source: Manitoba Centre for Health Policy, 2006

Figure 26: Asthma Algorithm #20: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 3 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 27: Asthma Algorithm #28: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 5 Years



Source: Manitoba Centre for Health Policy, 2006

The Venn diagram for algorithm #6 (Figure 24) showed that there were more than 77,000 disease cases identified when this algorithm was applied to the Manitoba administrative data for 2002/03. Overall, only a very small percent of disease cases were identified exclusively from the hospital data (0.2%), and there was minimal overlap between either the hospital separations and physician claims data (0.04%) or the hospital separations and prescription records data (0.5%) using one year of data. About 10% of disease cases were identified only from physician billing claims, but more than half (54.5%) were identified only from the prescription drug data. There was substantial overlap between the physician and prescription data sources using one year of data (33.3%).

Figures 25 to 27 show that as the number of years of data increased, there was greater overlap among the three sources. For example, in one year of data, only 0.9% of cases had a contact in all three of the data sources. This increased to 1.5% of cases using two years of data and 2.8% of cases using five years of administrative data. The percent overlap between the hospital and prescription data sources remained constant across the algorithms (i.e., approximately 0.4%) but the percent overlap between the physician and prescription data increased over time, from 38.8% using two years of administrative data to 43.6% using five years of data.

Regression Analyses for Cross-Sectional Prevalence Estimates

In the regression analyses for all age groups, full and reduced models were specified that included the algorithms based on one or more contacts in the prescription drug records or one or more contacts in hospital separations or physician billing claims or prescription drug records in one, two, three, or five years of administrative data (i.e., algorithms #3, #6, #9, #13, #16, #19, #23, and #27). Thus, a total of eight algorithms were included in the regression models.

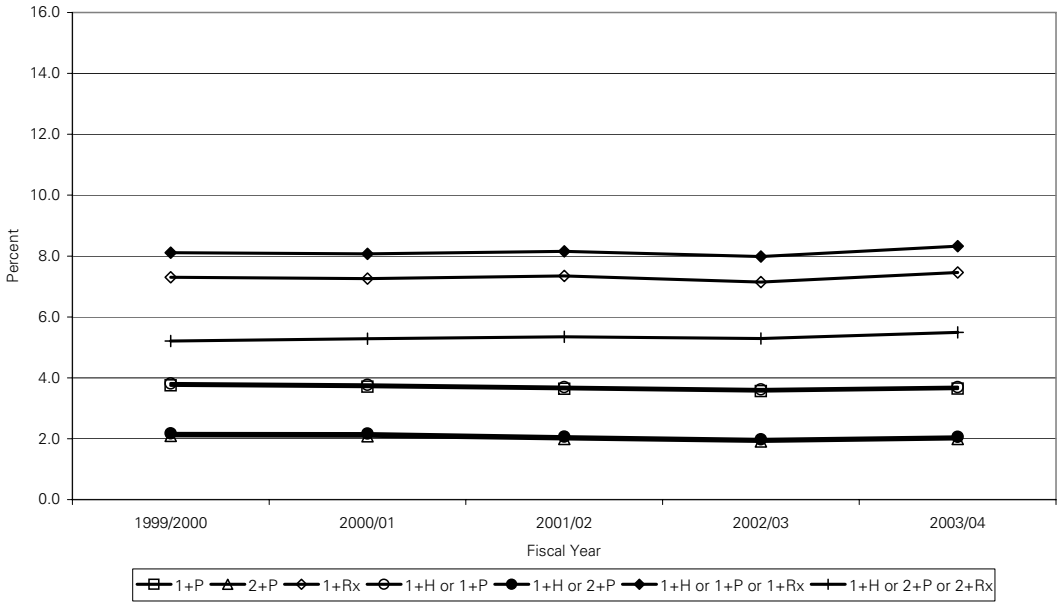
The LRT for the full model, which contained the main effects of algorithm, age, sex, region, and income quintile in addition to selected two-way interactions, and the reduced model, which contained main effects only, was not statistically significant (χ^2 = 87.2; df = 91, p = .5932). A non-significant LRT for the full and reduced models was also observed for the 12 to 18 years age group (χ^2 = 4.62; df = 49, p = .99), the 19 to 49 years age group (χ^2 =87.2; df = 81, p = .5932), and the 50+ years age group (χ^2 = 89.2; df = 91, p = .5932). This result indicates that the RR of asthma for different algorithms did not vary with the sociodemographic characteristics of the population.

In all four of the reduced models there was a statistically significant algorithm main effect: χ^2 =1222.6; df = 7, p < .0001 for all age groups; χ^2 = 729.9; df = 7, p < .0001 for 12 to 18 years; χ^2 =1222.6; df = 7, p < .0001 for 19 to 49 years; χ^2 =1222.6; df = 7, p < .0001 for 50+ years. This result indicates that there were statistically significant differences in the prevalence estimates for different algorithms. For the model for all age groups, there were significant differences between the prevalence estimate for algorithm #6 (reference) and the estimates for algorithms #3 and #10. For the 12 to 18 years age group, there was no significant difference in the prevalence estimate for algorithm #6 (reference) and the estimates for any of the other one-year algorithms. However, all other algorithms produced estimates which were significantly different from the estimate for algorithm #6. For the 19 to 49 years age group, this same result was observed. For the 50+ years age group, there was no significant difference between the prevalence estimate for algorithm #6 (reference) and the estimates for algorithms #3 and #10.

Longitudinal Prevalence Estimates

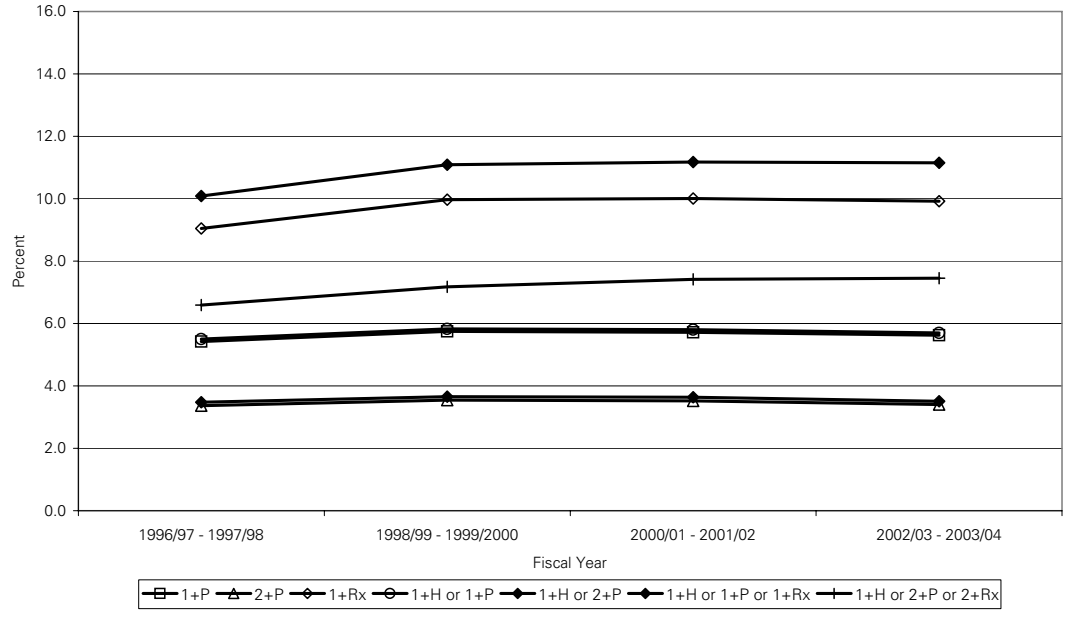
Trends in the prevalence of asthma for all groups for the one-, two-, and three-year algorithms are reported in Figures 28, 29, and 30, respectively. The corresponding age-specific prevalence trend estimates are not reported because they resulted in similar longitudinal profiles.

Figure 28: Provincial Trends in Asthma Prevalence for One-Year Algorithms, All Ages, 1999/2000 – 2003/04



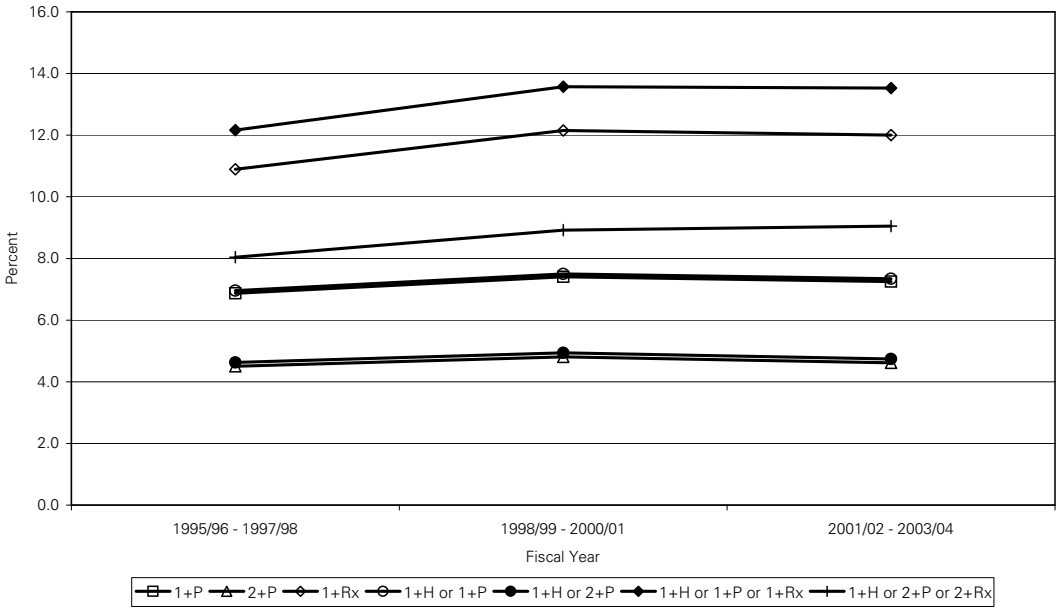
Source: Manitoba Centre for Health Policy, 2006

Figure 29: Provincial Trends in Asthma Prevalence for Two-Year Algorithms, All Ages, 1996/97 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Figure 30: Provincial Trends in Asthma Prevalence for Three-Year Algorithms, All Ages, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

The estimates depicted in Figure 28 show some change in the overall prevalence of asthma. All three of the algorithms based on prescription drug records showed a sharp increase between 2002/03 and 2003/04. This increase was not observed for the algorithms based solely on hospital and physician data.

The two-year algorithms depicted in Figure 29 showed almost no change in the overall prevalence of asthma in the population for the last three time periods. However, the algorithms that incorporated prescription drug data show a substantial increase between the period 1996/97–1997/98 and 1998/99–1999/2000. This increase is not evident in the algorithms based only on hospital separations and physician billing claims data. This same pattern was evident in Figure 30, which describes the trend in asthma prevalence for the three-year algorithms between 1995/96 and 2003/04.

Regression Analyses for Longitudinal Prevalence Estimates

Regression analyses were conducted for the longitudinal asthma prevalence estimates, to test for differences among the algorithms in the RR over time, and also to test whether the longitudinal estimates for different algorithms varied across the sociodemographic variables of age, sex, region of residence, and income quintile. Three sets of analyses were conducted, which focused on the one-, two-, and three-year algorithms separately for all age groups as well as for each of the three age groups. The results for the LRTs for the full, partial, and reduced models are summarized in Table 19.

Table 19: Summary of likelihood ratio test (LRT) results for models of longitudinal arthritis prevalence

# Years	Model Comparison	All Ages	12-19 Years	19-49 Years	50+ Years
1	Full & Reduced Model	χ^2 =1527.9 df = 175 $p < .0001$	χ^2 = 126.8 df = 97 $p = .0226$	χ^2 = 255.0 df = 123 $p < .0001$	χ^2 = 1054.2 df = 123 $p < .0001$
	Partial & Reduced Model	χ^2 = 27.2 df = 6 $p = .0001$	χ^2 = 1.8 df = 6 $p = .9387$	χ^2 = 9.3 df = 6 $p = .1555$	χ^2 = 44.5 df = 6 $p < .0001$
2	Full & Reduced Model	χ^2 = 655.8 df = 175 $p < .0001$	χ^2 =100.1 df = 97 $p = .3946$	χ^2 =220.4 df = 123 $p < .0001$	χ^2 =511.2 df = 123 $p < .0001$
	Partial & Reduced Model	χ^2 = 27.2 df = 6 $p = .0001$	χ^2 = 6.0 df = 6 $p = .4239$	χ^2 = 9.7 df = 6 $p = .1377$	χ^2 = 37.3 df = 6 $p < .0001$
3	Full & Reduced Model	χ^2 = 362.2 df = 175 $p < .0001$	χ^2 = 100.2 df = 97 $p = .3914$	χ^2 =170.0 df = 123 $p = .0033$	χ^2 = 313.1 df = 123 $p < .0001$
	Partial & Reduced Model	χ^2 = 27.0 df = 6 $p = .0001$	χ^2 = 8.26 df = 6 p = .2194	χ^2 =12.8 df = 8 $p = .0462$	χ^2 = 26.0 df = 6 $p = .0002$

Note: Full model contains main effects of algorithm, age, sex, region, and income quintiles and time period, two-way interactions of algorithm x age, algorithm x sex, algorithm x region, algorithm x time period, and three-way interactions of algorithm x age x time, algorithm x sex x time, algorithm x region x time, and algorithm x quintile x time; Partial model contains all main effects and algorithm x time; Reduced model contains main effects only. Results which are statistically significant indicate an improvement in model fit with the addition of one or more interaction terms.

Source: Manitoba Centre for Health Policy, 2006

The table reveals that LRTs for the partial and reduced models, as well as the full and reduced models were statistically significant for the combined age group, as well as for the 50+ years age-specific group. For the 12 to 19 years age group, the LRTs for the partial and reduced models were not statistically significant. The LRT for the full and reduced models for the one-year algorithms was statistically significant ($p = .0226$) but the tests for the full and reduced models for the two-year and three-year algorithms were not. For the 19 to 49 years age group, the LRTs for the partial and reduced models for the one- and two-year sets of algorithms were not statistically significant. The LRTs for all full and reduced models were statistically significant for this age group. Further investigation of the interaction effects in the full models revealed that the algorithm x region effect was often statistically significant ($p < .05$) but other interactions were not. These results indicate that the RR of asthma for different algorithms varied over time as well as with the sociodemographic characteristics of the population. In particular, the regression models suggest that the RR of asthma prevalence for different algorithms varies across regions of the province.

4.5 Chapter Summary

The validation results indicate that administrative data exhibit fair to good agreement with survey data for identifying cases of asthma.

The validation results indicate that administrative data exhibit fair to good agreement with survey data for identifying cases of asthma. The highest agreement was observed for the algorithm based on one or more hospital separations or two or more physician billing claims or two or more prescription drug records. Agreement was shown to vary with age, the presence of comorbid conditions, and income level, but not with sex or region of residence.

A sensitive algorithm for identifying asthma cases from administrative data can be obtained by using prescription drug records alone, or in combination with hospital separations and physician billing claims. There is substantial advantage that can be gained from using multiple years of administrative data to obtain a valid algorithm. There is very little trade-off between sensitivity and specificity for different algorithms; the latter was very high for all of the algorithms that were investigated. The analysis revealed very similar estimates of sensitivity and specificity for the 12 to 18 years, 19 to 49 years, and 50+ years age groups, indicating that the algorithms have similar validity for the identification of asthma cases in different age groups. There was a slight loss in sensitivity when self-reported cases of COPD and emphysema were excluded from the validation cohort, but the differences were very small.

The inferential analyses of the cross-sectional data revealed that prevalence estimates did not co-vary with the sociodemographic characteristics of the population. Moreover, the estimates for the two most sensitive algorithms in a single year of data were not significantly different from one another. However, the algorithm based on one or more contacts in hospital separations or physician billing claims or prescription drug records in one year of data produced estimates which were significantly different from the prevalence estimates based on the corresponding algorithms in two, three, and five years of data. Analyses of the longitudinal estimates revealed that for many of the age groups, there were significant differences in the estimates of asthma prevalence over time from different algorithms. As well, these longitudinal differences in prevalence often varied with the geographic region of residence of the population.

CHAPTER 5: CORONARY HEART DISEASE

5.1 Introduction and Review of Literature

Heart disease is the leading cause of death in both men and women in Canada, although it has been declining for several decades.

This chapter focusses on the identification of cases of non-fatal ischemic heart disease, commonly referred to as coronary heart disease (CHD) from administrative data. Heart disease is the leading cause of death in both men and women in Canada, although it has been declining for several decades. This finding may be the result of reduced incidence, increased survival, or a combination of the two factors. Vital statistics data have been used extensively to produce valid estimates of CHD mortality. However, these data do not provide information on non-fatal events nor on incidence. There is little published data about the prevalence of non-fatal CHD in the Canadian population. NPHS data from 1996/97 produced an estimate for 4% for both men and women 35 years of age and older (Heart and Stroke Foundation of Canada, 2000).

Table A.7 in Appendix A summarizes five studies that used administrative data to identify CHD cases. The literature review excluded studies that focused narrowly on only one form of CHD, such as acute myocardial infarction (AMI) or angina. Almost all of the studies used ICD-9-CM codes 410 to 414 to identify heart disease cases from administrative data, although Shah et al. (2000) excluded 412 (history of myocardial infarction). The ICES (1999) report also excluded ICD-9-CM 412. The retrieved studies primarily used hospital separations to identify CHD cases, although O'Connor et al. (1998) used physician claims data. None of the studies used prescription drug data to identify CHD cases from administrative data.

5.2 Description of Heart Disease Algorithms

Table 20 lists the algorithms that were selected for the validation study. All of the algorithms used ICD-9-CM codes 410 to 414 to identify heart disease cases. The second-level ATC codes that were selected for the research, based on the results of a literature review and consultations with clinical experts were C01 (cardiac therapy), C07 (beta-blocking agents), C08 (calcium channel blockers), and C09 (agents acting on the renin-angiotensin system). All of the DINs associated with these ATC codes in the MCHP Master Formulary were included in the analysis.

The algorithms investigated in this report are based on one, two, three, or five years of data. Two algorithms in each set are based only on the physician data, two are based on either hospital or physician data, and a single algorithm used all three data sources. None of the algorithms relied exclusively on prescription drug data for identifying disease cases; prescription drug records had to appear in combination with one physician billing claim for

an individual to be identified as a CHD case. This is because several of the drugs selected for this research are not used exclusively in the treatment or management of CHD. An algorithm that relied only on the prescription drug data was expected to have low specificity. To verify this, we examined six algorithms, which were based on: (1) one or more hospital separations or one or more physician billing claims or one or more prescription drug records (i.e., 1+ H or 1+ P or 1+ Rx), and (2) one or more hospital separations or one or more physician billing claims or two or more prescription drug records (i.e., 1+ H or 1+ P or 2+ Rx) in one, two, and three years of administrative data. The validation indices for these additional algorithms are reported in Appendix G. Analysis revealed that while sensitivity increased substantially (i.e., it ranged between 78.7% and 87.1%), specificity decreased and the PPV of a CHD diagnosis dropped substantially, and rarely exceeded 25%. As well, the prevalence of non-fatal heart disease was substantially higher for these algorithms than for the remaining algorithms—it was as high as 21.6% for the Manitoba population 19 years of age and older.

Table 20: Heart disease algorithms selected for validation

# Years	Algorithm	Hospital Separations or	Physician Claims or	Physician Claims and Prescription Drug Records
1	1		1 or more	
	2		2 or more	
	3	1 or more	1 or more	
	4	1 or more	2 or more	
	5	1 or more	2 or more	1 and 2 or more
2	6		1 or more	
	7		2 or more	
	8	1 or more	1 or more	
	9	1 or more	2 or more	
	10	1 or more	2 or more	1 and 2 or more
3	11		1 or more	
	12		2 or more	
	13	1 or more	1 or more	
	14	1 or more	2 or more	
	15	1 or more	2 or more	1 and 2 or more
5	16		1 or more	
	17		2 or more	
	18	1 or more	1 or more	
	19	1 or more	2 or more	
	20	1 or more	2 or more	1 and 2 or more

Source: Manitoba Centre for Health Policy, 2006

5.3 Validation Results

Table 21 contains the point estimates for the six validation indices for the 20 algorithms that were investigated for CHD. The 95% CIs for each of these estimates are reported in Appendix D, in Table D.8.

It is important to note at the outset, that individuals who indicated that they had CHF were excluded from the CCHS validation data (N = 19). We did not include the ICD-9-CM code for CHF in our algorithms (i.e., ICD-9-CM 428). Furthermore, and as expected, preliminary analyses revealed that estimates of sensitivity were improved when these individuals were excluded from the data. CCHS respondents who indicated that they had been diagnosed with heart disease were also asked to indicate whether they had angina or had previously had a heart attack. For the present report, we did not conduct separate validations for these specific forms of CHD (N = 178 for history of heart attack; N = 130 for angina), although that work could be undertaken in a subsequent study.

As Table 21 reveals, there was fair to moderate agreement between survey and administrative data, with values of κ ranging from 0.37 to 0.55. The highest value of the κ statistic was observed for several algorithms, including the three-year algorithm based on one or more hospital separations or two or more physician billing claims or one physician billing claim and two or more prescription drug records. It was equally high for the same algorithm based on five years of data, as well as for the five-year algorithm based on one or more hospital separations or two or more prescription drug records.

Sensitivity of the algorithms ranged from 28.6% to 67.9%. It was highest for the algorithm based on one or more hospital separations or one or more physician billing claims in five years of data. While sensitivity improved when five years of data were used instead of three years of data, the increase was modest (i.e., less than 7%) for all of the algorithms.

Specificity was very high (i.e., above 96%) for all of the investigated algorithms. Youden's index ranged from 0.27 to 0.63. The highest value was for the algorithm based on one or more hospital separations or one or more physician billing claims in five years. However, the algorithm based on one or more hospital separations or two or more physician billing claims or one physician billing claim and two or more prescription drug records resulted in almost an equivalent value of this summary index (0.62).

The PPV of a heart disease diagnosis ranged from 54.8% to 62.4% and was highest for the algorithm based on two or more physician billing claims in a single year of data. The NPV of a heart disease diagnosis was greater than 95% for all of the investigated algorithms.

Table 21: Estimates of agreement, sensitivity, specificity, and predictive values for heart disease algorithms

# Years		Algorithm	κ	Sens (%)	Spec (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.44	40.2	97.9	0.38	57.1	95.8
	2	2+ P	0.37	28.6	98.8	0.27	62.4	95.1
	3	1+ H or 1+ P	0.46	42.9	97.7	0.41	57.0	96.0
	4	1+ H or 2+ P	0.40	32.3	98.6	0.31	61.5	95.3
	5	1+ H or 2+ P or (1 P & 2+ Rx)	0.46	41.0	98.1	0.31	60.3	95.9
2	6	1+ P	0.51	52.8	96.9	0.50	55.1	96.7
	7	2+ P	0.47	41.8	98.1	0.40	60.5	95.9
	8	1+ H or 1+ P	0.52	55.5	96.7	0.52	54.8	96.8
	9	1+ H or 2+ P	0.49	45.6	97.8	0.43	59.7	96.2
	10	1+ H or 2+ P or (1 P & 2+ Rx)	0.52	53.1	97.3	0.50	57.9	96.7
3	11	1+ P	0.52	58.5	96.2	0.55	52.4	97.0
	12	2+ P	0.50	48.2	97.7	0.46	59.5	96.4
	13	1+ H or 1+ P	0.53	61.5	96.0	0.57	52.4	97.2
	14	1+ H or 2+ P	0.53	53.4	97.4	0.51	59.1	96.7
	15	1+ H or 2+ P or (1 P & 2+ Rx)	0.55	60.1	96.6	0.57	55.9	97.1
5	16	1+ P	0.52	65.2	95.2	0.60	49.2	97.5
	17	2+ P	0.54	56.6	96.9	0.54	56.8	96.9
	18	1+ H or 1+ P	0.53	67.9	95.0	0.63	49.0	97.6
	19	1+ H or 2+ P	0.55	60.4	96.6	0.57	55.9	97.2
	20	1+ H or 2+ P or (1 P & 2+ Rx)	0.55	66.6	95.7	0.62	52.6	97.6

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

In addition to undertaking the validation for the entire set of ICD-9-CM codes from 410 to 414, we conducted a separate validation for each of the individual codes in this set. The values of κ for this sub-analysis revealed that using algorithm #15, there was moderate agreement between survey and administrative data for 414 (other forms of chronic ischemic heart disease; κ = 0.51; 95% CI: 0.46 – 0.55), and 413 (angina pectoris; κ = 0.37; 95% CI: 0.32 – 0.42). However, for both 410 (acute myocardial infarction; κ = 0.16; 95% CI: 0.12 – 0.21) and 412 (history of myocardial infarction; κ = 0.08; 95% CI: 0.05 – 0.12) there was very low agreement. For 411 (other acute and subacute forms of ischemic heart disease) agreement was fair (κ = 0.25; 95% CI: 0.20 – 0.31).

Agreement Between Survey and Administrative Data

Logistic regression analysis was used to test the sociodemographic variables associated with agreement between survey and administrative data for coronary heart disease. The model contained the main effects of age, sex, region of residence, income adequacy quintile, and the presence of comorbid conditions, including diabetes and hypertension. The model was applied to the data for algorithm #20. The Hosmer-Lemeshow test indicated that the model fit the data well.

The following variables were statistically significant in the logistic regression model: age (χ^2 = 238.0, df = 5, p < .0001), sex (χ^2 = 4.2, df = 1, p = .0394), and presence of comorbid conditions (χ^2 = 27.9, df = 1, p < .0001). Table 22 contains the ORs and 95% CIs for the explanatory variables in this model. The odds of agreement between survey and administrative data were higher for individuals in younger age groups than for those in the oldest age group, lower for males than for females, and higher for individuals who did not have the comorbid conditions of diabetes and hypertension than for individuals who did have these comorbid conditions.

Table 22: Odds Ratio (OR) estimates and 95% CIs for predictors of agreement between administrative and survey data for heart disease

Predictors	OR	95% CI
Age		
19 – 54 years	25.9	(16.4, 40.9)
55 – 64 years	6.4	(4.1, 10.1)
65 – 74 years	2.7	(1.8, 3.9)
75 – 84 years	2.6	(1.7, 4.1)
85+ years	Ref	--
Sex		
Males	0.8	(0.6, 1.0)
Females	Ref	--
Region of Residence		
North Rural RHAs	1.2	(0.8, 1.9)
South Rural RHAs	1.2	(0.9, 1.6)
Winnipeg RHA	Ref	--
Comorbidity		
Absent	1.9	(1.5, 2.4)
Present	Ref	--
Income Quintile		
Lowest	0.9	(0.4, 1.8)
Low Middle	1.2	(0.7, 2.0)
Middle	0.9	(0.6, 1.4)
Upper Middle	0.9	(0.6, 1.4)
Not Stated	1.1	(0.6, 1.9)
Highest	Ref	--

Source: Manitoba Centre for Health Policy, 2006

5.4 Provincial Prevalence Estimates

Cross-Sectional Prevalence Estimates

Cross-sectional prevalence estimates for the 20 algorithms are reported in Table 23. Estimates ranged from 2.4% to 7.9%. The algorithm that had the highest value of the κ statistic (i.e., algorithm #15) resulted in a prevalence estimate of 5.8% for the Manitoba population 19 years of age and older. The algorithm with the highest value of Youden's index (i.e., algorithm #18) resulted in a crude prevalence estimate of 7.9%.

Table 23: Crude provincial prevalence estimates for heart disease algorithms, 1998/99 – 2002/03

# Years		Algorithm	Prevalence Estimate (%)
1	1	1+ P	3.5
	2	2+ P	2.4
	3	1+ H or 1+ P	3.7
	4	1+ H or 2+ P	2.7
	5	1+ H or 2+ P or (1 P & 2+ Rx)	3.5
2	6	1+ P	4.9
	7	2+ P	3.6
	8	1+ H or 1+ P	5.2
	9	1+ H or 2+ P	4.0
	10	1+ H or 2+ P or (1 P & 2+ Rx)	4.8
3	11	1+ P	6.0
	12	2+ P	4.5
	13	1+ H or 1+ P	6.3
	14	1+ H or 2+ P	4.9
	15	1+ H or 2+ P or (1 P & 2+ Rx)	5.8
5	16	1+ P	7.6
	17	2+ P	5.8
	18	1+ H or 1+ P	7.9
	19	1+ H or 2+ P	6.2
	20	1+ H or 2+ P or (1 P & 2+ Rx)	7.2

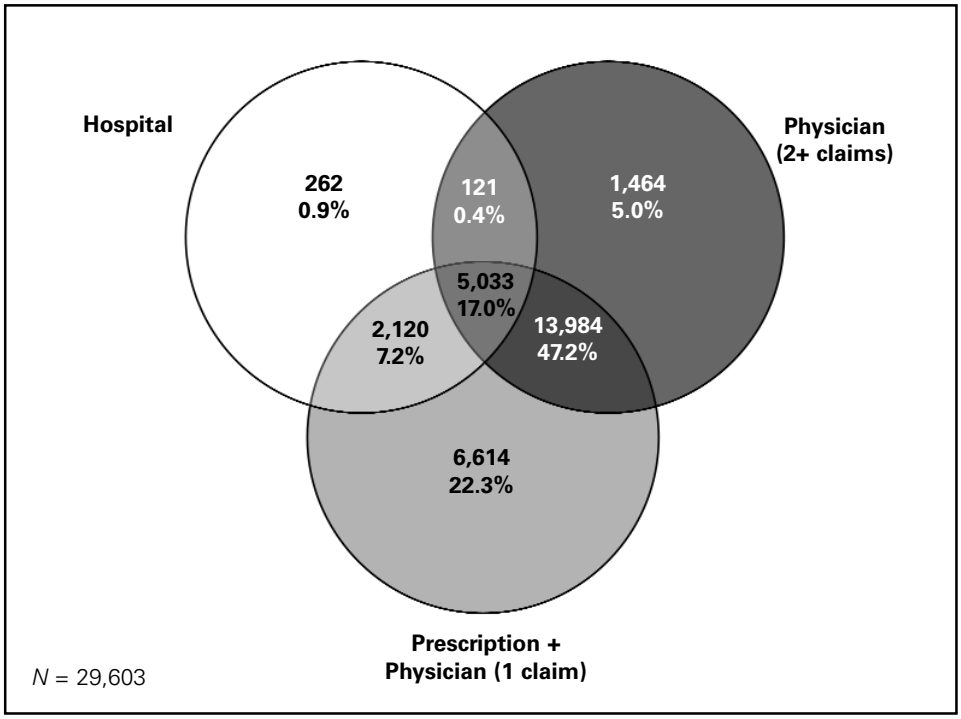
Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 1-year estimates are for 2002/03, 2-year estimates are for 2001/02 – 2002/03, 3-year estimates are for 2000/01 – 2002/03, 5-year estimates are for 1998/99 – 2002/03.

Source: Manitoba Centre for Health Policy, 2006

Venn Diagrams

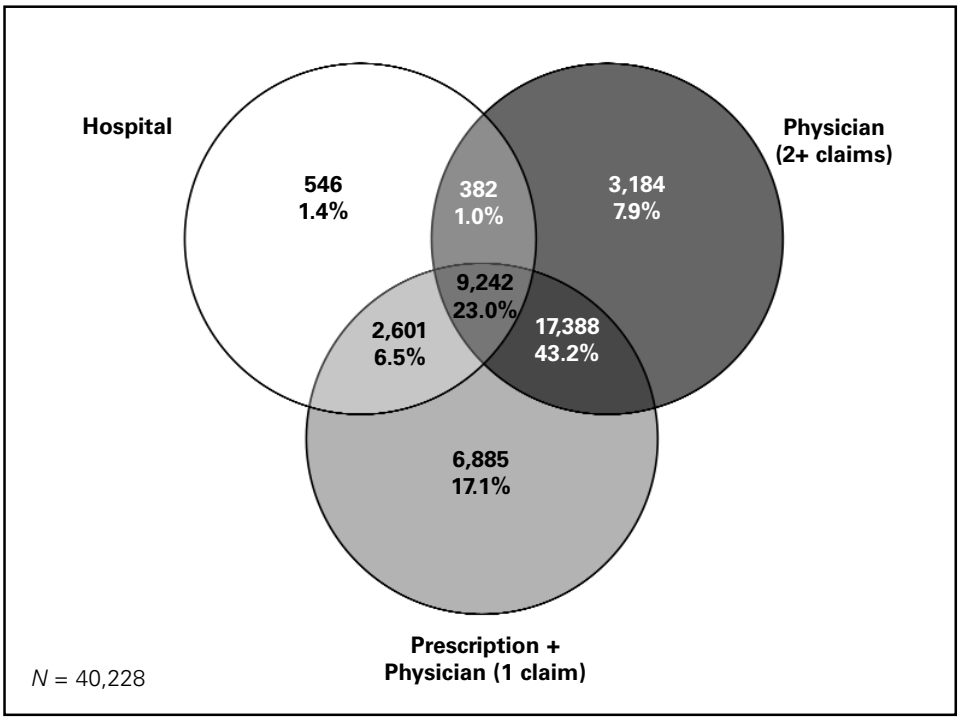
Venn diagrams for algorithms #5, #10, #15, and #20 are provided in Figures 31 to 34, respectively. Figure 31 shows that using a single year of data, almost 30,000 Manitoba residents 19 years of age and older were identified as non-fatal CHD cases. Only 5.0% were identified as having two or more physician billing claims, while 17.0% were identified in all three administrative data sources. Almost one-half (47.2%) of CHD cases were identified in both the physician and prescription drug data.

Figure 31: Coronary Heart Disease Algorithm #5: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 1 Year



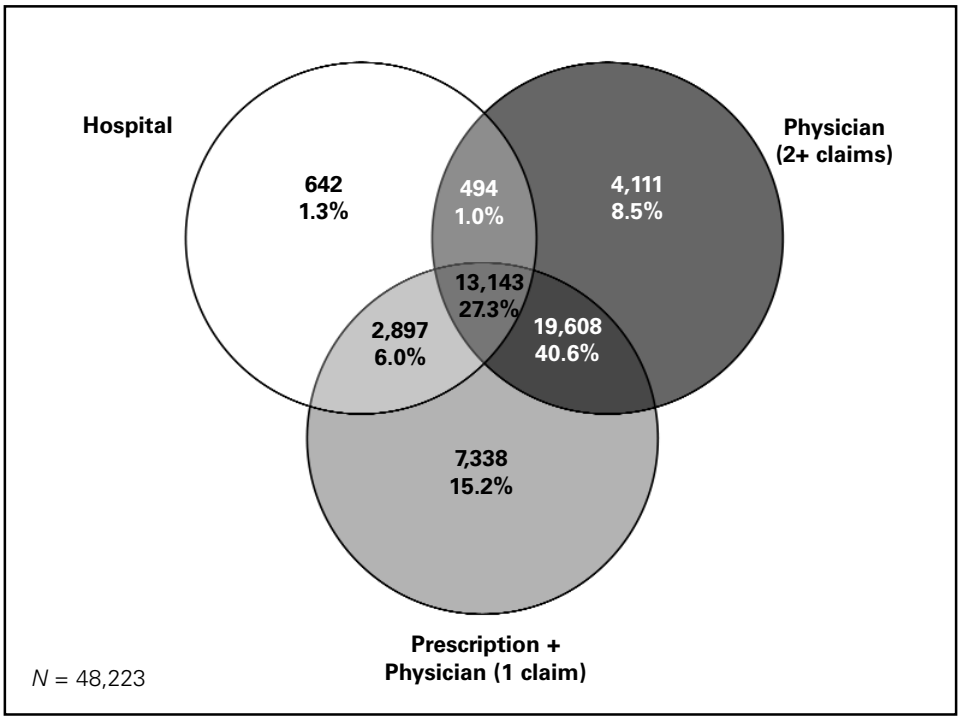
Source: Manitoba Centre for Health Policy, 2006

Figure 32: Coronary Heart Disease Algorithm #10: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 2 Years



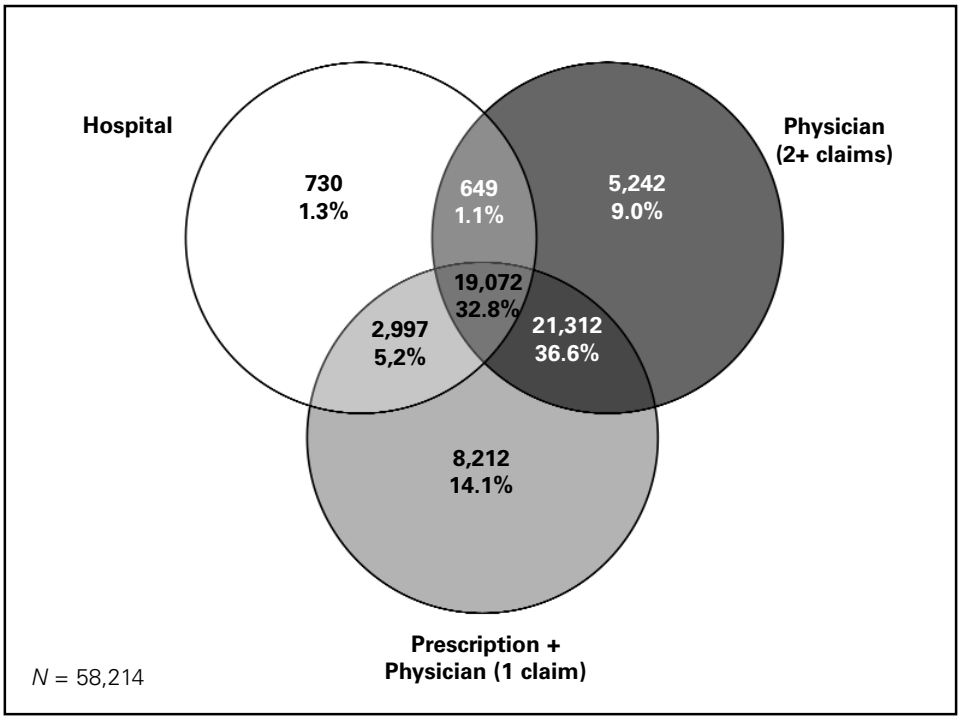
Source: Manitoba Centre for Health Policy, 2006

Figure 33: Coronary Heart Disease Algorithm #15: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 3 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 34: Coronary Heart Disease Algorithm #20: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 5 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 32 describes case counts for algorithm #10. Almost one-quarter (23%) of cases were identified in all three administrative data sources. More than 15% of individuals had one physician billing claim with a heart disease diagnosis in addition to two or more prescription drug claims with a relevant medication code. For algorithm #15, which shows the corresponding results for three years of data, the number of individuals with at least one contact in all three administrative data sources increased to 27.3%, and for five years of data this figure increased to 32.8%

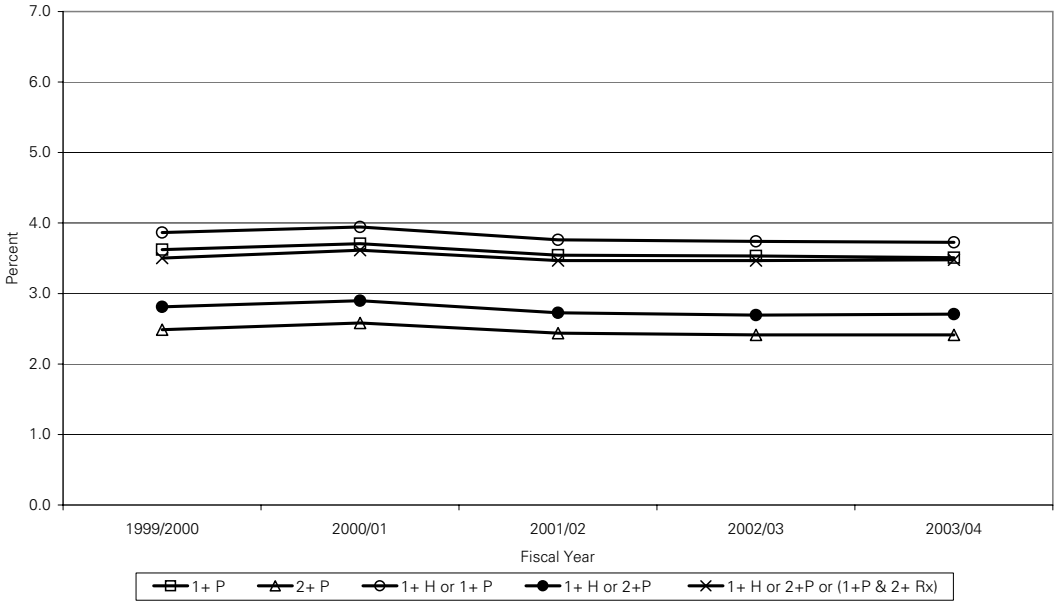
Regression Analyses for Cross-Sectional Prevalence Estimates

Regression analyses for the cross-sectional prevalence estimates were conducted for algorithms #3, #8, #10, #13, and #15. The LRT for the full and reduced models was statistically significant ($\chi^2 = 74.7$; df = 48; $p = .008$). Both the algorithm x age and algorithm x sex interactions were significant ($p < .0001$). This result indicates that the RR of heart disease for different algorithms varied by both the age and sex of the population.

Longitudinal Prevalence Estimates

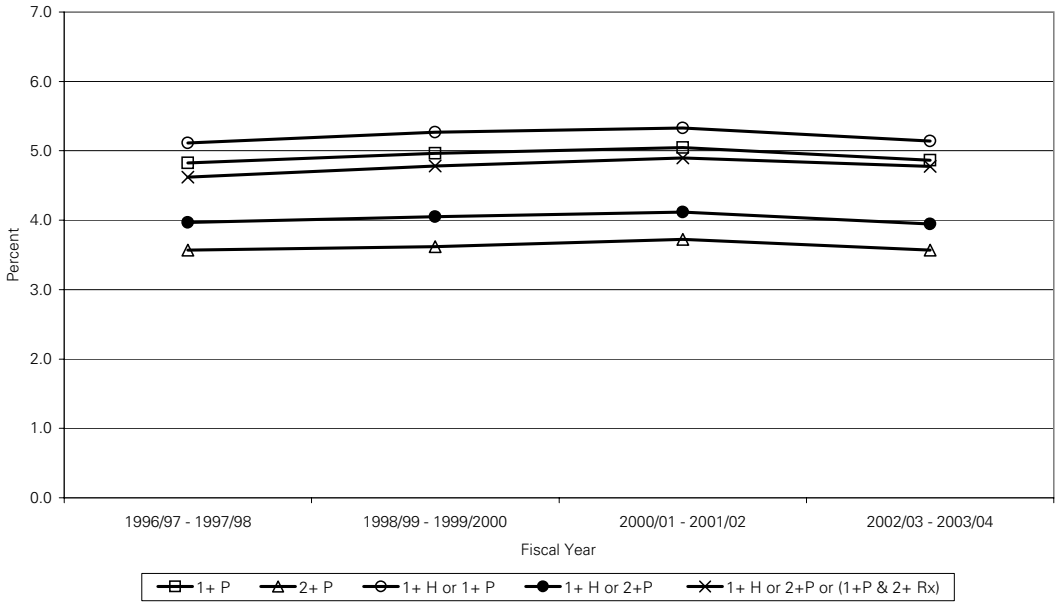
Estimates of the trend in the crude prevalence of non-fatal CHD are shown in Figures 35 to 37, for the algorithms based on one, two, and three years of administrative data, respectively. The trends for the one-year algorithms suggest that the crude prevalence dropped between 1999/2000 and 2001/02, and then remained relatively stable. This same pattern is evident in the graph of the trends for the two-year algorithms and three-year algorithms. In each figure, the trend lines were roughly parallel, indicating that each algorithm provided a similar picture of the trend in prevalence over time.

Figure 35: Provincial Trends in Prevalence of Coronary Heart Disease for One-Year Algorithms, 1999/2000 – 2003/04



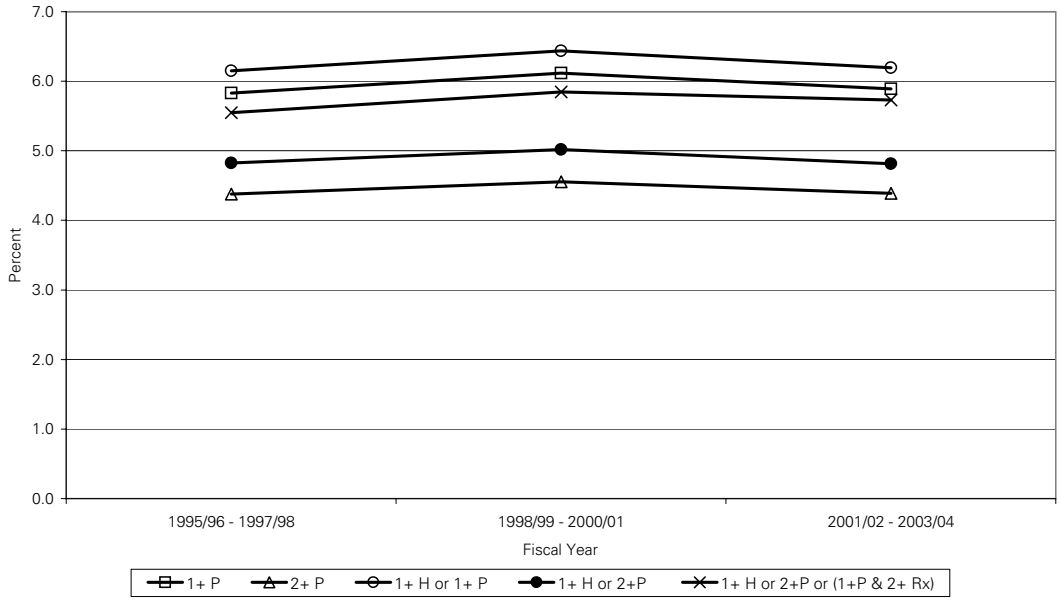
Source: Manitoba Centre for Health Policy, 2006

Figure 36: Provincial Trends in Prevalence of Coronary Heart Disease for Two-Year Algorithms, 1996/97 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Figure 37: Provincial Trends in Prevalence of Coronary Heart Disease for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Regression Analyses for Longitudinal Prevalence Estimates

Regression analyses were conducted for these longitudinal prevalence estimates, to test for differences among the algorithms in the RR over time, and also to test whether the trends for the algorithms varied across the sociodemographic variables. As in previous analyses, three regression models were defined, for the one-, two-, and three-year sets of algorithms.

For the one-year algorithms, the LRT for the partial and reduced models was not statistically significant (χ^2 = 3.0; df = 4; p = .5608), which indicates that the RR of heart disease over time did not vary for the one-year algorithms. However, the LRT for the full and reduced models was statistically significant (χ^2 = 419.0; df = 112; p < .0001). The GEE results for the full model revealed that the following two-way and three-way interactions were statistically significant: algorithm x sex (p = .0041), algorithm x region (p = .0303), algorithm x time x sex (p < .0001), algorithm x time x quintile (p = .0308), and algorithm x time x region (p = .0031). These results suggest that while the RR of heart disease does not vary over time for different algorithms, the RR of heart disease for the different algorithms varied with the demographic, socioeconomic, and geographic characteristics of the population.

The results for the two-year algorithms were similar. They showed that the LRT for the partial and reduced models was not statistically significant (χ^2 = 3.7; df = 4; p = .4504), but the LRT for the full and reduced models was statistically significant (χ^2 = 403.3; df = 112; p < .0001). Several two- and three-way interactions were significant in the GEE model results for the full model, including algorithm x age (p = .0194), algorithm x sex (p = .0007), algorithm x age x time (p < .0001), algorithm x quintile x time (p < .0001), algorithm x sex x time (p < .0001), and algorithm x region x time (p < .0001). This demonstrates that the RR of heart disease for the two-year algorithms varied over time with the demographic, socioeconomic, and geographic characteristics of the population.

For the three-year algorithms, the LRT for the partial and reduced models was not statistically significant (χ^2 = 3.3; df = 4; p = .5082), but the LRT for the full and reduced models was statistically significant (χ^2 = 384.3; df = 112; p < .0001). In the GEE model results for the full model, the following two-way and three-way interactions were statistically significant: algorithm x age (p = .0003), algorithm x sex (p < .0001), algorithm x sex x time (p < .0001), algorithm x region x time (p < .0001), and algorithm x quintile x time (p < .0001). This finding indicates that the RR of heart disease for the three-year algorithms varied over time with the demographic, socioeconomic, and geographic characteristics of the Manitoba population.

5.5 Chapter Summary

The results reported in this chapter reveal that there was moderate agreement between survey and administrative data for identifying cases of coronary heart disease.

The results reported in this chapter reveal that there was moderate agreement between survey and administrative data for identifying cases of coronary heart disease. Agreement, sensitivity and Youden's index increased slightly with increasing numbers of years of data. Several three-year and five-year algorithms produced similar results for the validation indices; some of these were based on just hospital separations or physician billing claims, but others were based on hospital separations or physician billing claims or prescription drug records. Agreement between the administrative data and survey data was predicted by age, sex, and the presence of comorbid chronic diseases.

Crude prevalence estimates of non-fatal coronary heart disease in Manitoba ranged from 2.4% to 7.9% for the investigated algorithms. The prevalence estimates for the algorithms with the highest values of Youden's index were 5.8% and 7.9%. According to the trend estimates for the algorithms, the prevalence of non-fatal heart disease in the Manitoba population changed very little in recent years, although a slight decrease was evident for all of the investigated algorithms. The algorithms did not result in significantly different estimates of the rate of change in prevalence. Finally, the analyses indicated that the relative rate of heart disease for different algorithms varied with the sociodemographic characteristics of the Manitoba population.

Validation of algorithms for identifying cases with specific forms of non-fatal CHD, including angina and AMI, could be undertaken using the data available in the CCHS. Alternatively, the research could be broadened to investigate the validity of administrative data for identifying all cases of cardiovascular disease, including congestive heart failure and hypertension, in addition to coronary heart disease, in a single algorithm. This was the approach adopted by Maio et al. (2005) in their study of the validity of prescription drug data for identifying chronic disease cases.

CHAPTER 6: DIABETES

6.1 Introduction and Review of Literature

Prevalence estimates from the National Diabetes Surveillance System (NDSS) for 1998/99 were 4.8% for Canada and 5.1% for Manitoba.

Diabetes is a significant public health concern in part because it has the potential to lead to a number of health complications. Prevalence estimates from the National Diabetes Surveillance System (NDSS) for 1998/99 were 4.8% for Canada and 5.1% for Manitoba (Health Canada, 2002). The crude prevalence for Manitoba for the period 1998/99–2000/01 was estimated at 5.7% using administrative data (Martens et al., 2003). It is important to note that separate estimates of Type I and Type II diabetes have not been produced. Moreover, cases of gestational diabetes have not been distinguished in the computation of these estimates.

Table A.4 in Appendix A summarizes published studies that used administrative data to identify diabetes cases. The studies reported in this table do not include those summarized in the comprehensive review conducted by Saydah et al. (2004). These authors reviewed the results of 16 studies published between 1966 and mid 2002 and referenced in MEDLINE; the studies selected for review reported one or more measures of the validity of administrative data for diabetes case identification. Saydah et al. (2004) also reviewed several articles that validated death certificate and survey data for identifying cases of diabetes. The median sensitivity reported in the 16 articles that validated administrative data was 81.5%. Sensitivity ranged from 46% to 97% and specificity ranged from 95% to 100%. Kappa values ranged from 0.67 to 0.96 and PPV ranged from 60% to 98% with a median of 92%. Overall, the authors support the use of administrative data for identifying cases of diabetes, noting that "we found these data sources [administrative databases and surveys] are adequately sensitive… and highly specific" (p. 514).

The algorithms examined in the studies reported in Table 24 were primarily based on hospital (i.e., inpatient) and physician (i.e., out-patient) data, although at least three studies also used medication codes in prescription drug data to identify diabetes cases. The majority of these studies used a single diagnostic code, ICD-9-CM code 250, for case identification. One study used ICD-10 diagnostic codes. ATC codes were not specified in any of the studies which used prescription drug data. Most studies relied on a single occurrence of a diagnostic code in administrative data to define diabetes cases. However, a few tested the effect on sensitivity, specificity, and predictive values of requiring multiple occurrences of ICD-9-CM code 250 in administrative data.

6.2 Description of Diabetes Algorithms

In keeping with the vast majority of previous studies that have used administrative data to estimate the prevalence of diabetes, the current study used ICD-9-CM code 250 to define cases from Manitoba's hospital and physician data. A single second-level ATC code, A10 (drugs used in diabetes) was used to identify diabetes cases from Manitoba's prescription drug data. All of the DINs with this ATC code were selected from the MCHP Master Formulary.

Table 24 enumerates the 18 diabetes algorithms that were evaluated. These algorithms were based on one, two, or three years of administrative data. Algorithms based on five years of administrative data were not investigated in this study because they were not investigated in previous research. All algorithms required at least one occurrence of a diagnostic code in hospital separations for an individual to be classified as a diabetes case. However, the algorithms varied in the number of occurrences of a diagnostic code in the physician claims and the number of occurrences of a medication code in prescription drug data for an individual to be classified as a diabetes case. The selection of these algorithms for testing and validation was based on the work of Blanchard et al. (1996), Hux et al. (2002), Martens et al. (2003), and Robinson et al. (1997).

Table 24: Diabetes algorithms selected for testing and validation

Years	Algorithm #	Hospital Separations or	Physician Claims or	Prescription Drug Records
1	1	1 or more	1 or more	
	2	1 or more	2 or more	
	3	1 or more	1 or more	1 or more
	4	1 or more	2 or more	1 or more
	5	1 or more	1 or more	2 or more
	6	1 or more	2 or more	2 or more
2	7	1 or more	1 or more	
	8	1 or more	2 or more	
	9	1 or more	1 or more	1 or more
	10	1 or more	2 or more	1 or more
	11	1 or more	1 or more	2 or more
	12	1 or more	2 or more	2 or more
3	13	1 or more	1 or more	
	14	1 or more	2 or more	
	15	1 or more	1 or more	1 or more
	16	1 or more	2 or more	1 or more
	17	1 or more	1 or more	2 or more
	18	1 or more	2 or more	2 or more

Source: Manitoba Centre for Health Policy, 2006

6.3 Validation Results

Table 25 contains the results for the six validation indices for the 18 diabetes algorithms. The 95% CIs for these estimates are reported in Appendix D in Table D.9.

Overall agreement between administrative and survey data was very good to excellent, with values of κ ranging from 0.73 to 0.86. The highest estimate was obtained for two of the two-year algorithms: (a) one or more hospital separations or two or more physician billing claims or one or more prescription drug records, and (b) one or more hospital separations or two or more physician billing claims or two or more prescription drug records. However, several other two- and three-year algorithms produced estimate of κ that were higher than 0.80.

Sensitivity was very high for all of the algorithms, and ranged from 85.5% to 90.5%. There was little or no variation in sensitivity for the same set of diagnostic/medication codes applied to one, two, or three years of administrative data. For example, the algorithm based on one or more hospital separations or two or more physician billing claims had an estimated sensitivity of 85.8% in one year of data, and 86.9% in two or three years of data. Overall, the algorithms with the highest sensitivity were based on having one or more contacts in hospital, or physician, or prescription drug databases in one, two, or three years of data.

Specificity was very high and also showed little variation across the algorithms. It ranged from 97.3% to 99.5%. Given the high and limited range of values for both sensitivity and specificity, Youden's index was also high and ranged, from 0.63 to 0.88. It was highest for the following three algorithms: (a) one or more hospital separations or one or more physician billing claims or one or more prescription drug records in two years, (b) one or more hospital separations or one or more physician billing claims or one or more prescription drug records in three years, and (c) one or more hospital separations or two or more physician billing claims or two or more prescription drug records in three years.

The PPV of a diabetes diagnosis showed a slightly greater range than sensitivity and specificity, from 68.2% to 89.5%. The maximum estimate was achieved for the algorithm based on one or more hospital separations or two or more physician billing claims in one year of data. NPV approached its upper bound for all algorithms; it attained values as high as 99.4%.

Table 25: Estimates of agreement, sensitivity, specificity, and predictive values for diabetes algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ H or 1+ P	0.77	76.9	98.7	0.76	79.2	98.5
	2	1+ H or 2+ P	0.73	63.2	99.5	0.63	89.5	97.7
	3	1+ H or 1+ P or 1+ Rx	0.81	85.8	98.6	0.84	79.4	99.1
	4	1+ H or 2+ P or 1+ Rx	0.84	80.7	99.4	0.80	88.9	98.8
	5	1+ H or 1+ P or 2+ Rx	0.81	85.5	98.6	0.84	79.8	99.1
	6	1+ H or 2+ P or 2+ Rx	0.84	80.1	99.4	0.80	89.4	98.7
2	7	1+ H or 1+ P	0.78	85.2	98.1	0.83	74.0	99.0
	8	1+ H or 2+ P	0.82	79.5	99.3	0.79	87.9	98.7
	9	1+ H or 1+ P or 1+ Rx	0.80	89.6	97.9	0.88	73.7	99.3
	10	1+ H or 2+ P or 1+ Rx	0.86	86.6	99.1	0.86	86.1	99.1
	11	1+ H or 1+ P or 2+ Rx	0.80	89.3	98.0	0.87	74.0	99.3
	12	1+ H or 2+ P or 2+ Rx	0.86	86.1	99.2	0.85	86.8	99.1
3	13	1+ H or 1+ P	0.75	87.8	97.4	0.84	68.7	99.2
	14	1+ H or 2+ P	0.83	84.9	99.0	0.84	83.9	99.0
	15	1+ H or 1+ P or 1+ Rx	0.76	90.5	97.3	0.88	68.2	99.4
	16	1+ H or 2+ P or 1+ Rx	0.84	88.4	98.8	0.87	82.1	99.3
	17	1+ H or 1+ P or 2+ Rx	0.76	90.2	97.4	0.88	68.6	99.4
	18	1+ H or 2+ P or 2+ Rx	0.85	88.1	98.8	0.87	83.0	99.2

Note: H = Hospital Separation; P = Physician Claim; Rx = Prescription Drug Claim; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

Agreement Between Survey and Administrative Data

The logistic regression models to test the sociodemographic variables associated with agreement between survey and administrative data contained the main effects of age, sex, region of residence, income adequacy quintile, and comorbidity (i.e., presence of heart disease or hypertension). The data for algorithm #10, which was based on one or more hospital separations or two or more physician billing claims or two or more prescription drug records was selected for this analysis. Two-way interactions were tested but were not statistically significant and were therefore excluded from the final model. The Hosmer-Lemeshow test indicated that the model fit the data well.

The following variables were statistically significant predictors of agreement between the two data sources: age (χ^2 = 46.5, df = 5, p < .0001) and presence of comorbid conditions (χ^2 = 4.2, df = 1, p = .0406). Table 26 contains

the ORs and 95% CIs for the explanatory variables in the model. The results indicate that the odds of agreement between the two data sources was higher for individuals in younger age groups than in the oldest age group, and was higher for individuals with no comorbid conditions than for those with a comorbid condition.

Table 26: Odds Ratio (OR) estimates and 95% CIs for predictors of agreement between administrative and survey data for diabetes

Predictors	OR	95% CI
Age		
19 – 44 years	5.9	(2.7, 13.1)
45 – 54 years	4.1	(1.7, 9.5)
55 – 64 years	1.3	(0.6, 2.8)
65 – 74 years	1.2	(0.6, 2.5)
75 – 84 years	1.2	(0.6, 2.4)
85+ years	Ref	–
Sex		
Males	1.0	(0.7, 1.3)
Females	Ref	--
Region of Residence		
North Rural RHAs	0.9	(0.5, 1.6)
South Rural RHAs	1.5	(1.0, 2.1)
Winnipeg RHA	Ref	--
Comorbidity		
Absent	1.5	(1.0, 2.2)
Present	Ref	--
Income Quintile		
Lowest	1.2	(0.5, 2.9)
Low Middle	1.2	(0.6, 2.3)
Middle	1.2	(0.7, 2.1)
Upper Middle	1.3	(0.8, 2.2)
Not Stated	3.8	(1.3, 11.3)
Highest	Ref	--

Source: Manitoba Centre for Health Policy, 2006

6.4 Provincial Prevalence Estimates

Cross-Sectional Prevalence Estimates

Table 27 summarizes the cross-sectional population-based prevalence estimates that were calculated for each of the 18 diabetes algorithms for the Manitoba population 19 years of age and older. They range from 5.8% to 8.2%. The algorithms with the highest estimates of κ (i.e., algorithms #10 and #12) resulted in prevalence estimates of 6.3% and 6.2%, respectively. The algorithm with the highest value of Youden's index and corresponding highest sensitivity (i.e., algorithm #15) resulted in a prevalence estimate of 8.2%. However, algorithm #9 had an equally high value of Youden's index, and it resulted in a prevalence estimate of 7.5%.

Table 27: Crude provincial prevalence estimates for diabetes algorithms, 2000/01 – 2002/03

# Years		Algorithm	Prevalence Estimate (%)
1	1	1+ H or 1+ P	5.8
	2	1+ H or 2+ P	4.4
	3	1+ H or 1+ P or 1+ Rx	6.5
	4	1+ H or 2+ P or 1+ Rx	5.6
	5	1+ H or 1+ P or 2+Rx	6.4
	6	1+ H or 2+ P or 2+ Rx	5.5
2	7	1+ H or 1+ P	7.1
	8	1+ H or 2+ P	5.6
	9	1+ H or 1+ P or 1+ Rx	7.5
	10	1+ H or 2+ P or 1+ Rx	6.3
	11	1+ H or 1+ P or 2+Rx	7.4
	12	1+ H or 2+ P or 2+ Rx	6.2
3	13	1+ H or 1+ P	7.9
	14	1+ H or 2+ P	6.3
	15	1+ H or 1+ P or 1+ Rx	8.2
	16	1+ H or 2+ P or 1+ Rx	6.8
	17	1+ H or 1+ P or 2+Rx	8.1
	18	1+ H or 2+ P or 2+ Rx	6.6

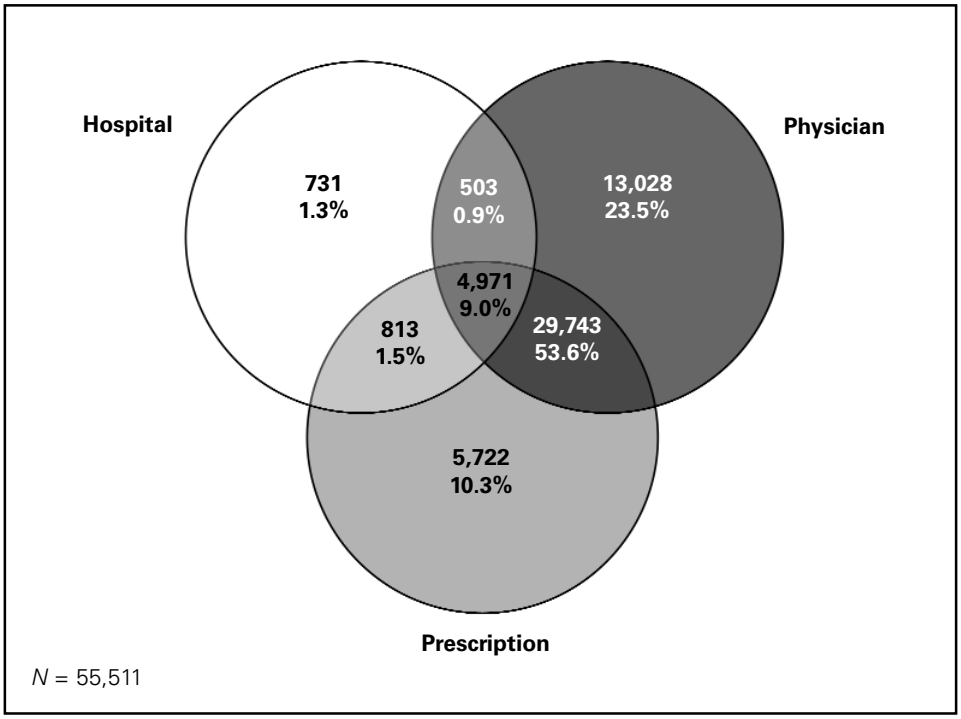
Note: H = hospital separation; P = physician claim, Rx = prescription drug record; One-year algorithms are from 2002/03, two-year algorithms are from 2001/02 – 2002/03, and three-year algorithms are from 2000/01 – 2002/03.

Source: Manitoba Centre for Health Policy, 2006

Venn Diagrams

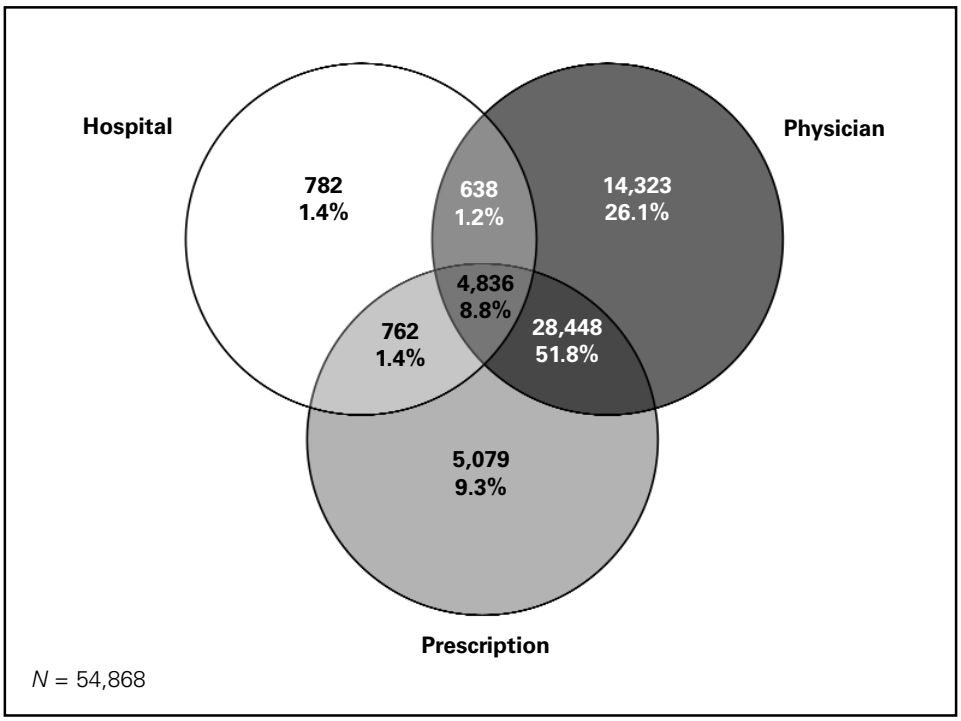
Figures 38 to 41 contain the Venn diagrams for algorithms #3, #5, #9, and #15. Algorithm #3, which was based on one or more contacts in the hospital, physician, or prescription drug data in a one-year period resulted in more than 55,000 diabetes cases. More than half of these individuals (53.6%) were identified from both the physician and prescription drug data. Ten percent of disease cases were identified solely from the prescription drug data, and 9.0% had a contact in all three data sources.

Figure 38: Diabetes Algorithm #3: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 1 Year



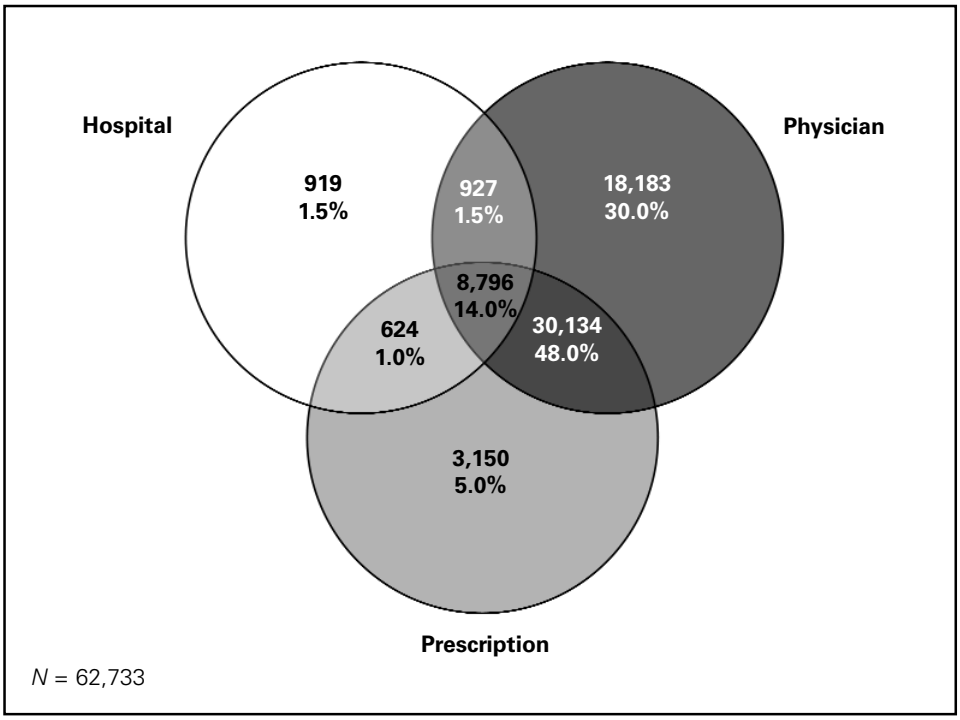
Source: Manitoba Centre for Health Policy, 2006

Figure 39: Diabetes Algorithm #5: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 1 Year



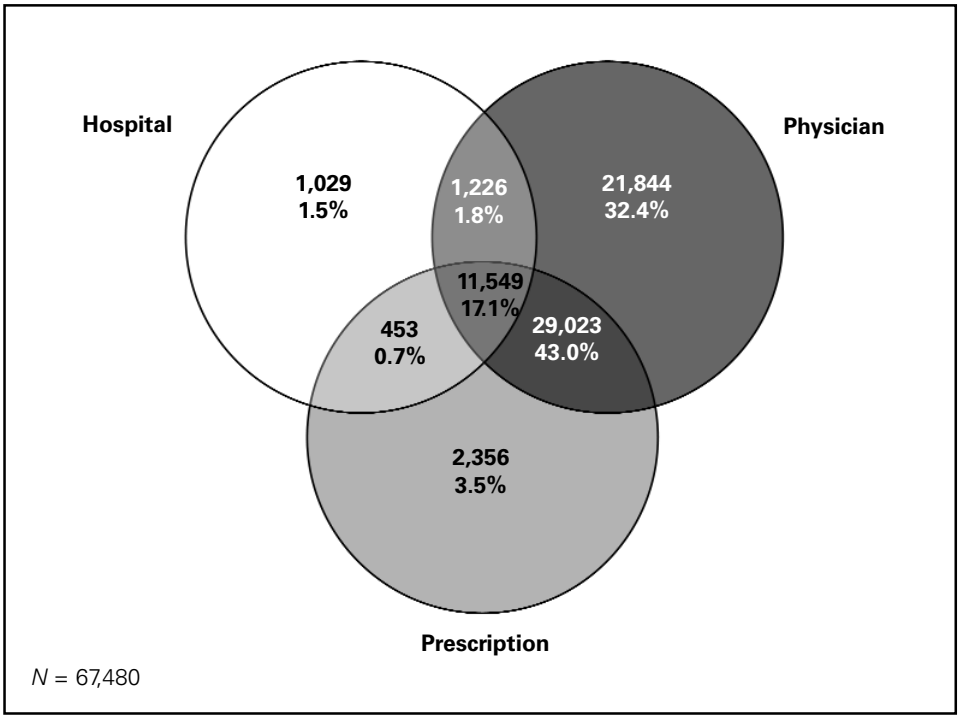
Source: Manitoba Centre for Health Policy, 2006

Figure 40: Diabetes Algorithm #9: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 2 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 41: Diabetes Algorithm #15: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 3 Years



Source: Manitoba Centre for Health Policy, 2006

Almost 63,000 individuals were identified as diabetes cases using algorithm #9 (Figure 40), which was based on two years of administrative data. Close to half (48.0%) of these individuals were identified in both the physician and prescription drug data, and 14.0% were identified in all three data sources. Using three years of data (Figure 41), the percentage of cases identified in only prescription drug data dropped substantially, to 3.5% of total cases, and the percentage of individuals identified in all three data sources increased, to 17.1%.

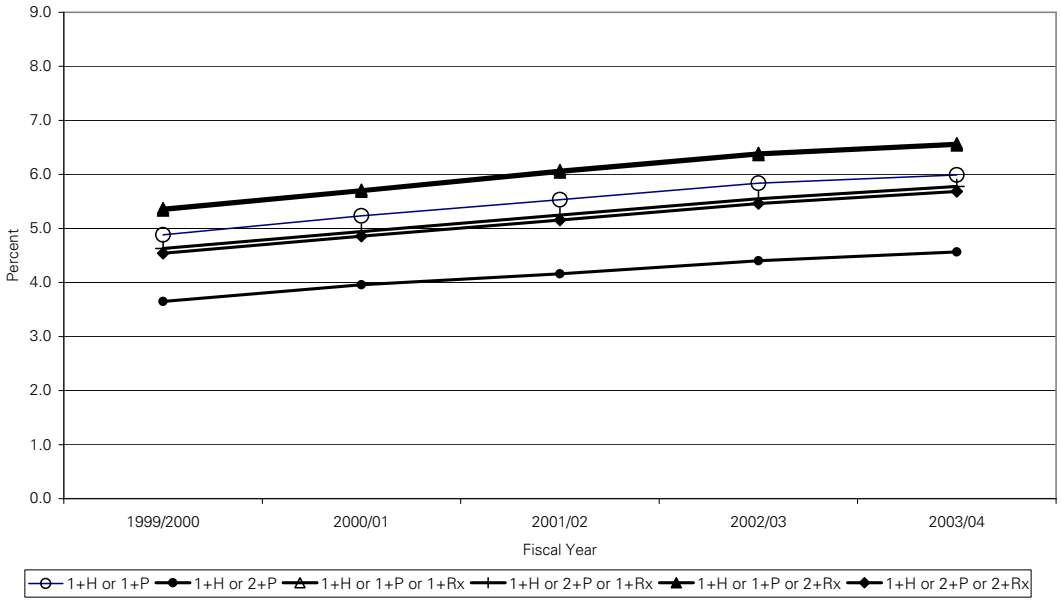
Regression Analyses for Cross-Sectional Prevalence Estimates

The regression models for the cross-sectional prevalence estimates were applied to six algorithms: #1, #5, #7, #11, #13, and #17. The LRT for the full and reduced models was non-significant ($\chi^2 = 63.2$; df = 60; $p = .3640$). These results indicate that the RR of diabetes prevalence for different algorithms did not vary across the sociodemographic characteristics of the population. Further analysis of the reduced model revealed that the algorithm main effect was statistically significant ($\chi^2 = 67.32$; df = 5; $p < .0001$) indicating that there were significant differences in the prevalence estimates for the six algorithms. In fact, further analysis revealed that all of the algorithms produced RRs which were significantly different from the estimate for algorithm #1 (reference).

Longitudinal Prevalence Estimates

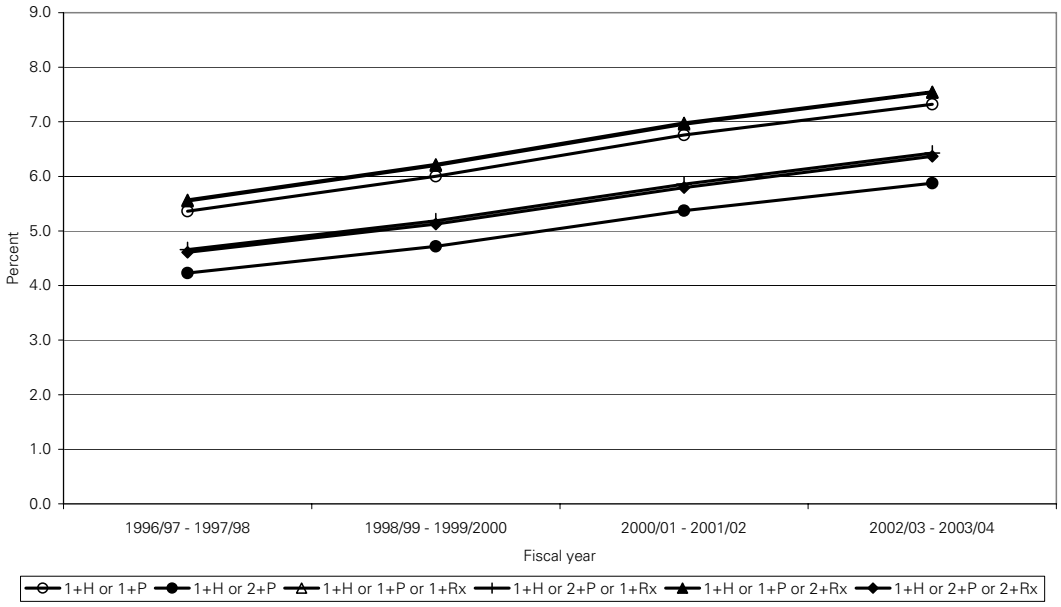
Figures 42, 43, and 44 summarize the trends in prevalence estimates that were obtained for the one-, two- and three-year algorithms, respectively. All three show a similar pattern of increasing prevalence over time. The trend lines for the six algorithms plotted in each graph are approximately parallel, indicating that all algorithms provided a similar picture of the change in prevalence over time.

Figure 42: Provincial Trends in Diabetes Prevalence for One-Year Algorithms, 1999/2000 – 2003/04



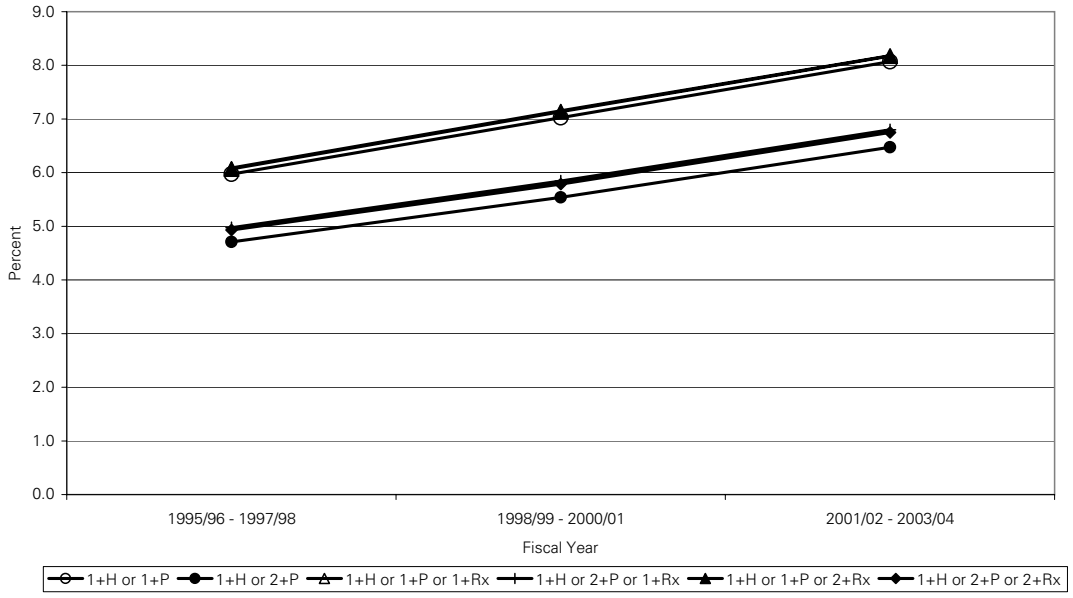
Source: Manitoba Centre for Health Policy, 2006

Figure 43: Provincial Trends in Diabetes Prevalence for Two-Year Algorithms, 1996/97 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Figure 44: Provincial Trends in Diabetes Prevalence for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Regression Analyses for Longitudinal Prevalence Estimates.
Regression analyses were conducted for the longitudinal prevalence estimates, to test for significant differences in the estimates from different algorithms over time, and also to test whether these longitudinal estimates varied across the sociodemographic variables of age, sex, region of residence, and income quintile. For the one-year algorithms, the LRT for the partial and reduced models was not statistically significant (χ^2 = 1.6; df = 5; p = .9032). The LRT for the full and reduced models was also not statistically significant (χ^2 = 127.0; df = 137; p < .7187). These results indicate that the RR of change in prevalence for different one-year algorithms did not vary over time, nor did it vary across the sociodemographic characteristics of the population. However, in the GEE analysis for the reduced model, there was both a significant time main effect (p < .0001), and an algorithm main effect (p < .0001). For the former, the RR was 1.04 (95% CI: 1.04 – 1.05). The results for the algorithm main effect revealed that the RRs for all algorithms were significantly different from the estimate for algorithm #1 (reference), with the exception of algorithm #4.

For the two-year algorithms, the LRT for the partial and reduced models was not statistically significant (χ^2 = 0.97; df = 5; p = .9653). The LRT for the full and reduced models was also not statistically significant (χ^2 = 126.6; df = 137; p = .7270). These results indicate that the RR of diabetes preva-

lence for different two-year algorithms did not vary over time or with the sociodemographic characteristics of the population. In the GEE models, the main effects of both time and algorithm were statistically significant ($p <$ .0001). The RR of change across two-year time periods was 1.09 (95% CI = 1.08 – 1.10). The prevalence estimates for all algorithm, with the exception of algorithms #9 and #10, were not significantly different from the estimate for algorithm #7 (reference).

Finally, for the three-year algorithms, the LRT for the partial and reduced models was also not statistically significant ($\chi^2 = 0.72$; df = 5; $p = .9817$). The LRT for the full and reduced models was also not statistically significant ($\chi^2 = 119.7$; df = 137; $p < .8536$). These results indicate that the RR of diabetes prevalence for different three-year algorithms did not vary over time or with the sociodemographic characteristics of the population. However, just like in the previous models, the effects of time and algorithm were both statistically significant ($p <$ .0001). The results for the time main effect revealed that the RR was 1.14 (95% CI = 1.13 – 1.14), indicating a statistically significant increase in prevalence over time. All algorithms produced prevalence estimates that were significantly higher than algorithm #13 (reference) with the exception of algorithms #15 and #17.

6.5 Chapter Summary

The results of this chapter confirm that administrative data are a valid tool for identifying diabetes cases.

The results of this chapter confirm that administrative data are a valid tool for identifying diabetes cases. There was high agreement between the two data sources. Agreement did, however, vary with age and the presence of comorbid conditions.

There is some value associated with using prescription drug data, in addition to hospital and physician data, to identify diabetes cases, although the gains were not large. Estimates of κ and sensitivity were highest for the algorithms based on all three administrative data sources, but they were not appreciably larger than the estimates for the algorithms based on hospital separations or physician billing claims. The crude estimates of prevalence for the algorithms with the highest agreement between the two data sources or the highest values of Youden's index ranged from 6.2% to 8.2%.

Further analyses revealed that the relative rate of diabetes did not vary across the sociodemographic characteristics of the population. All of the investigated algorithms showed an increasing prevalence of diabetes over time. However, different algorithms did not provide different estimates of the relative rate of increase.

CHAPTER 7: HYPERTENSION

7.1 Introduction and Review of Literature

Hypertension is a significant public health concern because it is the most common risk factor for cardiovascular disease.

Hypertension is a significant public health concern because it is the most common risk factor for cardiovascular disease. It is also a risk factor for other diseases, including diabetes, kidney failure and stroke. Martens et al. (2004) used administrative data to estimate the prevalence of hypertension as approximately 22% for Manitoba in 1998/99–2000/01. This was slightly higher than the provincial estimate of 20% for the period 1993/94 – 1995/96. The same increasing trend has been identified in U.S. survey data for the period 1991 to 2000 (Hajjar and Kotchen, 2003). Wolff et al. (1999) used 1995 survey data from Nova Scotia and estimated the crude prevalence to be 26.6%. Wolf-Maier et al. (2003) compared survey data from Canada, United States, Germany, Finland, Sweden, England, Spain, and Italy and demonstrated that the prevalence of hypertension is substantially higher in European countries than in North American countries. For the former, the age-adjusted prevalence was 44% while for the latter it was only 28%.

Table A.5 in Appendix A summarizes several studies that used administrative health data for identifying cases of hypertension. Overall, the results of the literature review suggest that administrative data are a valid tool for estimating the prevalence of hypertensive disease. Quam et al. (1993) examined the agreement between administrative data (i.e., physician claims and prescription drug records) and both patient survey and medical chart data. Cases were identified by one or more physician claims with an ICD-9-CM code of 401 (essential hypertension) or one or more prescriptions for diuretics, beta-blockers, calcium antagonists, or ACE inhibitors. Agreement between administrative and survey data was 43.4% using an algorithm based only on physician claims, 62.9% using an algorithm based only on prescription drug data, and 95.7% when both physician and prescription data were compared to survey data. The comparisons between administrative data and medical chart data produced similar levels of agreement. Muhajarine et al. (1997) also found good agreement between health survey data and physician billing claims using ICD-9-CM 401 or 402.

7.2 Description of Hypertension Algorithms

The current study used a single ICD-9-CM code, 401, to define hypertensive cases from Manitoba's hospital and physician data. The literature summarized in Table A.5 in Appendix A shows that both narrow and broad ranges of ICD-9-CM codes have been used in previous research; for example Muhajarine et al. (1997) used only 401 and 402, Rector et al. (2004) used 401 to 404, and Robinson et al. (1997) used 401 to 405 in addition to sev-

eral other ICD-9-CM codes. Our empirical examination of the administrative data for 2002/03 revealed that 99.5% of individuals who had a hospital separation or physician claim with an ICD-9-CM code in the range 401 to 405 were assigned the single ICD-9-CM code of 401. Given these results as well as the literature review results, we used this single diagnostic code to identify cases of hypertension

Based on the literature review and consultations with pharmacists and clinical experts, five second-level ATC codes were used to identify cases from Manitoba's pharmaceutical data. These were C02 (anti-hypertensives), C03 (diuretics), C07 (beta blocking agents), C08 (calcium channel blockers), and C09 (agents acting on the renin-angiotensin system). All of the DINs associated with these ATC codes were selected from the MCHP Master Formulary.

Table 28 enumerates the 18 hypertension algorithms that were evaluated in this research. These algorithms were based on one, two, or three years of administrative data. Some of the algorithms were based solely on the physician data, others relied on both the hospital and physician data, and the remainder combined all three data sources. Five years of data were not used to define disease cases because no previous studies used this many years of data to define hypertension algorithms. The algorithms varied in terms of the number of occurrences of a diagnostic code in the physician claims and the number of occurrences of a medication code in prescription drug data.

Table 28: Hypertension algorithms selected for validation

Years of Data	Algorithm #	Hospital Separations or	Physician Claims or	Prescription Drugs
1	1		1 or more	
	2		2 or more	
	3	1 or more	1 or more	
	4	1 or more	2 or more	
	5	1 or more	1 or more	1 or more
	6	1 or more	1or more	2 or more
2	7		1 or more	
	8		2 or more	
	9	1 or more	1 or more	
	10	1 or more	2 or more	
	11	1 or more	1or more	1 or more
	12	1 or more	1or more	2 or more
3	13		1 or more	
	14		2 or more	
	15	1 or more	1 or more	
	16	1 or more	2 or more	
	17	1 or more	1or more	1 or more
	18	1 or more	1or more	2 or more

Source: Manitoba Centre for Health Policy, 2006

7.3 Validation Results

Validation Indices

The results for the six validation indices for each of the 18 hypertension algorithms are reported in Table 29. The 95% CIs for these algorithms are reported in Appendix Table D.10.

There was moderate to good agreement between administrative and survey data, with values of κ ranging from 0.54 to 0.70 for the algorithms based on one year of data. For the three-year algorithms the range of values of κ was narrower, from 0.64 to 0.70. The algorithms with the highest values of κ were based on: (a) one or more hospital separations or one or more physician billing claims or two or more prescription drug records in one year, and (b) one or more hospital separations or two or more physician billing claims in three years.

Sensitivity ranged from 48.4% to 92.8%. It was highest for the algorithm based on one or more hospital separations or one or more physician billing claims or one or more prescription drug records in three years. However, sensitivity was above 90% for other two- and three-year algorithms.

Specificity ranged from 88.7% to 97.5% for the algorithms that were based on one year of data. The range of specificity values was similar for the algorithms based on two and three years of data. Overall, the lowest specificity (83.7%) was for the algorithm based on one or more contacts in hospital separations or physician billing claims or prescription drug records in three years of data.

Youden's index ranged from 0.46 to 0.79. Two algorithms resulted in the maximum value of this index: (a) one or more hospital separations or one or more physician billing claims or one or more prescription drug records in one year, and (b) one or more hospital separations or one or more physician billing claims or two or more prescription drug records in one year.

The PPV of a hypertension diagnosis ranged from 56.8% to 81.4%. The algorithm for which it was highest was algorithm #2, which was based on two or more physician claims in a single year of data. The NPV of a hypertension diagnosis was always very high, and the estimate only fell below 90% for algorithm #4 (i.e., 1+ H or 2+ P).

Table 29: Estimates of agreement, sensitivity, specificity, and predictive values for hypertension algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.65	68.4	94.7	0.63	74.6	93.0
	2	2+ P	0.54	48.4	97.5	0.46	81.4	89.3
	3	1+ H or 1+ P	0.66	70.1	94.4	0.65	73.9	93.3
	4	1+ H or 2+ P	0.56	51.1	97.1	0.48	79.9	89.7
	5	1+ H or 1+ P or 1+ Rx	0.68	89.9	88.7	0.79	64.4	97.5
	6	1+ H or 1+ P or 2+ Rx	0.70	89.0	89.9	0.79	66.6	97.3
2	7	1+ P	0.67	79.4	91.9	0.71	69.1	95.2
	8	2+ P	0.66	66.3	95.6	0.62	77.5	92.6
	9	1+ H or 1+ P	0.68	81.2	91.6	0.73	68.8	95.6
	10	1+ H or 2+ P	0.67	69.4	95.2	0.65	76.8	93.2
	11	1+ H or 1+ P or 1+ Rx	0.64	91.9	86.0	0.78	59.8	97.9
	12	1+H or 1+ P or 2+ Rx	0.66	91.2	87.3	0.78	62.0	97.8
3	13	1+ P	0.67	83.2	90.3	0.74	66.0	95.9
	14	2+ P	0.68	72.4	94.8	0.68	76.0	93.8
	15	1+ H or 1+ P	0.67	84.9	89.9	0.75	65.7	96.3
	16	1+ H or 2+ P	0.70	75.6	94.4	0.71	75.2	94.5
	17	1+ H or 1+ P or 1+ Rx	0.62	92.8	84.0	0.77	56.8	98.1
	18	1+H or 1+ P or 2+ Rx	0.64	92.2	85.7	0.78	59.4	98.0

Note: H = Hospital Separation; P = Physician Claim; Rx = Prescription Drug Claim; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

Agreement Between Survey and Administrative Data

The logistic regression model for hypertension was applied to the data for algorithm #6, which was based on one year of data. The model contained the main effects of age, sex, region of residence, presence of comorbid conditions and income adequacy quintile. The Hosmer-Lemeshow test revealed that the main effects model fit the data well.

The following explanatory variables were statistically significant in the logistic regression model: age (χ^2 = 161.1, df = 5, p < .0001) and the presence of comorbid conditions (χ^2 = 72.2, df = 1, p < .0001). The ORs and 95% CIs reported in Table 30 reveal that the odds of agreement between survey and administrative data were higher for individuals in younger age groups than for individuals in the oldest age group and were higher for individuals who did not have a comorbid condition than for individuals who did have a comorbid condition.

Table 30: Odds Ratio (OR) estimates and 95% CIs for predictors of agreement between administrative and survey data for hypertension

Predictors	OR	95% CI
Age		
19 – 44 years	9.2	(6.2, 13.6)
45 – 54 years	4.6	(3.1, 7.0)
55 – 64 years	2.8	(1.9, 4.2)
65 – 74 years	2.5	(1.7, 3.6)
75 – 84 years	1.9	(1.3, 2.8)
85+ years	Ref	–
Sex		
Males	0.8	(0.7, 1.0)
Females	Ref	–
Region of Residence		
North Rural RHAs	0.9	(0.6, 1.3)
South Rural RHAs	0.9	(0.7, 1.2)
Winnipeg RHA	Ref	–
Comorbidity		
Absent	2.6	(2.1, 3.2)
Present	Ref	–
Income Quintile		
Lowest	0.8	(0.5, 1.4)
Low Middle	0.9	(0.6, 1.4)
Middle	0.9	(0.7, 1.3)
Upper Middle	1.1	(0.8, 1.4)
Not Stated	0.8	(0.5, 1.2)
Highest	Ref	–

Source: Manitoba Centre for Health Policy, 2006

7.4 Provincial Prevalence Estimates

Cross-Sectional Prevalence Estimates

The crude prevalence estimates for each of the algorithms are summarized in Table 31. Algorithm #6, which resulted in the highest overall agreement between survey and administrative data, produced a prevalence estimate of 21.6%.

The algorithm with the highest sensitivity, which is based on one or more hospital separations, physician billing claims, or prescription drug records in a three-year period, resulted in a prevalence estimate of 27.4%. A prevalence estimate of 22.5% was produced by the algorithm with the highest value of Youden's index (i.e., algorithm #5).

Table 31: Crude provincial prevalence estimates for hypertension algorithms, 2000/01 – 2002/03

# of Years		Algorithm	Prevalence Estimate (%)
1	1	1+ P	14.2
	2	2+ P	9.3
	3	1+ H or 1+ P	14.7
	4	1+ H or 2+ P	10.0
	5	1+ H or 1+ P or 1+ Rx	22.5
	6	1+ H or 1+ P or 2+ Rx	21.6
2	7	1+ P	18.4
	8	2+ P	13.5
	9	1+ H or 1+ P	19.0
	10	1+ H or 2+ P	14.3
	11	1+ H or 1+ P or 1+ Rx	25.3
	12	1+ H or 1+ P or 2+ Rx	24.2
3	13	1+ P	21.0
	14	2+ P	15.9
	15	1+ H or 1+ P	21.6
	16	1+ H or 2+ P	16.8
	17	1+ H or 1+ P or 1+ Rx	27.4
	18	1+ H or 1+ P or 2+ Rx	26.2

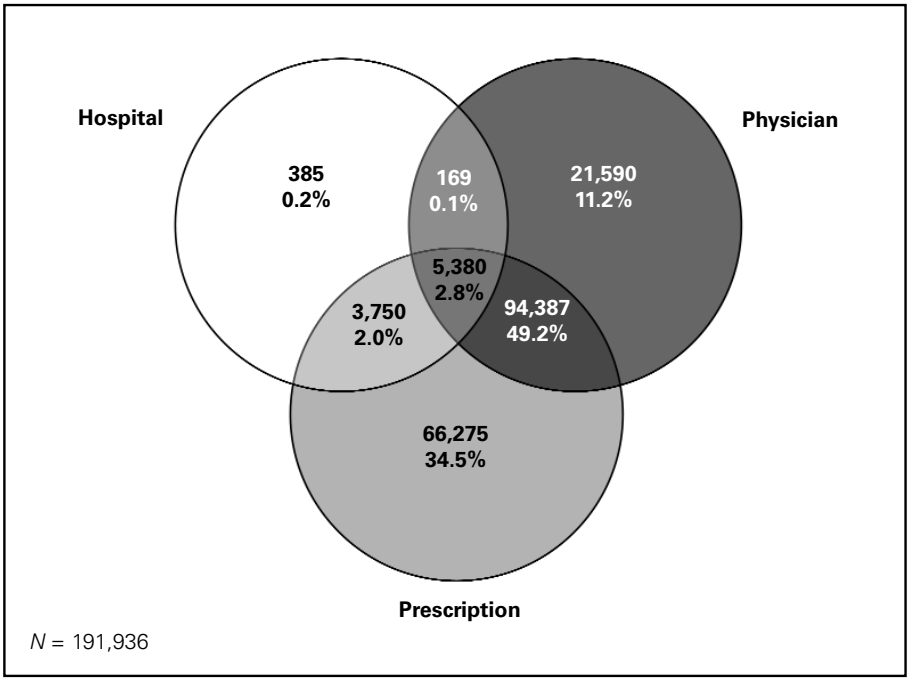
Note: Estimates for one-year algorithms are based on 2002/03; Estimates for two-year algorithms are based on 2001/02 – 2002/03; Estimates for three-year algorithms are based on 2000/01 – 2002/03.

Source: Manitoba Centre for Health Policy, 2006

Venn Diagrams

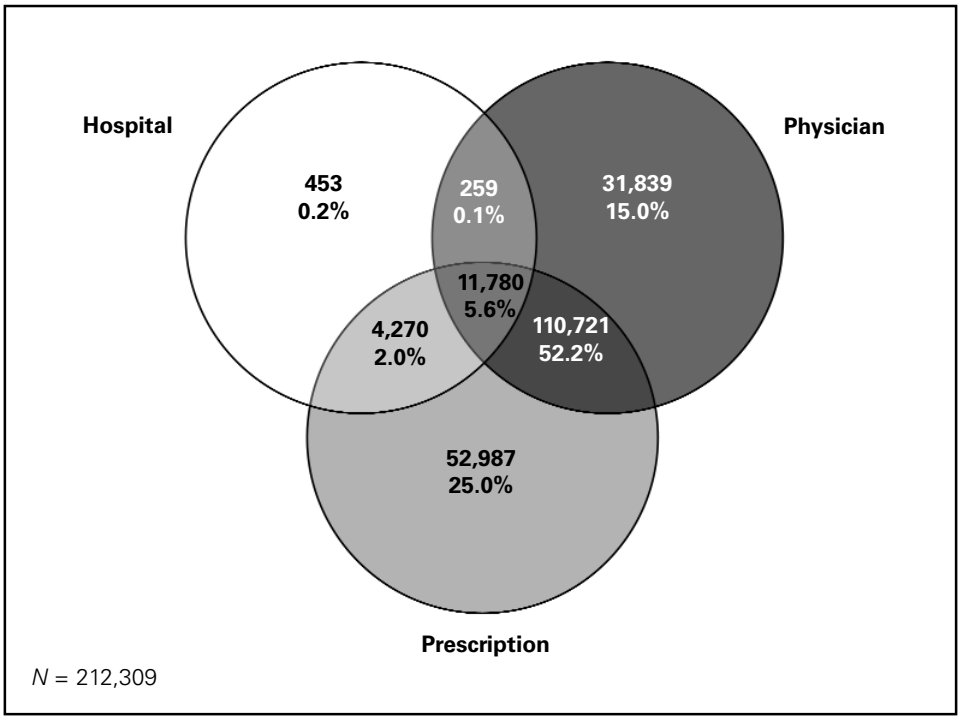
Figures 45, 46, and 47 contain Venn diagrams for algorithms #5, #11, and #17, respectively. Figure 45 shows that of the more than 190,000 cases of hypertension identified using 2002/03 administrative data, almost half (49.2%) were identified in both physician and prescription drug data. However, more than one third (34.5%) of cases were identified solely from prescription drug data. As expected, very few cases were identified only from hospital data (0.2%).

Figure 45: Hypertension Algorithm #5: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 1 Year



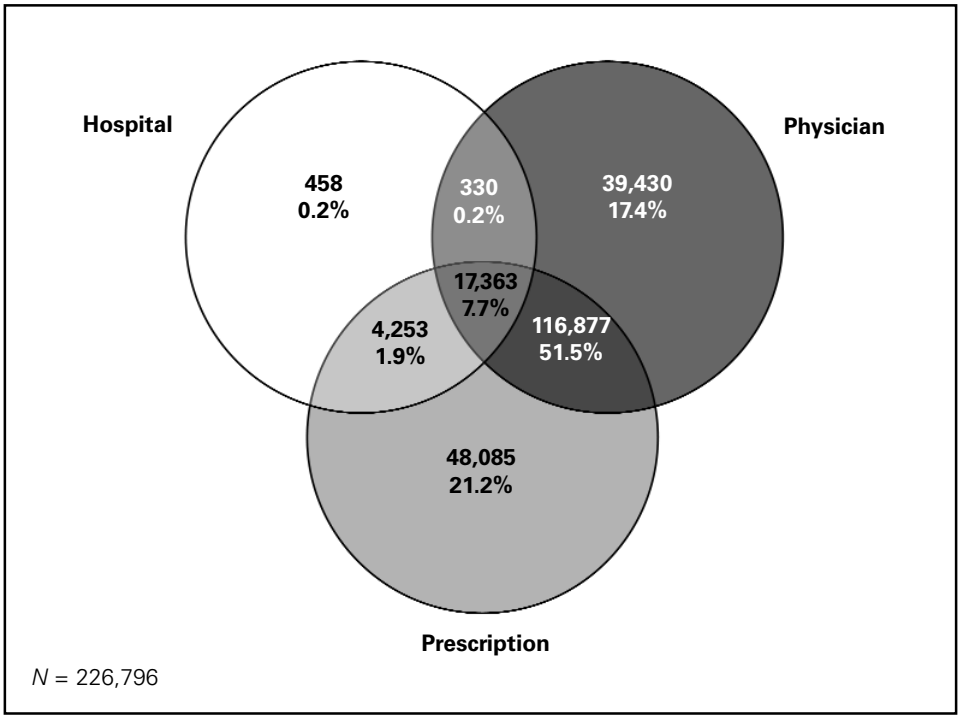
Source: Manitoba Centre for Health Policy, 2006

Figure 46: Hypertension Algorithm #11: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 2 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 47: Hypertension Algorithm #17: 1+ Hospital Separations or 1+ Physician Visits or 1+ Prescriptions, 3 Years



Source: Manitoba Centre for Health Policy, 2006

The number and percentage of cases identified from each of the three administrative data sources are similar for the two-year algorithm (Figure 46). One-quarter of cases were identified only from prescription drug records, while more than half (52.2%) were identified from both physician and prescription drug data. The algorithm based on three years of data (Figure 47) resulted in the identification of more than half (51.5%) of hypertension cases from both prescription and physician data and an additional 7.7% of cases appeared in all three data sources.

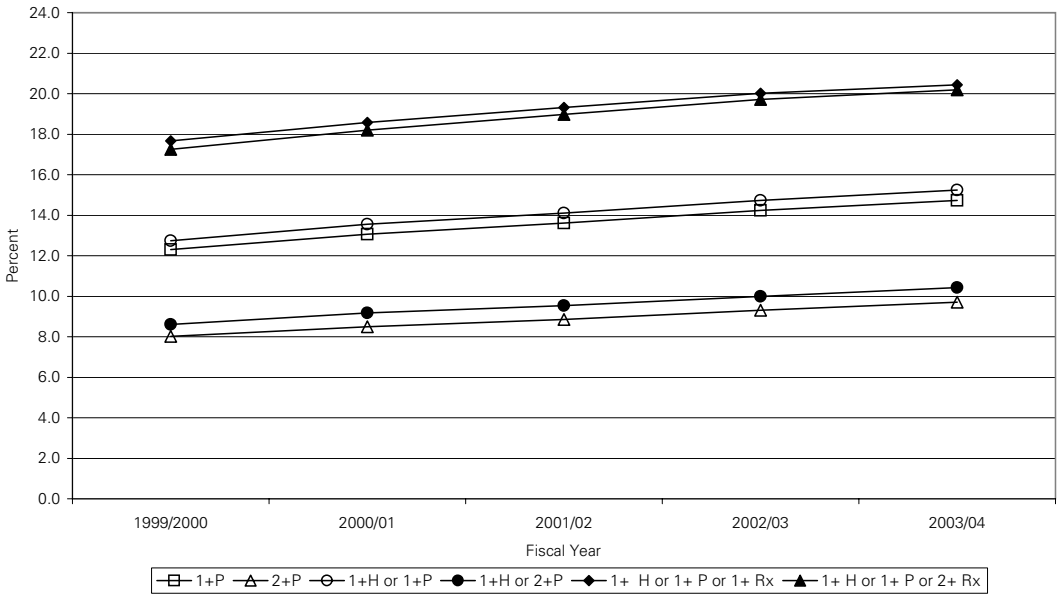
Regression Analyses for Cross-Sectional Prevalence Estimates

The regression models for the cross-sectional prevalence estimates included the following algorithms: #3, #5, #9, #11, #15, and #17. Thus, a total of six algorithms were investigated. The LRT for the full model, which contained the main effects of algorithm, age, sex, region of residence, and income quintile, in addition to two-way interactions, and the reduced model, which contained main effects only, was statistically significant ($\chi^2 = 421.0$; df = 60; $p < .0001$). Further investigation revealed that the algorithm x sex interaction was not statistically significant ($p = .1019$), but the algorithm x age ($p < .0001$), algorithm x region ($p = .0001$), and algorithm x quintile ($p = .0075$) effects were significant. These results indicate that the RR of hypertension prevalence for different algorithms varied across all of the sociodemographic characteristics of the population with the exception of sex.

Longitudinal Prevalence Estimates

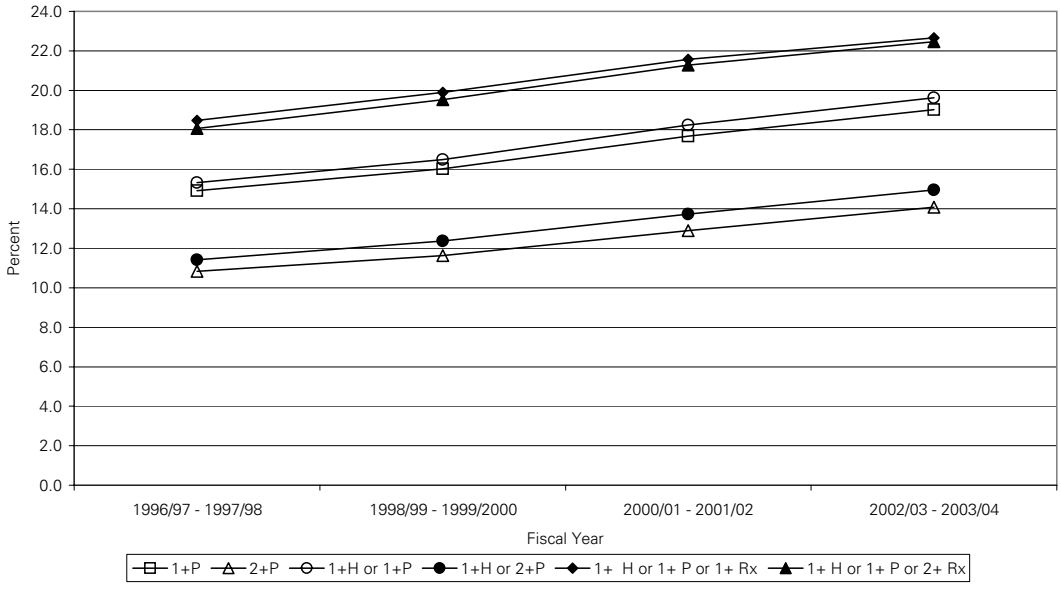
Figures 48, 49, and 50 depict the trends in prevalence estimates that were obtained for the one-year, two-year, and three-year algorithms, respectively. All three figures showed an increasing trend in prevalence, and all of the trend lines were approximately parallel, which suggests that each algorithm provides a similar picture of the change in the crude prevalence of hypertension over time.

Figure 48: Provincial Trends in Hypertension Prevalence for One-Year Algorithms, 1999/2000 – 2003/04



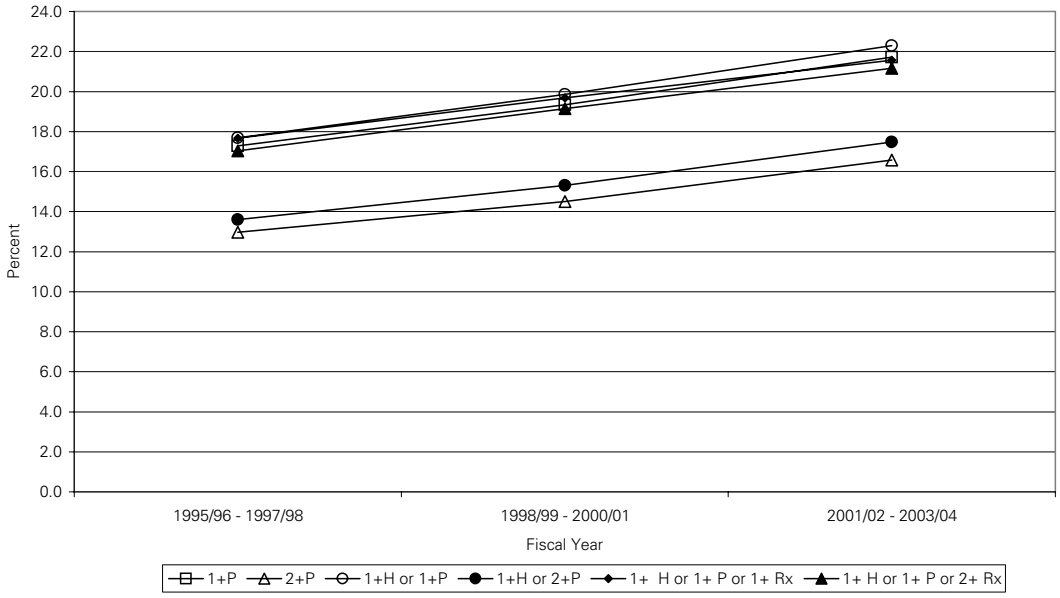
Source: Manitoba Centre for Health Policy, 2006

Figure 49: Provincial Trends in Hypertension Prevalence for Two-Year Algorithms, 1996/97 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Figure 50: Provincial Trends in Hypertension Prevalence for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

Regression Analyses for Longitudinal Prevalence Estimates

Regression analyses were conducted for the longitudinal prevalence estimates, to test for differences in the estimates from the different algorithms over time, and also to test whether the longitudinal estimates varied with the sociodemographic characteristics of the population (i.e., age, sex, region of residence, and income). For the one-year algorithms, the LRT for the partial and reduced models was not statistically significant (χ^2 = 1.7, df = 5, p = .8880), which indicates that there was no difference in the RR of change for different algorithms. However, the LRT for the full model, which contained main effects in addition to selected two-way and three-way interactions, and the reduced model, which contained main effects only, was statistically significant (χ^2 = 1927.6, df = 137, p < .0001). An assessment of the GEE model results revealed that all of the main effects were statistically significant (p < .0001), as were the following interactions: algorithm x age (p < .0001), algorithm x sex (p < .0001), algorithm x region (p = .0217), algorithm x time x age (p < .0001), algorithm x time x sex (p < .0001), and algorithm x time x region (p < .0001). These results indicate that the RR of change in the prevalence estimates for different algorithms varied with age, sex, and region of residence of the population over time.

The regression analyses for the two-year algorithms produced different results. The LRT for the partial and reduced models was statistically significant (χ^2 =13.0, df = 5, p = .0234), as was the LRT for the full and reduced models (χ^2 =679.2, df = 137, p < .0001). The former result indicates that the RR of change in hypertension varied across the algorithms. The GEE

results for the full model revealed that all of the main effects were statistically significant ($p < .0001$) in addition to the following two-way and three-way interaction terms: algorithm x time ($p = .0234$), algorithm x age ($p < .0001$), algorithm x sex ($p < .0001$), algorithm x region ($p = .0217$), algorithm x age x time ($p < .0001$), and algorithm x region x time ($p < .0001$). This finding demonstrates that the RR of change in prevalence estimates for different algorithms did vary with the sociodemographic characteristics of the population over time.

Finally, the regression analyses for the three-year algorithms showed that the LRT for the partial and reduced models was not statistically significant ($\chi^2 = 6.3$, df = 5, $p < .2815$), but the LRT for the full and reduced models was significant ($\chi^2 = 1780.8$; df = 137, $p < .0001$). The former result indicates that the RR of change in hypertension prevalence was not significantly different for different algorithms. With respect to the full model, the GEE analyses revealed that all of the main effects were statistically significant ($p < .0001$) in addition to the following two-way and three-way interactions: algorithm x time ($p = .0015$), algorithm x age ($p < .0001$), algorithm x sex ($p < .0001$), time x algorithm x age ($p < .0001$), algorithm x time x sex ($p < .0001$), algorithm x time x quintile ($p < .0001$), and algorithm x time x region ($p = .0003$). This result indicates that the RR of change in prevalence estimates for different algorithms did vary with the sociodemographic characteristics of the population over time.

7.5 Chapter Summary

This study confirms the results of previous studies that showed administrative data have good validity for identifying cases of hypertension.

This study confirms the results of previous studies that showed administrative data have good validity for identifying cases of hypertension. Two different algorithms resulted in very good agreement between survey and administrative data; one of these algorithms was based on a combination of hospital separations and physician billing claims and prescription drug records in one year, while the other relied on only hospital separations and physician billing claims in three years.

The algorithm based on one or more contacts in hospital separations, physician billing claims, or prescription drug records in a one-year period had the highest value of Youden's index. The algorithm based on one or more contacts in hospital separations or physician billing claims or prescription drug records in three years had the highest sensitivity. The positive predictive value was slightly lower for the former algorithm than for an algorithm which was based on one or more contacts in either hospital separations or prescription drug records or two or more contacts in physician billing claims.

Agreement between survey and administrative data was predicted by age and the presence of comorbid conditions.

Prevalence estimates were very similar (i.e., approximately 22%) for the algorithms with the highest values for kappa and Youden's index. The regression results indicated that the relative rate of hypertension prevalence for the investigated algorithms varied across the sociodemographic characteristics of the population. The analyses of the longitudinal trends revealed that the relative rate of change over time varied for some of the algorithms.

CHAPTER 8: STROKE[1]

8.1 Introduction and Review of Literature

Stroke is ranked third as a cause of death in Canada following only heart disease and cancer.

This chapter focusses on the use of administrative data for identifying cases of stroke, from administrative data. Stroke is ranked third as a cause of death in Canada following only heart disease and cancer (http://www.statcan.ca/english/Pgdb/health36.htm). The societal burden associated with non-fatal stroke is estimated to eclipse other chronic disorders (Verbrugge et al., 1989; Dobkin, 2003).

Recognition of the importance of stroke surveillance in Canada has resulted in several initiatives. In 2000, the Heart and Stroke Foundation of Canada released the report *The Changing Face of Heart Disease and Stroke in Canada* which highlights risk factors, use of health care services, and health outcomes associated with stroke. The Canadian Stroke Network, established in 1999, include researchers from across the country who seek to decrease the physical, social and economic consequences of stroke. Part of this initiative is a national registry of stroke patients.

The validity of administrative data for identifying both fatal and non-fatal cases of stroke has been investigated in a large number of studies. Several of these studies are summarized in Appendix Table A.6. A significant issue in the research has been the choice of diagnostic codes to identify disease cases. Some studies adopted the broadest possible set of codes, which included ICD-9-CM 430 to 438 (cerebrovascular disease). However, others excluded specific codes, like 437, which represent stroke of undetermined causes. In general, however, the consensus of these studies is that administrative data can provide a valid tool for identifying stroke cases using one or more ICD-9-CM codes.

While the methodology associated with stroke case identification has been the subject of multiple investigations, a limitation of all of the studies identified in Appendix A is that they have only relied on hospital separations to identify stroke cases. As Kokotail and Hill (2005) observe, this results in a bias towards the identification of only the most severe cases of stroke from administrative data. The validity of physician claims and prescription drug records for stroke surveillance, to identify less severe non-fatal stroke cases has been unexplored.

[1]Material contained in this chapter was prepared by Dr. David Moore in partial fulfillment of the course 93.741 Directed Readings in Community Health Sciences, University of Manitoba. It has been edited by Dr. Lisa Lix, principal investigator, to conform to the form and style used in other chapters of this report.

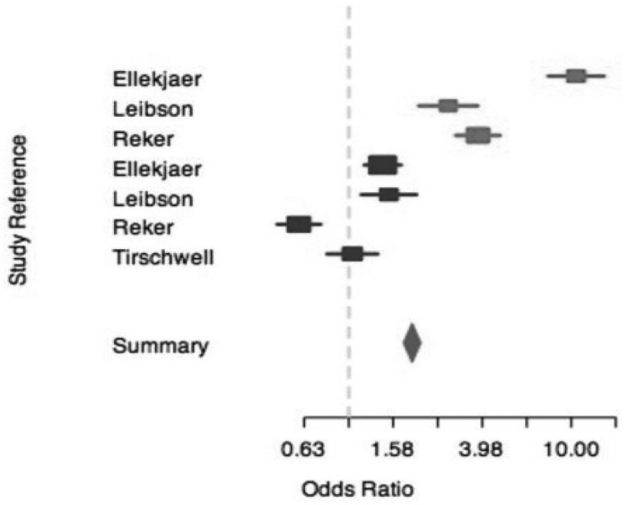
8.2 Description of Stroke Algorithms

The algorithms selected for validation and testing in this research were based on a meta-analysis. A MEDLINE search on the terms administrative database *AND* stroke *OR* cerebrovascular disease returned a total of twenty-eight references from the period 1965 to 2005. After full review of these articles, four were selected and compared by meta-analysis. The selected studies each included a numerical validation of stroke case identification from administrative databases using either an independently maintained stroke registry or an independent prospective or retrospective chart review. All of the studies selected for the meta-analysis used only hospital separations for the identification of stroke cases. Studies excluded from the meta-analysis either did not validate stroke algorithms or did not report sufficient data to produce the validation indices of sensitivity, specificity, PPV, and NPV.

The selected studies included a sensitive algorithm based on ICD-9-CM codes in the range from 430 to 438, allowing inclusion of all patients identified as having cerebrovascular disease, and a more specific algorithm limited to ICD-9-CM codes 430, 431, 434, 435 and 436. Using the stated results in an article, a 2 x 2 classification table was derived for each study, allowing calculation of a Mantel-Haenszel statistic. The Mantel-Haenszel statistic is used to compute a pooled OR. A pooled OR of 1.0 indicates that the probability of the event (i.e., stroke case identification) is equally likely in both the gold standard data and administrative data. A value greater than 1.0 indicates an overestimate of stroke cases by the administrative data compared to the stroke cases identified using the gold standard data, while OR values less than 1.0 indicate that stroke case identification is less likely in the gold standard data source. A total of three data sets were compared for the sensitive algorithm and four for the specific algorithm.

The Mantel-Haenszel OR for the summary meta-analysis combining both the sensitive and specific algorithms was 1.9 (95% CI 1.8 – 2.1). In Figure 51, it is seen that overall, the more specific ICD9-CM stroke diagnostic codes more consistently represent stroke cases as determined by the gold standard. The OR ratio of 1.0 is indicated as the dashed vertical line in the figure. The three points on the right in Figure 51 represent the ORs for the sensitive stroke algorithm while the four points on the left represent the ORs for the specific algorithms.

Figure 51: Plot of the Meta-Analysis ORs for Stroke Cases



Source: Manitoba Centre for Health Policy, 2006

On the basis of the results of the meta-analysis it was expected that the specific ICD-9 codes would give a more accurate estimate of the true prevalence of stroke in the population. In this report, both sensitive and specific sets of algorithms were investigated using broad (ICD-9-CM 430-438) and narrow (ICD-9-CM 430, 431, 434, 435, 436) sets of diagnostic codes, respectively.

Based on consultations with clinical experts, the following drug categories were used to identify prescription drug records for inclusion in the research: anti-platelet agents such as aspirin (ASA) at 81 or 325 mg once a day, clopidogrel, ticlopidine, dipyridamole, and combination agents such as Aggrenox (ASA 25mg dipyridamole 200mg slow release) and oral anti-coagulants such as warfarin, pheninidione, and nicoumalone. The fifth level ATC codes selected for the research were B01AA02, B01AA03, B01AA07, B01AC07, B01AB01, B01AC30, B01AC05, B01AC06, B01AC04, B01AB09, B01AB04, B01AB10. Thrombolytic agents such as rt-PA (recombinant tissue plasminogen activator) and intravenous anti-platelet agents (anti GP 2b/3a) such as abciximab, tirofiban and eptifibatide are markers for stroke therapy when administered on an inpatient basis. However, they can not be used as markers of stroke therapy when used on an outpatient basis. Therefore prescription drug records with DINs for these drugs were not included in the algorithms. Parenteral anticoagulants such as heparin and low molecular weight heparins such as enoxiparin are also given as stroke therapy on an inpatient basis, but are less likely to be markers of stroke therapy when used on an outpatient basis. Therefore, prescription drug records with DINs for these drugs were not included in the algorithms. The DINs for all drugs with the ATC codes noted above were obtained from the MCHP Master Formulary.

Table 32 lists the 24 algorithms that were selected for this research. Algorithms #1, #2, and #3, are based on the specific set of ICD-9-CM codes and algorithms #1f, #2f, and #3f are based on the sensitive (or full) set of diagnostic codes in one year of data. For two, three, and five years of administrative data, the same ordering has been used to define the specific and sensitive algorithms. For example, algorithms #4 and 4f are based on a single hospital separation with the specific and sensitive diagnostic codes, respectively, in two years.

It is important to note that none of the algorithms relied solely on prescription drug records to identify non-fatal stroke cases. This is because the drugs selected are not used exclusively as markers for stroke. Using one or more contacts in prescription drug records to identify stroke cases was expected to result in low specificity, as we have verified for other chronic diseases.

Table 32: Stroke algorithms selected for validation

# Years	Algorithm	Hospital Separations or	Physician Claims or	Physician Claims and Prescription Drug Records
1	1	1 or more		
	2	1 or more	1 or more	
	3	1 or more	2 or more	1 and 2 or more
	1f	1 or more		
	2f	1 or more	1 or more	
	3f	1 or more	2 or more	1 and 2 or more
2	4	1 or more		
	5	1 or more	1 or more	
	6	1 or more	2 or more	1 and 2 or more
	4f	1 or more		
	5f	1 or more	1 or more	
	6f	1 or more	2 or more	1 and 2 or more
3	7	1 or more		
	8	1 or more	1 or more	
	9	1 or more	2 or more	1 and 2 or more
	7f	1 or more		
	8f	1 or more	1 or more	
	9f	1 or more	2 or more	1 and 2 or more
5	10	1 or more		
	11	1 or more	1 or more	
	12	1 or more	2 or more	1 and 2 or more
	10f	1 or more		
	11f	1 or more	1 or more	
	12f	1 or more	2 or more	1 and 2 or more

Note: The first three algorithms in each year are based on a specific set of ICD-9-CM codes (430, 431, 434, 435, 436), while the last three algorithms are based on a sensitive (i.e., full) set of codes (430 – 438).

Source: Manitoba Centre for Health Policy, 2006

8.3 Validation Results

Validation Indices

Table 33 contains the point estimates for the six validation indices for the 24 algorithms. The 95% CIs for each of these estimates are reported in Appendix D, in Table D.11.

Table 33: Estimates of agreement, sensitivity, specificity, and predictive values for stroke algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ H	0.12	7.3	99.8	0.07	42.1	98.3
	2	1+ H or 1+ P	0.35	28.4	99.4	0.28	49.2	98.7
	3	1+ H or 2+ P or (1 P & 2+ Rx)	0.35	26.6	99.6	0.26	52.7	98.6
	1f	1+ H	0.22	14.7	99.7	0.14	51.6	98.4
	2f	1+ H or 1+ P	0.39	34.9	99.2	0.34	46.3	98.8
	3f	1+ H or 2+ P or (1 P & 2+ Rx)	0.40	33.9	99.4	0.33	52.9	98.8
2	4	1+ H	0.24	15.6	99.8	0.15	56.7	98.4
	5	1+ H or 1+ P	0.44	43.1	99.1	0.42	46.5	98.9
	6	1+ H or 2+ P or (1 P & 2+ Rx)	0.43	38.5	99.3	0.38	51.2	98.9
	4f	1+ H	0.33	23.8	99.7	0.24	56.5	98.6
	5f	1+ H or 1+ P	0.42	40.5	98.7	0.39	40.5	99.0
	6f	1+ H or 2+ P or (1 P & 2+ Rx)	0.43	42.2	99.1	0.41	46.9	98.9
3	7	1+ H	0.26	18.4	99.7	0.18	50.0	98.5
	8	1+ H or 1+ P	0.42	47.7	98.7	0.46	40.3	99.0
	9	1+ H or 2+ P or (1 P & 2+ Rx)	0.44	44.0	99.0	0.43	45.7	99.0
	7f	1+ H	0.38	29.4	99.6	0.29	55.2	98.7
	8f	1+ H or 1+ P	0.42	54.1	98.2	0.52	36.6	99.1
	9f	1+ H or 2+ P or (1 P & 2+ Rx)	0.45	49.5	98.8	0.48	43.2	99.0
5	10	1+ H	0.41	33.0	99.6	0.33	58.1	98.8
	11	1+ H or 1+ P	0.45	60.6	98.1	0.59	37.1	99.2
	12	1+ H or 2+ P or (1 P & 2+ Rx)	0.46	54.1	98.6	0.53	42.1	99.1
	10f	1+ H	0.45	39.4	99.4	0.39	53.7	98.9
	11f	1+ H or 1+ P	0.43	66.1	97.5	0.64	33.0	99.4
	12f	1+ H or 2+ P or (1 P & 2+ Rx)	0.47	61.5	98.3	0.60	39.9	99.3

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; PPV = Positive Predictive Value; NPV = Negative Predictive Value; 95% confidence intervals for all estimates are reported in Appendix D.

Source: Manitoba Centre for Health Policy, 2006

There was low to moderate agreement between the survey and administrative data, with values of κ ranging from 0.12 to 0.47. The highest estimate of κ was obtained for the sensitive algorithm based on one or more hospital separations or two or more physician billing claims, or one physician billing claim and two or more prescription drug records in five years (i.e., algorithm #12f). However, the corresponding five-year algorithm based on the specific ICD-9-CM codes produced almost the same estimate (i.e., $\kappa = 0.46$).

As the results indicate, using an algorithm based on one or more hospital separations in one year of administrative data to identify non-fatal stroke cases resulted in low agreement and sensitivity regardless of whether a specific (7.3%) or sensitive (14.7%) set of ICD-9-CM codes was selected (i.e., algorithms #1 and #1f). Sensitivity improved substantially when the physician billing claims were used in addition to the hospital data to identify stroke cases. For example, with a single year of data, the difference in sensitivity between algorithm #1f, based on hospital separations only, and #2f, based on one or more contacts in either hospital separations or physician claims, was 20.2%. While there was a slight drop in sensitivity when the algorithms based on contacts in one or more of the administrative data sources were adopted, the PPV of a stroke diagnosis improved slightly. For example, with a single year of data, the PPV increased by 6.6% when algorithms #2f and #3f were compared. Specificity was very high (i.e., above 97%) for all of the algorithms.

The algorithm which resulted in the highest combined estimate of sensitivity and specificity (0.64) was algorithm #11f, which was based on one or more hospital separations or one or more physician claims in five years of data. However, the PPV of a stroke diagnosis was much lower for this algorithm (33.0%) than for algorithm #12f (39.9%), which was based on one or more hospital separations or two or more physician billing claims or one physician billing claim and two or more prescription drug records in five years, which produced only a slightly lower estimate of Youden's index (0.60).

The PPV of a stroke diagnosis ranged from 33.0% to 58.1%. NPV was consistently above 98% for all of the algorithms.

Agreement Between Survey and Administrative Data

Logistic regression analysis was used to test the sociodemographic variables associated with agreement between survey and administrative data for stroke. The logistic model was applied to the data for algorithm #12f. The model contained the main effects of age, sex, region of residence, income adequacy quintile, and comorbid conditions (i.e., presence of heart disease or diabetes). Two-way interactions were tested but were not statistically significant and were therefore excluded from the final model. The Hosmer-Lemeshow test indicated that the main effects model fit the data well.

The following variables were statistically significant: age (χ^2 = 91.8, df = 4, p < .0001), sex (χ^2 = 4.4, df = 1, p = .0353), presence of comorbid conditions (χ^2 = 18.8, df = 1, p < .0001), and income adequacy quintile (χ^2 = 13.3, df = 5, p = .0201). Table 34 contains the ORs and 95% CIs for all explanatory variables. The odds of agreement between the two data sources were higher

for individuals in younger age groups than for individuals in the oldest age group, and were lower for males than for females. The odds of agreement were also higher for individuals with no comorbid conditions than for those with comorbid conditions and lower for individuals in low income adequacy quintiles than for individuals in the highest quintile.

Table 34: Odds Ratio (OR) estimates and 95% CIs for predictors of agreement between administrative and survey data for stroke

Predictors	OR	95% CI
Age		
19 – 49 years	28.9	(13.0, 64.3)
50 – 64 years	10.0	(4.9, 20.1)
65 – 74 years	2.9	(1.6, 5.0)
75 – 84 years	1.6	(1.0, 2.7)
85+ years	Ref	–
Sex		
Males	0.7	(0.5, 1.0)
Females	Ref	–
Region of Residence		
North Rural RHAs	0.8	(0.4, 1.7)
South Rural RHAs	1.2	(0.8, 1.8)
Winnipeg RHA	Ref	–
Comorbidity		
Absent	2.4	(1.6, 3.5)
Present	Ref	–
Income Quintile		
Lowest	0.1	(0.1, 0.4)
Low Middle	0.2	(0.1, 0.8)
Middle	0.2	(0.1, 0.8)
Upper Middle	0.3	(0.1, 0.9)
Not Stated	0.2	(0.1, 0.5)
Highest	Ref	–

Source: Manitoba Centre for Health Policy, 2006

8.4 Provincial Prevalence Estimates

Cross-Sectional Prevalence Estimates

Prevalence estimates for all 24 algorithms are summarized in Table 35. There was substantial variability both within and across the one-year, two-year, three-year, and five-year sets of algorithms. The algorithms based on a single hospital separation in one year of data resulted in estimates of 0.2% and 0.3% when the specific and sensitive diagnostic codes were adopted, respectively. The difference between the corresponding algorithms based on the sensitive and specific sets of ICD-9-CM codes was never greater than 0.6%. Descriptive analyses of the data revealed that a single ICD-9-CM code, 438 (late effects of cerebrovascular disease), was responsible for the majority of the difference in prevalence estimates between the sensitive and specific algorithms, as more than 20% of individuals were assigned this code. There were very few cases coded using ICD-9-CM 432 (other and unspecified intracranial hemorrhage).

For algorithm #12f, which had the highest value of the κ statistic, the crude prevalence of non-fatal stroke in Manitoba was estimated to be 2.9%. For algorithm #11f, which had the highest value of Youden's index, prevalence was estimated to be 3.8%.

Table 35: Crude provincial prevalence estimates for stroke

# Years		Algorithm	Prevalence Estimate (%)
1	1	1+ H	0.2
	2	1+ H or 1+ P	1.0
	3	1+ H or 2+ P or (1 P & 2+ Rx)	0.7
	1f	1+ H	0.3
	2f	1+ H or 1+ P	1.3
	3f	1+ H or 2+ P or (1 P & 2+ Rx)	1.0
2	4	1+ H	0.4
	5	1+ H or 1+ P	1.7
	6	1+ H or 2+ P or (1 P & 2+ Rx)	1.2
	4f	1+ H	0.6
	5f	1+ H or 1+ P	2.1
	6f	1+ H or 2+ P or (1 P & 2+ Rx)	1.6
3	7	1+ H	0.6
	8	1+ H or 1+ P	2.2
	9	1+ H or 2+ P or (1 P & 2+ Rx)	1.7
	7f	1+ H	0.9
	8f	1+ H or 1+ P	2.8
	9f	1+ H or 2+ P or (1 P & 2+ Rx)	2.1
5	10	1+ H	0.9
	11	1+ H or 1+ P	3.2
	12	1+ H or 2+ P or (1 P & 2+ Rx)	2.4
	10f	1+ H	1.3
	11f	1+ H or 1+ P	3.8
	12f	1+ H or 2+ P or (1 P & 2+ Rx)	2.9

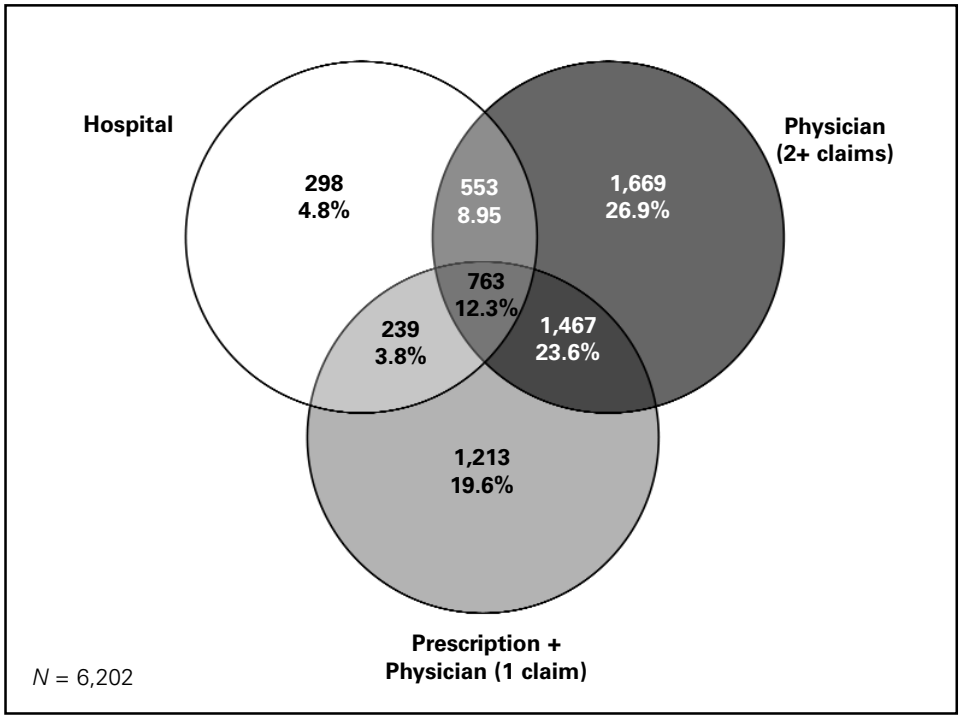
Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 1-year estimates are for 2002/03, 2-year estimates are for 2001/02 – 2002/03, 3-year estimates are for 2000/01 – 2002/03, 5-year estimates are for 1998/99 – 2002/03.

Source: Manitoba Centre for Health Policy, 2006

Venn Diagrams

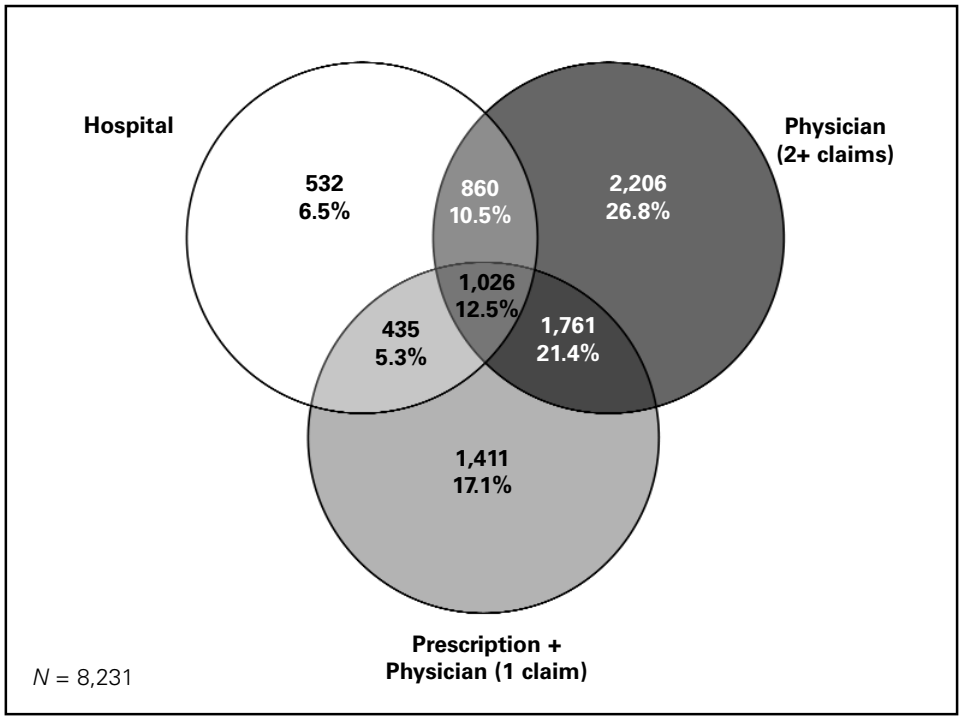
Venn diagrams for algorithms #3, #3f, #12, and #12f are reported in Figures 52 to 55. The first of these figures shows that when a single year of data and a specific set of diagnostic codes were used, slightly more than 6,200 cases of non-fatal stroke were identified in Manitoba's administrative data using a combination of hospital separations, physician billing claims, and prescription drug records. More than one-quarter were identified from physician billing claims with two or more contacts. More than 10% were identified in all three data sources. When the sensitive set of diagnostic codes was used, the total number of cases identified increased by approximately 2,000. The proportions of stroke cases identified from the physician and prescription data remained about the same, but greater numbers of cases were identified using only hospital separations, or hospital separations, or physician billing claims, or prescription drug records.

Figure 52: Stroke Algorithm #3: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 1 Year



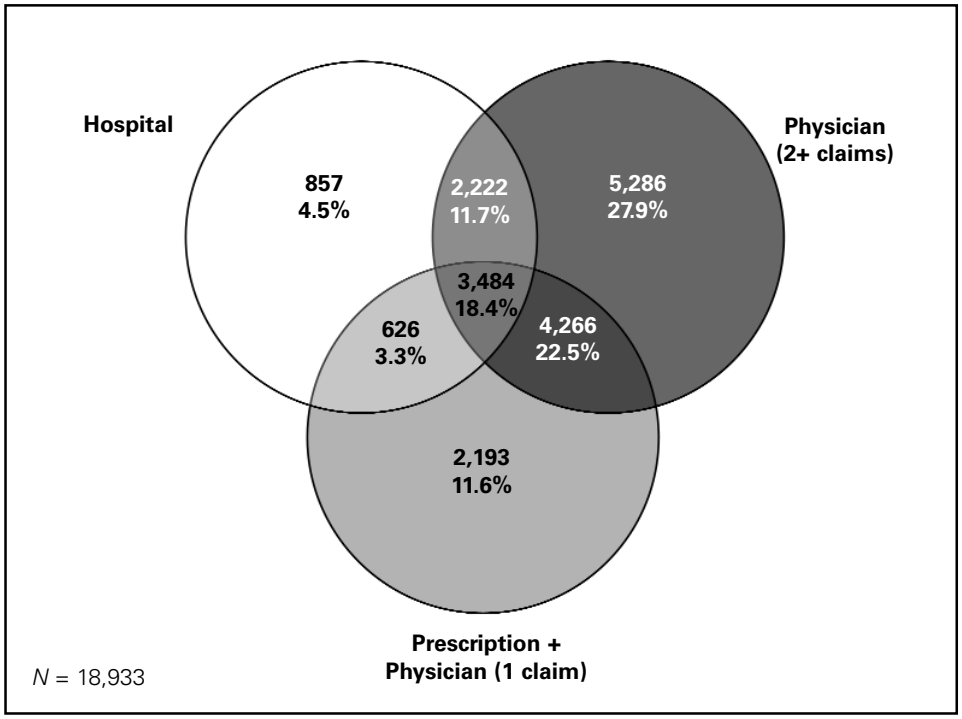
Source: Manitoba Centre for Health Policy, 2006

Figure 53: Stroke Algorithm #3f: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 1 Year



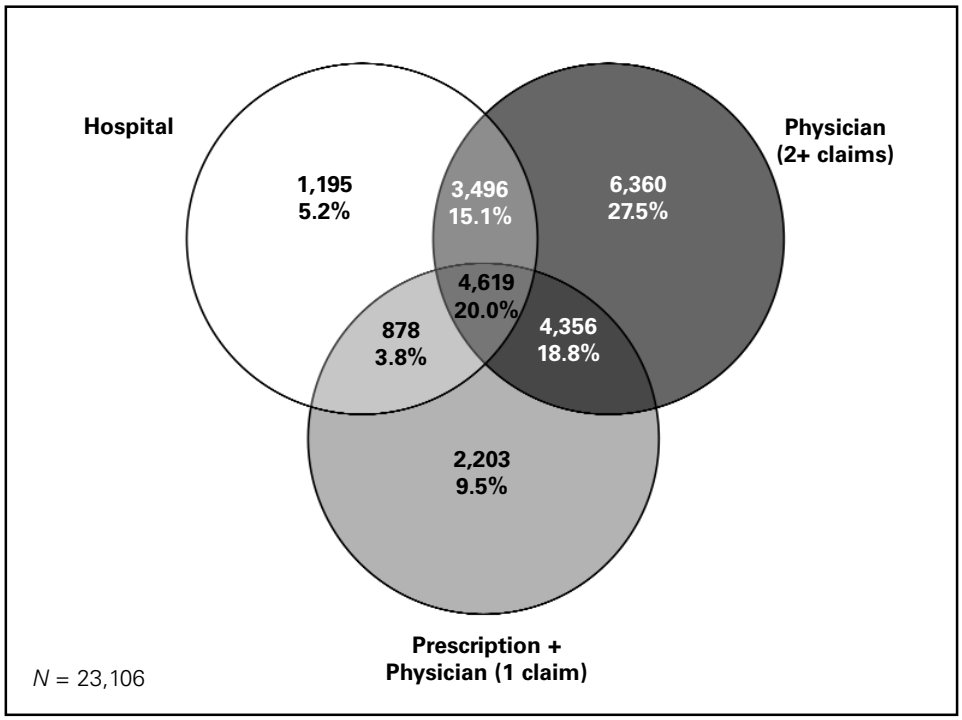
Source: Manitoba Centre for Health Policy, 2006

Figure 54: Stroke Algorithm #12: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 5 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 55: Stroke Algorithm #12f: 1+ Hospital Separations or 2+ Physician Visits or 1 Physician Visit & 2+ Prescriptions, 5 Years



Source: Manitoba Centre for Health Policy, 2006

Figure 54 shows that when the number of years of administrative data was increased to five, the total number of cases tripled. The percent of stroke cases identified with one physician billing claim and two or more prescription drug records in five years was much lower than the number identified using the corresponding one-year algorithm (i.e., decrease from 19.6% to 11.6%). Comparing the algorithms based on one and five years of administrative data, the percent of cases identified in all three data sources also increased substantially, to 18.5%. The same pattern was observed when algorithm #12f was compared to algorithm #3f (see Figure 55).

Regression Analyses for Cross-Sectional Prevalence Estimates

The regression analyses were used to test for differences in prevalence using two different models. The first model was for the algorithms based on the specific set of ICD-9-CM diagnostic codes. It included algorithms #1 (reference), #3, #4, #6, #7, #9, #10, and #12. The second was for the algorithms based on the sensitive set of ICD-9-CM diagnostic codes. It included algorithms #1f (reference), #3f, #4f, #6f, #7f, #9f, #10f, and #12f. Thus, a total of eight algorithms were investigated in each of the regression models.

For the first model, based on the specific set of diagnostic codes, the LRT for the full and reduced models was statistically significant (χ^2 = 329.7, df = 84.0, p < .0001). The algorithm x age (p = .0018), algorithm x sex (p =

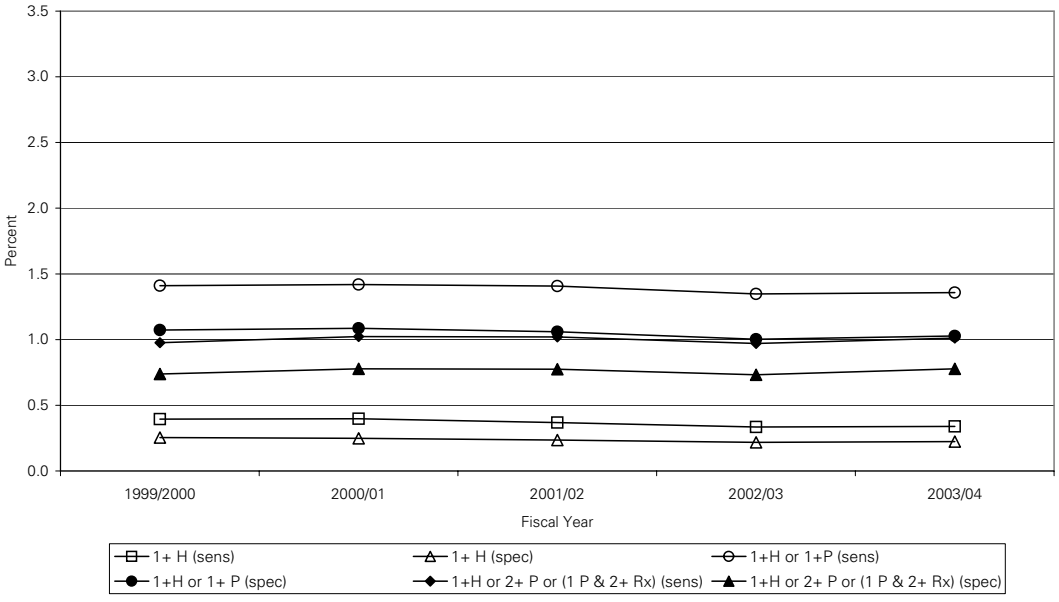
.0178), and algorithm x region ($p < .0001$) effects were all statistically significant. These results indicate that the relative prevalence estimates from different algorithms varied with the sociodemographic characteristics of the population.

For the second model, based on the sensitive set of diagnostic codes, the LRT for the full and reduced models was also statistically significant ($\chi^2 = 315.0$, df = 84.0, $p < .0001$). The algorithm x age ($p < .0001$), algorithm x sex ($p = .0009$), and algorithm x region ($p < .0001$) effects were all statistically significant. These results also indicate that the relative prevalence estimates from different algorithms varied with the sociodemographic characteristics of the population.

Longitudinal Prevalence Estimates

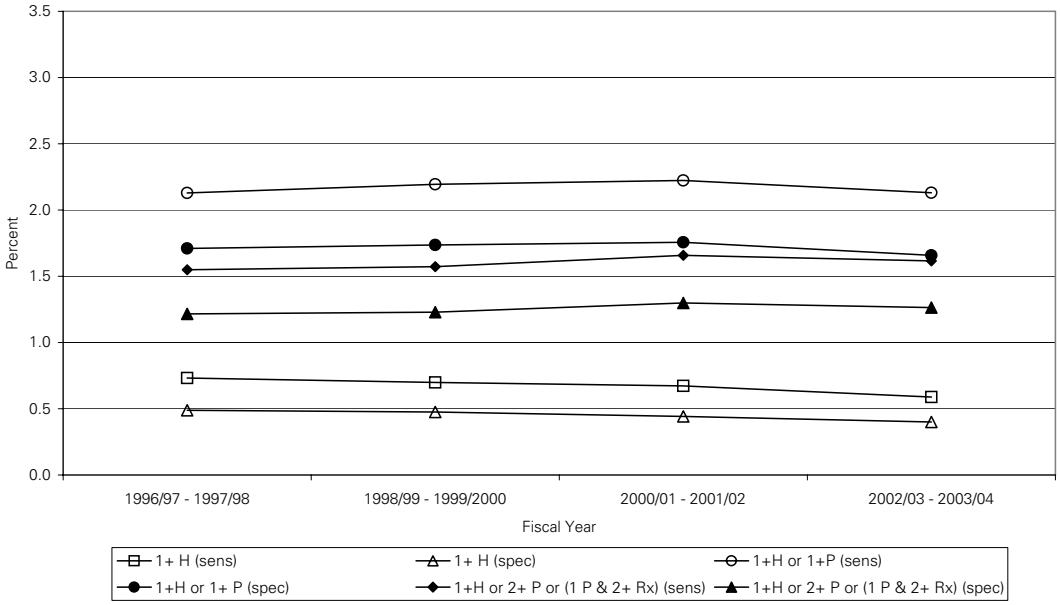
Figures 56, 57, and 58 depict the longitudinal change in the crude prevalence estimates for the stroke algorithms based on one, two, and three years of administrative data. Figure 56 reveals that according to the one-year algorithms, the prevalence of non-fatal stroke remained largely unchanged over time. The trend lines for all algorithms were approximately parallel, indicating each provided a similar picture of the change in the crude stroke prevalence in the most recent five years of the study period.

Figure 56: Provincial Trends in Stroke Prevalence for One-Year Algorithms, 1999/2000 – 2003/04



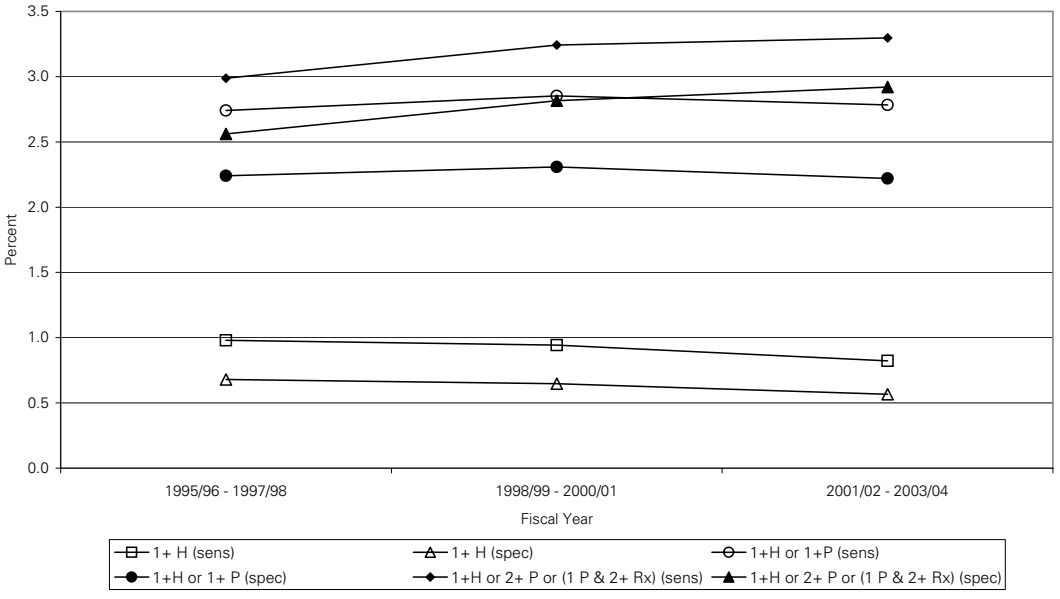
Source: Manitoba Centre for Health Policy, 2006

Figure 57: Provincial Trends in Stroke Prevalence for Two-Year Algorithms, 1996/97– 2003/04



Source: Manitoba Centre for Health Policy, 2006

Figure 58: Provincial Trends in Stroke Prevalence for Three-Year Algorithms, 1995/96 – 2003/04



Source: Manitoba Centre for Health Policy, 2006

On the other hand, the remaining two figures show that the two-year and three-year algorithms provide very different pictures of the change in stroke prevalence in Manitoba over time. The two algorithms based on only hospital separations resulted in a slight decrease over time, while the algorithms based on contacts in hospital separations, or physician billing claims, or prescription drug records resulted in a slight increase in prevalence over time. The two algorithms based on one or more contacts in either hospital separations or physician claims showed a modest decrease in prevalence.

Regression Analyses for Prevalence Estimates

Regression analyses were conducted for the longitudinal prevalence estimates, to test for differences in the RR of change for the estimates from different algorithms, and also to test whether these longitudinal estimates covaried with the sociodemographic variables of age, sex, region, and income quintile. Each of the one-, two-, and three-year models include the six algorithms that used the sensitive and specific sets of ICD-9-CM codes.

For the one-year algorithms, the LRT for the partial and reduced models was statistically significant (χ^2 = 32.7; df = 5; p < .0001), as was the LRT for the full and reduced models (χ^2 = 1131.0; df = 115; p < .0001). The former result indicates that relative rate (RR) of change in prevalence for different algorithms varied over time. For the latter finding, the GEE results indicated that the following interactions were statistically significant: algorithm x age (p < .0001), algorithm x sex (p = .0042), algorithm x region (p < .0001), algorithm x time (p < .0001), and algorithm x region x time (p < .0001). This result indicates that the RR of change in prevalence for different algorithms not only varied over time, but also across the sociodemographic characteristics of the population.

For the two-year algorithms, the LRT for the partial and reduced models was statistically significant (χ^2 = 60.0; df = 5; p < .0001), as was the LRT for the full and reduced models χ^2 = 1180.1; df = 115; p < .0001). Moreover, the GEE results indicated that the following interactions were statistically significant: algorithm x age (p < .0001), algorithm x sex (p = .0002), algorithm x region (p < .0001), algorithm x time (p < .0001), algorithm x region x time (p < .0001), algorithm x time x sex (p = .0287). These results indicate that the RR of change over time in the prevalence of stroke was different across the algorithms, and that this difference also varied across the sociodemographic characteristics of the population.

For the three-year algorithms the LRT for the partial and reduced models was statistically significant (χ^2 = 76.0; df = 5; p < .0001), as was the LRT for the full and reduced models (χ^2 = 1028.8; df = 115; p < .0001). The GEE model testing results revealed that the following two- and three-way

interactions were statistically significant: algorithm x sex (p = .0063), algorithm x region (p < .0001), algorithm x time (p < .0001), algorithm x sex x time (p = .0072), algorithm x time x quintile (p = .0090), and algorithm x time x region (p = .0398). Like the results for the two-year algorithms, these results for three-year algorithms indicate that the RR of change over time in the prevalence of stroke was different across the algorithms, and that this difference also varied across the sociodemographic characteristics of the population.

8.5 Chapter Summary

The results of our validation study indicated that using only a single hospital separation in a single year of data would result in an algorithm with very poor sensitivity to detect non-fatal stroke cases when self-report survey data were used as the gold standard.

This chapter focused on identifying non-fatal cases of stroke in administrative data, and used physician billing claims and prescription drug records, in addition to hospital separations, to identify disease cases. Previous research in this area has examined the effect of using broad or narrow sets of diagnostic codes to identify stroke cases. We compared the results for all validation indices when different sets of diagnostic codes were included in the algorithms.

The results of our validation study indicated that using only a single hospital separation in a single year of data would result in an algorithm with very poor sensitivity to detect non-fatal stroke cases when self-report survey data were used as the gold standard. Including physician billing claims and prescription drug records to identify cases of non-fatal stroke resulted in improved sensitivity but did not result in decreased specificity. Increasing the number of years of administrative data had a large impact on agreement between the two data sources and also on sensitivity. However, the positive predictive value of identifying stroke cases from the administrative data decreased as the number of years of administrative data decreased, and was quite low for the five-year algorithm with the highest sensitivity. Estimates of stroke prevalence for the algorithms with the highest estimates of agreement and the highest values of Youden's index ranged from 2.9% to 3.8%.

Agreement between survey and administrative data was predicted by several characteristics of individuals, including age, sex, the presence of comorbid conditions, and income level. Regression analyses revealed that for both cross-sectional and longitudinal data, the prevalence estimates from different stroke algorithms varied with the sociodemographic characteristics of the population, as well as over time. There was a strong region effect in both the cross-sectional and longitudinal models, which suggests that geographic differences may exist in the rate of hospitalization for stroke or the rate of physician visits for stroke. Surprisingly however, region of residence was not a predictor of agreement between survey and administrative data.

CHAPTER 9: CONCLUSIONS AND RECOMMENDATIONS

9.1 Summary of Findings

The results of this research indicate that administrative data can be used to validly identify cases of asthma, diabetes, and hypertension in Manitoba when survey data are adopted as the gold standard.

This report evaluated the use of Manitoba's administrative data to identify chronic disease cases. This work builds on previous research conducted in Manitoba that investigated the use of administrative data to identify cases of such diseases as inflammatory bowel disease, diabetes, dementia, and depression.

This research adds to the body of literature on validation studies of administrative data in a number of important ways. First, it systematically considered the added value of prescription drug data for identifying disease cases. For some diseases such as asthma, the validation results indicated that prescription drug data were essential to achieve maximum agreement between survey and administrative data, as well as maximum sensitivity. For other diseases, such as diabetes, the advantage associated with using prescription drug data, in addition to hospital and physician data, to identify disease cases was much smaller. This was because most cases could be identified using either of the latter two data sources. Second, this research investigated the characteristics of the individual that were associated with agreement between survey and administrative data. The predictors of agreement varied across the diseases, although age was a statistically significant predictor in all models. Sociodemographic and comorbidity variables were selected for this analysis because it is widely recognized that these factors are associated with health care utilization.

This research also examined the effect that different chronic disease algorithms have on trends in prevalence estimates. For some chronic diseases, like asthma and stroke, the choice of algorithms had important implications for understanding changes in disease prevalence over time. The research also demonstrated that the relative difference in prevalence estimates for different algorithms is associated with the sociodemographic characteristics of the population, including age, sex, region of residence, and income quintile.

Finally, the research has contributed to the literature on methods for identifying chronic disease cases from administrative data because it adopted a systematic approach that used a wide range of validation indices, in addition to a series of inferential analyses to model the factors that influence prevalence estimates derived from chronic disease algorithms applied to administrative data.

The results of this research indicate that administrative data can be used to validly identify cases of asthma, diabetes, and hypertension in Manitoba when survey data are adopted as the gold standard. The overall validity of

administrative data for identifying cases of arthritis, osteoarthritis, coronary heart disease, and stroke was fair to good, and for rheumatoid arthritis it was very poor.

However, these validation results may be influenced by the choice of a gold standard for the research. As indicated in the description of methods, we selected CCHS data as the gold standard because it is the only other source of population-based data in Manitoba for identifying cases of multiple chronic diseases. As well, the large sample size ensured that there were sufficient positive disease cases to validate administrative data algorithms for even rare diseases such as non-fatal stroke and rheumatoid arthritis. However, it is evident from the results of this investigation that there is bias in this data source for chronic disease case identification. Despite the fact that survey data have been used in previous research to validate arthritis algorithms, the self-report question in CCHS cycle 1.1 to identify individuals with rheumatoid arthritis was not effective in distinguishing those with the condition from those with another form of arthritis. More than 8.0% of CCHS respondents in the validation cohort reported a diagnosis of rheumatoid arthritis, which is significantly higher than the prevalence estimates reported using data from other surveys and from clinical registries. Research has previously demonstrated that administrative data can provide valid case identification for rheumatoid arthritis (Lacaille et al., 2005). Therefore, a validation study using an unbiased gold standard data source should be undertaken for rheumatoid arthritis.

9.2 Recommendations on Using the Research Results

This research focused on methods for identifying chronic disease cases from administrative health data. The results of this research will be of greatest benefit to analysts who are preparing reports about chronic disease prevalence in Manitoba, as well as to researchers who are conducting other methodological studies about the use of administrative data for defining disease cases. It also has relevantce to researchers from other jurisdictions who seek to develop chronic disease algorithms that can be applied to their own administrative data.

Analysts can use the results of this research to select one or more algorithms to generate chronic disease prevalence estimates for the Manitoba population. Depending on the goals of future reports, a chronic disease algorithm can be selected based on high agreement between survey and administrative data, high sensitivity to identify positive disease cases, high specificity to avoid identifying false disease cases, or the maximum combination of sensitivity and specificity. Thus, the validation results can be used like a menu, to select the algorithm that is best suited to the goal of a future study of chron-

ic disease prevalence. A high sensitivity algorithm would be selected, for example, if an analyst wanted to generate a prevalence estimate that would capture the maximum number of probable disease cases. On the other hand, a high specificity algorithm would be selected if an analyst wanted to generate a prevalence estimate that would avoid detecting false positive disease cases.

Table 36 summarizes the algorithms with the maximum estimates of κ, sensitivity, specificity, and Youden's index for each chronic disease. Crude provincial prevalence estimates are also provided for each algorithm. For some diseases more than one algorithm had equivalent (or near equivalent) maximum estimates of these validation indices. In this table, we report on the algorithm that had a high value of a statistic but required the fewest number of years of data, the fewest number of data sources, or the fewest number of contacts in administrative data.

The results reported in Table 36 indicate that for rheumatoid arthritis and osteoarthritis, the algorithm with the highest estimate of κ is also the algorithm with the highest sensitivity. For osteoarthritis, asthma, coronary heart disease and stroke, the algorithm with the highest estimate of sensitivity is also the algorithm with the highest value of Youden's index.

The crude provincial prevalence estimates for algorithms with the maximum estimate of κ and Youden's index vary substantially for some diseases. For example, the prevalence estimate for all forms of arthritis is 20.3% for the algorithm with the maximum estimate of κ and 31.5% for the algorithm with the maximum estimate of Youden's index. For coronary heart disease, however, the prevalence estimates for these two algorithms are very similar. This finding is consistent with our literature review results. For many of the diseases, such as arthritis, the estimate of the magnitude of the burden of a chronic disease on a population is strongly influenced by method used to identify disease cases.

Table 36: Crude provincial prevalence estimates for chronic disease algorithms with the maximum estimates of κ, sensitivity, specificity, and Youden's index

Chronic Disease	Algorithm	κ	Sens. (%)	Spec. (%)	Youden's Index	Prev. (%)
Arthritis	1+ H or 2+ P or (1 P & 2+ Rx), 2 yrs	**0.37**	51.7	84.9	0.37	20.3
	1+ P, 5 yrs	0.27	**78.1**	58.6	0.37	47.3
	2+ P, 1 yr	0.24	25.3	**93.8**	0.19	9.2
	1+ H or 2+ P, 5 yrs	0.35	63.7	75.9	**0.40**	31.5
Rheumatoid Arthritis	1+ P, 5 yrs	**0.17**	**11.3**	99.2	**0.11**	1.6
	1+ P, 1 yr	0.12	7.4	**99.8**	0.07	0.7
	1+ H or 2+ P or (1 P & 2+ Rx), 5 yrs	**0.17**	9.4	99.6	**0.11**	1.0
Osteoarthritis	1+ P, 5 yrs	**0.32**	**49.9**	88.7	**0.39**	13.2
	2+ P, 1 yr	0.16	12.3	**98.5**	0.11	2.3
Asthma (All Ages)	1+ H or 2+ P or 2+ Rx, 5 yrs	**0.59**	75.4	94.2	0.70	11.6
	1+ H or 1+ P or 1+ Rx, 5 yrs	0.50	**84.3**	88.6	**0.73**	17.5
	2+P, 1 yr	0.27	18.1	**99.5**	0.18	1.9
Coronary Heart Disease	1+ H or 2+ P or (1 P & 2+ Rx), 3 yrs	**0.55**	60.1	96.6	0.60	5.8
	1+ H or 1+ P, 5 yrs	0.53	**67.9**	95.0	**0.63**	7.6
	2+ P, 1 yr	0.37	28.6	**98.8**	0.27	2.4
Diabetes	1+ H or 1+ P or 2+ Rx, 2 yrs	**0.86**	86.1	99.2	0.85	6.3
	1+ H or 1+ P or 1+ Rx, 3 yrs	0.76	**90.5**	97.3	0.88	8.2
	1+ H or 2+ P, 1 yr	0.73	63.2	**99.5**	0.63	4.4
	1+ H or 1+ P or 1+ Rx, 2 yrs	0.80	89.6	97.9	**0.88**	7.5
Hypertension	1+ H or 1+ P or 2+ Rx, 1 yr	**0.70**	89.0	89.9	0.79	21.6
	1+ H or 1+ P or 1+ Rx, 3 yrs	0.62	**92.8**	84.0	0.77	27.4
	2+ P, 1 yr	0.54	48.4	**97.5**	0.46	9.3
	1+ H or 1+ P or 1+ Rx, 1 yr	0.68	89.9	88.7	**0.79**	22.5
Stroke	1+ H or 2+ P or (1 P & 2+ Rx), 5 yrs[a]	**0.47**	61.5	98.3	0.60	2.9
	1+ H or 1+ P, 5 yrs[a]	0.43	**66.1**	97.5	**0.64**	3.8
	1+ H, 1 yr[b]	0.12	7.3	**99.8**	0.07	0.2

Note: Values in bold are the maximum κ, sensitivity, specificity, or Youden's index values. All prevalence estimates are defined for the population 19 years of age and older except for asthma, which is defined for the population 12 years of age and older.

[a]Algorithm is based on the sensitive set of ICD-9-CM codes for stroke (430 – 438)

[b]Algorithm is based on the specific set of ICD-9-CM codes for stroke (430, 431, 434, 435, 436).

Source: Manitoba Centre for Health Policy, 2006

9.3 Future Research Opportunities

This section focusses on future research opportunities in the following areas: (1) chronic disease validation studies, and (2) research on methods to identify chronic disease cases from administrative data.

Chronic Disease Validation

Validation data were not available at the time this study was initiated for two chronic conditions that were of interest to the Working Group. These were renal disease and CHF. CCHS contains one question about the preva-

lence of CHF, but the number of positive cases was too small to conduct a validation analysis. CCHS does not contain any questions regarding renal disease. Other validation sources, including disease-specific clinical registries maintained in Manitoba, could also be investigated for their potential to be included in future studies.

Tables A.7 and A.8 in Appendix A summarize the results of our review of literature on methods to identify cases of CHF and renal disease from administrative data. While the number of validation studies is small, the results of the literature review indicate that in other jurisdictions, administrative data have been used to identify cases of both diseases.

The chronic disease algorithms investigated in this report could be validated using other data sources, including other survey data, clinical registry data, or chart review data. Further enhancement of population-based health survey data may contribute to the accuracy of future validation studies involving administrative data. The adoption of questionnaire content that will result in accurate information about diagnosed chronic diseases, date(s) of diagnosis, and specific types or forms of a disease is critical.

Methods to Identify Chronic Disease Cases from Administrative Data
As this report has demonstrated, the sensitivity and specificity of chronic disease algorithms and the agreement between survey and administrative data is influenced by several elements of administrative data, including the type of data source, the number of years of data, and the choice of diagnostic codes used to construct the algorithm. Sensitivity, specificity and agreement may also vary with sociodemographic characteristics of the population, such as age and region of residence. It is an overwhelming task to test all possible combinations of data features to identify the algorithm with the maximum sensitivity and specificity or agreement.

Pattern classification models (Duda et al., 2000), based on probability and decision theory, may be advantageous for identifying disease cases in administrative data. The models, which include neural networks, classification trees, support vector machines, and nearest neighbour methods, have been applied to classification problems like risk scoring for academic failure, speech recognition, medical diagnostics, and clinical decision rules (e.g., Shanker et al., 2000). These models search for patterns or associations among data features (i.e., predictor variables) to identify clusters of individuals or objects. A large set of data features can initially be included in the model, and tests of statistical significance conducted to systematically assess their contribution to the prediction model. The significance of interactions among predictor variables can also be tested. Estimates of the probability of misclassification can be obtained.

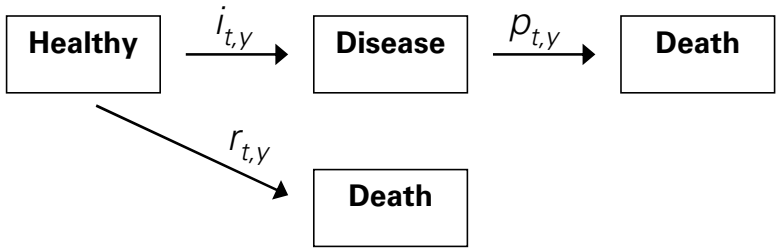
Previous research has demonstrated that no one classification model is uniformly superior (Balakrishnan et al., 1994; Song et al., 2004). The models rest on different assumptions about the data distribution, the nature of the association between the outcome and predictor variables, and the distribution of the predictor variables. They also vary in their sensitivity to the amount of noise (i.e., measurement error) in the data. Research is presently being undertaken by the first author to examine the feasibility of applying pattern classification models to administrative data for identifying cases of chronic disease. This research will also identify the techniques that provide optimal performance identifying disease cases.

Chronic Disease Incidence

Previous research has used administrative data to estimate chronic disease incidence as well as prevalence (Blanchard et al., 1996; Ellekjaer et al., 1999). The point in time at which the selected disease algorithm is first satisfied in administrative data is typically used to approximate the date of disease onset. This empirical method of estimating incidence requires access to multiple years of linked person-specific longitudinal data and is sensitive to the length of the study observation period. Estimates at the beginning of the study period may exhibit substantial bias compared to those later in the period. Brameld et al. (2003) developed a retrograde survival model for censored data to adjust empirical incidence estimates for bias due to variations in the length of the observation period. This model has only been tested with linked hospital data. It represents a possible gold standard for deriving incidence estimates from administrative data, but comparison with alternate methods is needed.

Another approach to estimate chronic disease incidence is to use health state models. These models only require aggregate data on prevalence and mortality to estimate incidence (Hill et al., 1999). Figure 59 depicts a simple three-state model. All individuals in the population are initially assumed disease free, and to transition either to the disease state or death. The incidence rate (i) is the transition from the healthy to diseased state and the mortality rate (r) is the transition rate from disease to death. Rates of disease-specific, all-cause mortality, and prevalence are used as model inputs. Incidence rates, assumed to co-vary with age and calendar year, are generated as outputs. Both deterministic models, based on the theory of differential equations, and stochastic models, which assume random variation in the relationships among the health states, have been applied to chronic disease data (Podgor and Leske, 1986; Brookmeyer and Gray, 2000). Health state models are appealing for predicting possible trends in incidence using a range of input values. However, the accuracy of these models in relation to empirical and other model-based estimates requires evaluation.

Figure 59: Three-State Model and the Rate of Transition Between States at Age *t* in Year *y*



Source: Manitoba Centre for Health Policy, 2006

9.4 Conclusions

Administrative data are an accessible and inexpensive source of data for monitoring the prevalence of chronic diseases.

Public health analysts and researchers at provincial and regional levels require tools that can be used to develop new chronic disease surveillance systems or enhance existing systems. Table 37 lists some of the major uses of surveillance systems, including detection of new public health problems and identifying risk factors.

Table 37: Uses of public health surveillance data

Uses
Providing quantitative estimates of the magnitude of a health problem
Detecting emergent health problems and epidemics
Documenting the distribution and spread of a health event geographically or among defined populations
Testing hypotheses
Facilitating planning
Facilitating epidemiologic and laboratory research
Describing the natural history of a condition
Monitoring change in risk factors for health-event occurrence
Detecting changes in health practices
Assessing control and prevention activities

Note: Table is reproduced from Thacker et al. (1995)

Source: Manitoba Centre for Health Policy, 2006

Administrative data are an accessible and inexpensive source of data for monitoring the prevalence of chronic diseases and have been incorporated into chronic disease surveillance systems for some diseases in Manitoba as well as in other jurisdictions. Manitoba benefits from the availability of population-based hospital separations, physician billing claims, and prescription drug records. This research demonstrates these three sources can be used to generate valid estimates of the prevalence of multiple chronic diseases for the entire population.

GLOSSARY

Administrative Data / Databases. Data collected, usually by government, for some administrative purpose (e.g., keeping track of the population eligible for certain benefits, paying doctors or hospitals), but not primarily for research or surveillance purposes.

Anatomical Therapeutic Chemical (ATC) Classification. The ATC system is maintained by the World Health Organization (WHO) Collaborating Centre for Drug Statistics Methodology, and was first published in 1976. Under this system, drugs are divided into groups at each of five levels according to the organ or system on which they act and/or their therapeutic and chemical characteristics. The levels are: (1) anatomical group, (2) therapeutic main group, (3) therapeutic/pharmacological subgroup, (4) chemical/therapeutic/pharmacological subgroup, and (5) subgroup for chemical substance.

Arthritis. (from Greek *arthro-*, joint + *-itis*, inflammation) A group of conditions that affect the health of the bone joints in the body. One in three adult Americans suffer from some form of arthritis and the disease affects about twice as many women as men. Arthritic diseases include rheumatoid arthritis and psoriatic arthritis, which are autoimmune diseases; septic arthritis, caused by joint infection; and the more common osteoarthritis, or degenerative joint disease. Arthritis can be caused from strains and injuries caused by repetitive motion, sports, overexertion, and falls. Unlike the autoimmune diseases, osteoarthritis largely affects older people and results from the degeneration of joint cartilage. See Table 1 for how arthritis was defined for this study.

- **Rheumatoid arthritis** (**RA**) is a chronic, inflammatory autoimmune disorder that causes the immune system to attack the joints. It is a disabling and painful inflammatory condition, which can lead to substantial loss of mobility due to pain and joint destruction. The disease is also systemic in that it often also affects many extra-articular tissues throughout the body including the skin, blood vessels, heart, lungs, and muscles.
- **Osteoarthritis** (**OA**) also known as degenerative arthritis or degenerative joint disease, and sometimes referred to as "arthrosis" or "osteoarthrosis"), is a condition in which low-grade inflammation results in pain in the joints, caused by wearing of the cartilage that covers and acts as a cushion inside joints. As the bone surfaces become less well protected by cartilage, the patient experiences pain upon weight bearing, including walking and standing. Due to decreased movement because of the pain, regional muscles may atro-

phy, and ligaments may become more lax. OA is the most common form of arthritis. The word is derived from the Greek word "*osteo*", meaning "of the bone", "*arthro*", meaning "joint", and "*itis*", mean ing inflammation, although many sufferers have little or no inflammation.

Asthma. A disease in which inflammation of the airways causes airflow into and out of the lungs to be restricted. See Table 1 for how asthma was defined for this study.

Canadian Community Health Survey (CCHS). The CCHS was conducted by Statistics Canada to provide regular and timely cross-sectional estimates of health determinants, health status and health system utilization for 136 health regions in Canada, including the territories. Survey respondents were sampled from 11 regions in Manitoba. Respondents were 12 years of age and older; the sampling methodology was designed to ensure over-representation of youth under 19 years of age and seniors 65 years of age and older.

Chronic Disease. Chronic diseases are those conditions that are generally incurable, are often caused by a complex interaction of factors, and usually have a prolonged clinical course.

Confidence Interval (CI) / Limits. An interval, calculated from data, which contains a population parameter, such as the population median or mean, with specified probability. For example, a 95% Confidence Interval (written as 95% CI) would have a 95% probability of containing the true population value.

Congestive Heart Failure (CHF). Also called congestive cardiac failure (CCF) or just heart failure, is the inability of the heart to pump a sufficient amount of blood throughout the body, or requiring elevated filling pressures in order to pump effectively. CHF is an abnormal cardiac condition that reflects impaired cardiac pumping and blood flow. The pooling of blood leads to congestion in body tissue.

Coronary Heart Disease (CHD). Coronary heart disease (CHD), also called coronary artery disease (CAD), ischemic heart disease, or atherosclerotic heart disease, is the end result of the accumulation of atheromatous plaques within the walls of the arteries that supply the myocardium (the muscle of the heart). While the symptoms and signs of coronary heart disease are noted in the advanced state of disease, most individuals with coronary heart disease show no evidence of disease for decades as the disease progresses before the first onset of symptoms, often a "sudden" heart attack, finally arise. After decades of progression, some of these atheromatous plaques may rupture and (along with the activation of the blood clotting

system) start limiting blood flow to the heart muscle. The disease is the most common cause of sudden death. See Table 1 for how CHD was defined for this study.

Cross-Sectional Study. A study that examines the relationship between diseases (or other health-related characteristics) and other variables of interest as they exist in a defined population at one point in time. The presence or absence of disease and the presence or absence of the other variables (or, if they are quantitative, their level) are determined in each member of the study population or in a representative sample at one particular time. The temporal sequence of cause and effect cannot necessarily be determined in a cross-sectional study. Consequently, disease prevalence rather than incidence is normally recorded in a cross-sectional study.

Descriptive Analyses. Descriptive statistics are procedures for summarizing, organizing, graphing, and, in general, describing quantitative information. Often contrasted with inferential statistics, which is used to make inferences about a population based on information about a sample drawn from that population.

Diabetes. A chronic condition in which the pancreas no longer produces enough insulin (Type I Diabetes) or when cells stop responding to the insulin that is produced (Type II Diabetes), so that glucose in the blood cannot be absorbed into the cells of the body. The most common endocrine disorder, Diabetes Mellitus affects many organs and body functions, especially those involved in metabolism, and can cause serious health complications including renal failure, heart disease, stroke, and blindness. See Table 1 for how diabetes was defined for this study.

Drug Identification Number (DIN). An 8 digit number, assigned by the Therapeutic Products Directorate of Health Canada, to each drug approved for use in Canada in accordance with the Food and Drug Regulation. The same drug (e.g. Amoxicillin, 250 mg capsules) can have several different DINs associated with it (due to different manufacturers).

Generalized Estimating Equations (GEE). A method of estimation used in the analysis of longitudinal data, which consists of repeated measures of an individual or cluster of individuals over time. These repeated measures from any one individual or cluster are correlated with each other and are therefore no longer independent. GEEs use the data to estimate the correlation between a single individual or cluster's response and provide a correct estimate of each effect's variance.

Generalized Linear Model (GLM). A unified class of models for regression analysis of independent observations of a discrete or continuous response. A characteristic feature of generalized linear models is that a suitable non-linear transformation of the mean response is a linear function of the convariates. Generalized linear models provide a unified method for analyzing diverse types of univariate responses (e.g., continuous, binary, counts). Generalized linear models are actually a collection of regression models and they include as special cases the standard linear regression for normally distributed continuous outcomes, logistic regression models for a binary outcome, or Poisson regression models for counts.

Hospital Discharge Database. Hospital abstracts are completed at the point of discharge for all separations from acute care facilities in Manitoba. They include up to 16 diagnosis codes based on the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM).

Hypertension. Primary hypertension is often referred to as high blood pressure. The "tension" in hypertension describes the vascular tone of the smooth muscles in the artery and arteriole walls. It accounts for over 90 percent of all cases of hypertension in the U.S. and develops without apparent causes. Hypertension is a major health problem, especially because it often has no symptoms. If left untreated, hypertension can lead to heart attack, stroke, enlarged heart, or kidney damage. See Table 1 for how hypertension was defined for this study

Incidence. The number of new cases of a specific disease / condition / event over a specified time period. The incidence rate uses new cases in the numerator; individuals with a history of the disease / condition are not included. The denominator for incidence rates is the population at risk. Even though individuals who have already developed the condition should be eliminated from the denominator, incidence rates are often expressed based on the average population rather than the population at risk. In the case of chronic conditions, where most people appear to be at risk, the distinction between populations at risk and the whole population appears to be less critical.

Inferential Analysis. Inferential statistics allow one to draw conclusions or inferences from data. Usually this means coming to conclusions about a population on the basis of data describing a sample. Statistical inference uses probability and information about a sample to draw conclusions ("inferences") about a population or about how likely it is that a result could have been obtained by chance.

Interaction Effect. The joint effect of two or more independent variables on a dependent variable. Interaction effects occur when independent variables not only have separate effects but also have combined effects on a dependent

variable. Put somewhat differently, interaction effects occur when the relation between two variables differs depending on the value of another variable.

Kappa. (κ), a measure of agreement between two sources, each of which is measured on a binary scale (i.e., disease present/absent).

Likelihood Ratio Test (LRT). As the name implies, the likelihood ratio is a ratio of two likelihoods. It is widely used as a test statistic, perhaps especially for relations among categorical variables displayed in contingency tables. The smaller the LR, the stronger the relationship. This is because (in comparison with the chi-square method) with the LR we attempt to accept a particular model, not reject a null hypothesis.

Logistic Regression. A regression model for describing the relationship between a response variable and one or more explanatory variables where the response variable follows a binomial distribution. Logistic regression is used to model the probability of occurrence of a binary or dichotomous outcome.

Longitudinal Study. A longitudinal survey describes or measures a population at several points in time.

Main Effect. The simple effect of an independent variable on a dependent variable; the effect of an independent variable uninfluenced by other variables. Used in contrast with the interaction effect of two or more independent variables on a dependent variable. There is some controversy about whether it is appropriate to try to interpret main effects in the presence of interaction effects.

Manitoba Formulary. The Manitoba Drug Benefits and Interchangeability Formulary lists therapeutically effective drugs of proven high quality that have been approved as eligible benefits under the Pharmacare drug benefit program. It also includes a list of interchangeable drugs. It is compiled with the advice of the Manitoba Drug Standards and Therapeutics Committee, assisted by Manitoba Health staff and outside consultants. The Minister of Health gives the final approval for benefits under the Pharmacare drug benefit program.

Manitoba Health Services Insurance Plan (MHSIP). The health insurance plan provided by Manitoba Health. It is financed from general revenues of the Province of Manitoba and with funds provided by the Government of Canada.

Mantel-Haenszel Test. Two groups are compared on a binary response, adjusting for control variables and applied to K strata of 2 x 2 tables where response totals are treated as fixed. Under the null hypothesis of conditional independence, this statistic has approximately a chi-squared distribution with df=1.

Negative Binomial Regression. Regression analyses for data that follows a negative binomial distribution, which occurs when an event is relatively rare, but is highly variable over the entire population.

Negative Predictive Value (NPV). The negative predictive value of a test is the probability that the patient will not have the disease when restricted to all patients who test negative. You can compute the negative predictive value as NPV = TN / (TN + FN) where TN and FN are the number of true negative and false negative results, respectively. Notice that the denominator for negative predictive value is the number of patients who test negative.

Odds Ratio. The ratio of the odds of an event occurring in one group to the odds of it occurring in another group, or to a data-based estimate of that ratio. These groups might be men and women, an experimental group and a control group, or any other dichotomous classification.

Offset Variable. An offset variable is used to adjust Poisson and Negative Binomial models for differential "exposure" in data records (e.g., different lengths of time periods, populations).

Physician Claims. These are the claims that are submitted to the provincial government by individual physicians for services they provide. Fee-for-service physicians receive payment based on these claims, while those submitted by salaried physicians are only for administrative purposes. The physician claims data file is part of the Population Health Research Data Repository.

Poisson Distribution. A probability density function that is often used as a mathematical model of the number of random events in a suitable interval of time and space, that has its mean equal to its variance, and that is used as an approximation to the binomial distribution. The distribution function has the form: $P(X = x) = e^{-\lambda}\lambda^{x} / x!$ where P is the probability that X is some value x, which acquires non-negative integral values, and λ is the mean number of occurrences in the specified interval.

Poisson Regression. Regression analyses for data that follow a Poisson distribution. Poisson regression is often the best choice for modeling counts of rare events, such as death.

Population Health Research Data Repository (PHRDR). A comprehensive database developed to describe and explain patterns of health care and profiles of health and illness. It is located at the Manitoba Centre for Health Policy (MCHP). The database contains anonymized encounter-based records of individual's interactions with the health care system, including physicians, hospitals, nursing homes, home care, and pharmaceutical prescriptions. The Repository also includes data from other agencies, for example, Statistics Canada data at the level of enumeration area. Subsets of the data are used in specific approved research projects.

Population Registry. Refers to the Research Registry, which contains data on the insured population organized by family registration numbers. The research registry contains information on dates of coverage, age, sex, and place of residence (by postal code and municipal code only; no addresses are contained in the file). Annual snapshots of this data have been received since 1970. Information on marital status has been constructed from the family registration information. A massive programming effort maintained over many years has joined these snapshot files together such that individual histories can be constructed over the entire period of the data base. This results in the creation of the longitudinal population registry; many checks have been done on this registry. Software has been developed to facilitate longitudinal follow-up or mobility, migration, and mortality.

Population Surveillance. Langmuir, one of the originators of the modern concept, defined population surveillance in 1963 as 'the continued watchfulness over the distribution and trends of incidence through the systematic collection, consolidation and evaluation of morbidity and mortality reports and other relevant data'. He pointed out that intrinsic to the concept of surveillance is the regular dissemination of information derived from the data to all who require it.

Positive Predictive Value (PPV). The positive predictive value of a test is the probability that the patient has the disease when restricted to those patients who test positive. This term is sometimes abbreviated as PPV. You can compute the positive predictive value as PPV = TP / (TP + FP) where TP and FP are the number of true positive and false positive results, respectively. Notice that the denominator for positive predictive value is the number of patients who test positive.

Prescription Drug Database - Drug Programs Information Network (DPIN). An electronic, on-line, point-of-sale prescription drug database. It links all community pharmacies (but not hospitals or nursing care homes/personal care homes [PCHs]) and captures information about all

Manitoba residents, including most prescriptions dispensed to status Indians. DPIN contains information such as unique patient identification, age, birth date, sex, medication history, over-the-counter medication history, patient postal code, new drug prescribed, date dispensed, and unique pharmacy identification number. DPIN is maintained by the Government of Manitoba's Ministry of Health.

Prevalence. The term prevalence refers to the proportion of the population that 'has' a given disease at a given time. The measure of a condition in a population at a given point in time is referred to as point prevalence. A second type of prevalence is called period prevalence. Over a period of time, such as five years, this measures the number of individuals with a particular condition in the population during that time period. Period prevalence is the most common measure of prevalence used in MCHP studies. Prevalence data provide an indication of the extent of a condition and may have implications for the provision of services needed in a community.

- **Period Prevalence.** The measure of a disease or condition in a population during a given point in time. It is a combination of point prevalence and incidence.
- **Point Prevalence.** The measure of a disease or condition in a population at a given point in time.

Regional Health Authority (RHA). In 1997, Manitoba established 11 RHAs as governance structures for northern and rural health services: South Eastman, South Westman, Brandon, Central, Marquette, Parkland, North Eastman, Interlake, Burntwood, Norman and Churchill. Winnipeg was originally divided into 2 additional authorities: the Winnipeg Community and Long Term Care Authority and the Winnipeg Hospital Authority. Each RHA has the responsibility for providing for the delivery and administration of health services in a specified geographic area.

Renal Disease. The kidneys are bean-shaped excretory organs. Part of the urinary system, the kidneys filter wastes (especially urea) from the blood and excrete them, along with water, as urine. The adjective meaning "kidney-related" is renal, from the Latin. Any diseases that affect the blood vessels, including diabetes, high blood pressure, and atherosclerosis (hardening of the arteries), can impair the kidneys' ability to filter blood and regulate fluids in the body. Disease and infection in other parts of the body can also trigger a kidney disorder. Because kidney impairment can be life-threatening, disorders and diseases that may affect the kidney deserve prompt attention. Kidney disease often causes no symptoms until late in its course and can lead to end-stage kidney failure, which is fatal unless a dialysis machine is used or a kidney transplant is performed. There are more than 100 disor-

ders, diseases, and conditions that can lead to progressive destruction of the kidneys.

Respondent Attrition. Loss of subjects (e.g,. respondents answering questions in a survey or interview) over the course of the research project. Attrition may be a source of bias if the subjects who are lost make the sample less representative of the population.

Sensitivity. One of two indices used to evaluate the accuracy of a test that predicts dichotomous outcomes (e.g. logistic regression). It is the number of "true positives" (those testing positive who have the disease), divided by all those with the disease.

Specificity. One of two indices used to evaluate the accuracy of a test that predicts dichotomous outcomes (e.g. logistic regression). It is the number of "true negatives" (those testing negative who do not have the disease), divided by all those without the disease.

Standard Error. In statistics, the standard error of a measurement, value or quantity is the standard deviation of the process by which it was generated, after adjusting for sample size. In other words the standard error is the standard deviation of the sample mean. The standard error of a sample from a population is the standard deviation of the sampling distribution and may be estimated by the formula: σ/n where σ is the standard deviation of the population distribution and n is the size (number of items) in the sample.

Stroke. A stroke occurs when there is a sudden death of brain cells due to a lack of oxygen when the blood flow to the brain is impaired by blockage or rupture of an artery to the brain. Symptoms of a stroke depend on the area of the brain affected. The most common symptom is weakness or paralysis of one side of the body with partial or complete loss of voluntary movement or sensation in a leg or arm. Other common symptoms include speech problems, weak facial muscles, numbness and tingling. A stroke involving the base of the brain can affect balance, vision, swallowing, breathing and consciousness.

Survey data. Collected through a research design in which a sample of subjects is drawn from a population and studied (usually interviewed) to make inferences about the population.

Validity. In statistics a valid measure is one which is measuring what it is supposed to measure. Validity implies reliability (accuracy). A valid measure must be reliable, but a reliable measure need not be valid. Validity refers to getting results that accurately reflect the concept being measured.

Venn diagram. A Venn diagram is an illustration of the relationships between and among sets, groups of objects that share something in common. Usually, Venn diagrams are used to depict set intersections (denoted by an upside-down letter U). This type of diagram is used in scientific and engineering presentations, in theoretical mathematics, in computer applications, and in statistics.

Vital Statistics. A Manitoba government department responsible for keeping records and registries of all births, deaths, marriages and stillbirths that take place in Manitoba.

Wald Test. The Wald statistic represents the square of the ratio between the regression coefficient and its standard error. This statistic follows a χ^2 distribution with one degree of freedom, which is equal to the standard normal distribution squared.

Youden's Index. The index is defined as sensitivity + specificity – 1, where sensitivity and specificity are calculated as proportions. Youden's index has minimum and maximum values of –1 and +1, respectively, with a value of +1 representing the optimal value for an algorithm.

REFERENCES

Altman DG. (1991). *Practical statistics for medical research*. London: Chapman & Hall.

Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. *Am Heart J* 2002; 144(2):290-296

Badley EM, DesMeules M, (eds.). *Arthritis in Canada: An Ongoing Challenge*. Ottawa, ON: Health Canada, 2003.

Badley EM, Wang, PP. Arthritis and the aging population: projections of arthritis prevalence in Canada 1991 to 2031. *J Rheumatol* 1998;25(1):138-144.

Balakrishnan PV, Cooper MC, Jacob VS, Lewis PA. A study of the classification capabilities of neural networks using unsupervised learning: A comparison with k-means clustering. *Psychometrika* 1994;59(4):509-525.

Benesch C, Witter DM, Wilder AL, Duncan PW, Samsa GP, Matchar DB. Inaccuracy of the International Classification of Disease (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. *Neurology* 1997;49(3):660-664.

Blanchard JF, Ludwig S, Wajda A, Dean H, Anderson K, Kendall O, Depew N. Incidence and prevalence of diabetes in Manitoba, 1986-1991. *Diabetes Care* 1996; 19(8):807-811.

Blanchard JF, Dean H, Anderson K, Wajda A, Ludwig S, Depew N. Incidence and prevalence of diabetes in children aged 0-14 years in Manitoba, Canada, 1985-1993. *Diabetes Care* 1997;20(4):512-515.

Bolen J, Sniezek J, Theis K, Helmick C, Hootman J, Brady T, Langmaid G. Racial/ethnic differences in the prevalence and impact of doctor-diagnosed arthritis --- United States, 2002. *MMWR* 2005;54(5):119-123.

Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR. Identifying hypertension-related comorbidities from administrative data: What's the optimal approach? *Am J Med Qual* 2004;19(5):201-206.

Brameld KJ, Holman CD, Lawrence DM, Hobbs MS. Improved methods for estimating incidence from linked hospital morbidity data. *Int J Epidemiol* 2003;32(4):617-624.

Brameld KJ, Thomas MA, Holman CD, Bass AJ, Rouse IL. Validation of linked administrative data on end-stage renal failure: application of record linkage to a 'clinical base population'. *Aust N Z J Public Health* 1999;23(5):464-467.

Brookmeyer R, Gray S. Methods for projecting the incidence and prevalence of chronic diseases in aging populations: application to Alzheimer's disease. *Stat Med* 2000;19:1481-1493.

Chronic Disease Prevention Alliance of Canada (CDPAC). *The Case for Change*. Available from URL: http://www.cdpac.ca/content/case_for_change/case_for_change.asp. Accessed on: January 19, 2006.

Chen Y, Johansen H, Thillaiampalam S, Sambell C. Asthma. *Health Reports* 2005 March; 16(2):43-53.

Clarke JL, Nash DB. The effectiveness of heart failure disease management: Initial findings from a comprehensive program. *Disease Management* 2002;5(4):215-223.

Cricelli C, Mazzaglia G, Samani F, Marchi M, Sabatini A, Nardi R, Ventriglia G, Caputi AP. Prevalence estimates for chronic diseases in Italy: Exploring the differences between self-report and primary care databases. *J Public Health Med* 2003;25(3):254-257.

Cujec B, Jin Y, Quan H, Johnson D. The province of Alberta, Canada avoids the hospitalization epidemic for congestive heart failure patients. *Int J Cardiol* 2004;96(2):203-210.

Dey AN, Bloom B. Summary health statistics for U.S. children: National Health Interview Survey, 2003. *Vital Health Stat* 2005;10(223). Available from URL: http://www.cdc.gov/nchs/data/series/sr_10/sr10_223.pdf.

Dobkin B. *The Clinical Science of Neurologic Rehabilitation. 2nd ed.* New York, NY: Oxford, 2003:375-376.

Duda RO, Hart PE, Stork DG. *Pattern Classification. 2nd ed.* New York, NY: Wiley, 2000.

Ellekjaer H, Holmen J, Kruger O, Terent A. Identification of incident stroke in Norway: Hospital discharge data compared with a population-based stroke register. *Stroke* 1999;30:56-60.

Erzen D, Roos LL, Manfreda J, Anthonisen NR. Changes in asthma severity in Manitoba. *Chest* 1995;108(1):16-23.

Feldman DE, Thivierge C, Guerard L, Dery V, Kapetanakis C, Lavoie G, Beck EJ. Changing trends in mortality and admissions to hospital for elderly patients with congestive heart failure in Montreal. *CMAJ* 2001;165(8):1033-1036.

Fitzmaurice G M, Laird N M, Ware J H. *Applied Longitudinal Analysis.* Hoboken, NJ: John Wiley & Sons, 2004.

Fowles J B, Fowler E J, Craft C. Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. *J Ambul Care Manage* 1998;21(1):24-34.

Frieden TR. Asleep at the switch: local public health and chronic disease. *Am J Public Health* 2004;94(12):2059-2061.

Goff DC, Jr., Pandey DK, Chan FA, Ortiz C, Nichaman MZ. Congestive heart failure in the United States: is there more than meets the I(CD code)? The Corpus Christi Heart Project. *Arch Intern Med* 2000;160(2):197-202.

Hajjar I, Kotchen TA. Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988-2000. *JAMA* 2003;290(2):199-206.

Hansell A, Hollowell J, McNiece R, Nichols T, Strachan D. Validity and interpretation of mortality, health service and survey data on COPD and asthma in England. *Eur Respir J* 2003;21(2):279-286.

Haroun MK, Jaar BG, Hoffman SC, Comstock GW, Klag MJ, Coresh J. Risk factors for chronic kidney disease: A prospective study of 23,534 men and women in Washington County, Maryland. *J Am Soc Nephrol* 2003;14(11):2934-2941.

Harrold LR, Yood RA, Andrade SE, Reed JI, Cernieux J, Straus W, Weeks M, Lewis B, Gurwitz JH. Evaluating the predictive value of osteoarthritis diagnoses in an administrative database. *Arthritis Rheum* 2000;43(8):1881-1885.

Health Canada. (n.d.). *It's Your Health: Type 2 Diabetes.* Available from URL: http://www.hc-sc.gc.ca/iyh-vsv/diseases-maladies/diabete_e.html. Accessed on: Sept 21, 2005.

Health Canada. *Diabetes in Canada, 2nd ed*. Ottawa, ON: Centre for Chronic Disease Prevention and Control, Population and Public Health Branch, 2002.

Health Surveillance Coordination Division. *Chronic Disease Surveillance in Canada: A Background Paper.* Ottawa, ON: Centre for Surveillance Coordination Population and Public Health Branch, Health Canada, 2003.

Heart and Stroke Foundation of Canada. *The Changing Face of Heart Disease and Stroke in Canada*. Ottawa, ON: Centre for Chronic Disease Prevention and Control Cardiovascular Disease, 2000.

Hill GB, Forbes WF, Kozak J. A simple method for estimating incidence from prevalence. *Chronic Dis Can* 1999;20(4):151-153.

Hootman JM, Helmick CG, Schappert SM. Magnitude and characteristics of arthritis and other rheumatic conditions on ambulatory medical care visits, United States, 1997. *Arthritis Rheum* 2002;47(6):571-581.

Huff L, Bogdan G, Burke K, Hayes E, Perry W, Graham L, Lentzner H. Using hospital discharge data for disease surveillance. *Public Health Rep* 1996;111(1):78-81.

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.

Huzel L, Roos LL, Anthonisen NR, Manfreda J. Diagnosing asthma: the fit between survey and administrative database. *Can Respir J* 2002;9(6):407-412.

Institute for Clinical Evaluative Sciences. *Cardiovascular Health and Services in Ontario: An ICES Atlas*. Toronto, ON: ICES, 1999.

Jollis JG, Ancukiewicz M, Delong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information-systems - Implications for outcomes research. *Ann Intern Med* 1993;119(8):844-850.

Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kieval RI, Lindsey SM, Roberts WN, Sheff DM, Spencer RT, Weaver AL, Baron JA. Sensitivity and positive predictive value of Medicare Part B physician claims for rheumatologic diagnoses and procedures. *Arthritis Rheum* 1997; 40(9):1594-1600.

Kokotailo RA and Hill MD. Coding of stroke and stroke risk factors using International Classification of Diseases, Revisions 9 and 10. *Stroke* 2005;36:1776-1781.

Kozyrskyj AL, Mustard CA, Becker AB. Identifying children with persistent asthma from health care administrative records. *Can Respir J* 2004;11(2):141-145.

Lacaille D, Anis AH, Guh DP, Esdaile JM. Gaps in care for rheumatoid arthritis: A population study. *Arthritis Rheum* 2005;53(2):241-248.

Lee DS, Donovan L, Austin PC, Gond T, Liu PP, Rouleau JL, Tu JV. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care* 2005;43(2):182-188.

Leibson CL, Maessens JM, Brown RD, Whisnant JP. Accuracy of hospital discharge abstracts for identifying stroke. *Stroke* 1994;25(12):2348-2355.

Leppala JM, Virtamo J, Neinonen OP. Validation of stroke diagnosis in the national hospital discharge register and the register of causes of death in Finland. *Euro J of Epidemiol* 1999;15:155-160.

Li SQ, Cunningham J, Cass A. Renal-related deaths in Australia 1997-1999. *Intern Med J* 2004;34(5):259-265.

Losina E, Barrett J, Baron JA, Katz JN. Accuracy of Medicare claims data for rheumatologic diagnoses in total hip replacement recipients. *J Clin Epidemiol* 2003;56(6):515-519.

Macy E, Schatz M, Gibbons C, Zeiger R. The prevalence of reversible airflow obstruction and/or methacholine hyperreactivity in random adult asthma patients identified by administrative data. *J Asthma* 2005;42(3):213-220.

Mahonen M, Salomaa V, Brommels M, Molarius A, Miettinen H, Pyorala K, Tuomilehto J, Arstila M, Kaarsalo E, Ketonen M, Kuulasmaa K, Lehto S, Mustaniemi H, Niemela M, Palomaki P, Torppa J, Vuorenmaa T. The validity of hospital discharge register data on coronary heart disease in Finland. *Euro J of Epidemiol* 1997;(13):403-415.

Maio V, Yuen E, Rabinowitz C, Louis D, Jimbo M, Donatini A, Mall S, Taroni F. Using pharmacy data to identify those with chronic conditions in Emilia Romagna, Italy. *J Health Serv Res Policy* 2005;10(4):232-238.

Martens P, Burchill C, Fransoo R, De Coster C, McKeen N, Ekuma O, *The Need to Know* Team, Prior H, Chateau D, Burland E, Robinson R, Jebamani L, Metge C. *Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study.* Winnipeg, MB: Manitoba Centre for Health Policy, 2004.

Martens PJ, Fransoo R, *The Need to Know* Team, Burland E, Jebamani L, Burchill C, Black C, Dik N, MacWilliam L, Derksen S, Walld R, Steinbach C, Dahl M. *The Manitoba RHA Indicators Atlas: Population-Based Comparisons of Health and Health Care Use*. Winnipeg, MB: Manitoba Centre for Health Policy, 2003.

Martin LM, Leff M, Calonge N, Garrett C, Nelson DE. Validation of self-reported chronic conditions and health services in a managed care population. *Am J Prev Med* 2000;18(3):215-218.

Martinez-Selles M, Garcia Robles JA, Prieto L, Serrano JA, Munoz R, Frades E, Almendral J. Annual rates of admission and seasonal variations in hospitalizations for heart failure. *Eur J Heart Fail* 2002;4(6):779-786.

Mayo ME, Chockalingam A, Reeder BA, Phillips S. Surveillance for stroke in Canada. *Health Reports* 1994;6(1):62-72.

McCulloch CE, Searle SR. *Generalized, Linear, and Mixed Models*. New York, NY: John Wiley and Sons, Inc., 2001.

Morgan CL, Currie CJ, Stott NCH, Smithers M, Butler CC, Peters JR. Estimating the prevalence of diagnosed diabetes in a health district of Wales: The importance of using primary and secondary care sources of ascertainment with adjustment for death and migration. *Diabet Medicine* 2000;17(2):141-145.

Maskarinec G. Diabetes in Hawaii: Estimating prevalence from insurance claims data. *Am J Public Health* 1997;(87):1717-1720.

Morrison DS, McLoone P. Changing patterns of hospital admission for asthma, 1981-97. *Thorax* 2001;56(9):687-690.

Muhajarine N, Mustard C, Roos LL, Young TK, Gelskey DE. Comparison of survey and physician claims data for detecting hypertension. *J Clin Epidemiol* 1997;50(6):711-718.

O'Connor PJ, Rush WA, Pronk NP, Cherney LM. Identifying diabetes mellitus or heart disease among health maintenance organization members: Sensitivity, specificity, predictive value, and cost of survey and database methods. *Am J of Managed Care* 1998;(4):335-342.

Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol* 2004;57(10):1096-1103.

Pajunen P, Koukkun H, Ketonen M, Jerkkola T, Immonen-Raiha P, Karja-Koskenkari P, Mahonen M, Neimela M, Kuulasmaa K, Palomaki P, Mustonen J, Lehtonen A, Arstilla M, Vuorenmaa T, Lehto S, Miettinen H, Torppa J, Tuomilehto J, Kesaniemi Y A, Pyorala K, Salomma V. The validity of the Finnish hospital discharge register and causes of death register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil* 2005;12(2):132-137.

Philbin EF, DiSalvo TG. Prediction of hospital readmission for heart failure: development of a simple risk score based on administrative data. *Rev Port Cardiol* 1999;18(9):855-856.

Podgor MJ, Leske MC. Estimating incidence from age-specific prevalence for irreversible diseases with differential mortality. *Stat Med* 1986;(5):573-578.

Powell KE, Diseker RA, III, Presley RJ, Tolsma D, Harris S, Mertz KJ, Viel K, Conn DI, McClellan W. Administrative data as a tool for arthritis surveillance: estimating prevalence and utilization of services. *J Public Health Manag Pract* 2003;9(4):291-298.

Quam L, Ellis LBM, Venus P, Clouse J, Taylor CG, Leatherman S. Using claims data for epidemiologic research - the concordance of claims-based criteria with the medical record and patient survey for identifying a hypertensive population. *Med Care* 1993;31(6):498-507.

Rawson N S, Malcolm E. Validity of the recording of ischaemic heart disease and chronic obstructive pulmonary disease in the Saskatchewan Health care datafiles. *Stat Med* 1995;14(24):2627-2643.

Rector TS, Wickstrom SL, Shah M, Thomas Greeenlee N, Rheault P, Rogowski J et al. Specificity and sensitivity of claims-based algorithms for identifying members of Medicare plus Choice health plans that have chronic medical conditions. *HSR* 2004;39(6):1839-1861.

Reker DM, Rosen AK, Hoenig H, Berlowitz DR, Laughlin J, Anderson L, Marshall DR, Rittman M. The hazards of stroke case selection using administrative data. *Med Care* 2002;40(2):96-104.

Rhodes L, Moorman JE, Redd SC, Mannino DM. Self-reported asthma prevalence and control among adults—United States, 2001. *JAMA* 2003;289:2639-2640.

Ritter PL, Stewart AL, Kaymaz H, Sobel DS, Block DA, Lorig KR. Self-reports of health care utilization compared to provider records. *J Clin Epidemiol* 2001;54(2):136-141.

Robinson JR, Young TK, Roos LL, Gelskey DE. Estimating the burden of disease. Comparing administrative data and self-reports. *Med Care* 1997;35(9):932-947.

Roos NP, Mustard CA. Variation in health and health care use by socioeconomic status in Winnipeg, Canada: Does the system work well? Yes and no. *Milbank Q* 1997;75(1):89-111.

Ruof J, Hulsemann JL, Mittendorf T, Handelmann S, Aultman R, der Schulenburg JM, Zeidler H, Merkesdal S. Comparison of estimated medical costs among patients who are defined as having rheumatoid arthritis using three different standards. *Eur J Health Econ* 2004;5(1):64-69.

Sartor F, Walckiers D. Estimate of disease prevalence using drug consumption data. *Am J Epidemiol* 1995;141(8):782-787.

Saydah SH, Geiss LS, Tierney E, Benjamin SM, Engelgau M, Brancati F. Review of the performance of methods to identify diabetes cases among vital statistics, administrative, and survey data. *Ann Epidemiol* 2004;12:507-516.

Schatz M, Nakahiro R, Crawford W, Mendoza G, Mosen D, Stibolt TB. Asthma quality-of-care markers using administrative data. *Chest* 2005;128(4):1968-1973.

Schultz SE, Kopec JA. Impact of chronic conditions. *Health Reports* 2003;14(4):41-53.

Shah B R, Hux J E, Zinman B. Increasing rates of ischemic heart disease in the native population of Ontario, Canada. *Arch Intern Med* 2000;160(12):1862-1866.

Shanker M, Hu MY, Hung MS. Estimating probabilities of diabetes mellitus using neural networks. *SAR QSAR Environ Res* 2000;11(2):133-147.

Singh JA, Holmgren AR, Noorbaloochi S. Accuracy of Veterans Administration databases for a diagnosis of rheumatoid arthritis. *Arthritis Rheum* 2004;51(6):952-957.

Song X, Mitnitski A, Cox J, Rockwood K. Comparison of machine learning techniques with classical statistical models in predicting health outcomes. *Medinfo* 2004;11(Pt 1):736-740.

Thacker S B, Stroup D F, Rothenberg R B. Public health surveillance for chronic conditions: A scientific basis for decisions. *Stat Med* 1995;14(5-7):629-641.

The National Asthma Control Task Force. *The Prevention and Management of Asthma in Canada: A Major Challenge Now and in the Future*. Ottawa, ON: Public Health Agency of Canada, 2000.

Tirschwell DL, Longstreth WT. Validating administrative data in stroke research. *Stroke* 2002;33(10):2465-2470.

Verbrugge L M, Lepkowski J M, Imanaka Y. Comorbidity and its impact on disability. *Milbank Q* 1989;67(3-4):450-484.

Wang PP, Elsbett-Koeppen R, Geng G, Badley EM. Arthritis prevalence and place of birth: Findings from the 1994 Canadian National Population Health Survey. *Am J Epidemiol* 2000;152(5):442-445.

Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol* 2004;57(2):131-141.

Wigertz A, Westerling R. Measures of prevalence: Which healthcare registers are applicable? *Scand J Public Health* 2001;29(1):55-62.

Wilson C, Susan L, Lynch A, Saria R, Peterson D. Patients with diagnosed diabetes mellitus can be accurately identified in an Indian Health Service Patients Registration Database. *Public Health Rep* 2001;116(1):45-50.

Wolff HK, Andreou P, Bata IR, Comeau DG, Gregor RD, Kephart G, MacLean DR, Sketris I. Trends in the prevalence and treatment of hypertension in Halifax County from 1985-1995. *CMAJ* 1999;161(6):699-704.

Wolf-Maier K, Cooper RS, Banegas JR, Giampaoli S, Hense HW, Joffres M, Kastarinen M, Poulter N, Primatesta P, Rodriguez-Artalejo F, Stegmayr B, Thamm M, Tuomilehto J, Vanuzzo D, Vescio F. Hypertension, prevalence and blood pressure levels in 6 European countries, Canada, and the United States. *JAMA* 2003;289(18):2363-2369.

Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3(1):32-35.

Young TK, Roos NP, Hammerstrand KM. Estimated burden of diabetes mellitus in Manitoba according to health insurance claims: A pilot study. *CMAJ* 1991;144(3):318-324.

APPENDIX A: RESULTS OF A LITERATURE REVIEW ON THE USE OF ADMINISTRATIVE DATA TO IDENTIFY CHRONIC DISEASE CASES

Appendix Table A.1: Summary of previous research on methods to identify arthritis cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Katz et al. (1997)	Country: USA Source: Medicare physicians claims Years: March 1, 1993 to October 31, 1993	Codes: ICD-9-CM RA: 714.0, 714.1, 714.2, 714.30, 714.31, 714.32, 714.33 OA: 715.0 or 715.09, 715.1, 715.15, 715.16, 715.2, 715.25, 715.26, 715.3, 715.35, 715.36, 715.8, 715.85, 715.86, 715.89, 715.9 Fibromyalgia: 729.0, 729.1, 729.2, 729.4, 729.5 Current Procedural Terminology (CPT) codes: Joint and soft tissue injection and aspiration: 20550, 20600, 20605, 20610	Age of cohort was not specified.	Patient office records (using a standardized data collection form). Medicare Part B physician claims: Max Sens = 90% (RA); 90% (OA) Max PPV = 95%(RA); 88% (OA in the knee)	"For OA, when specific 5-digit ICD-9-CM codes indicting the anatomic location of OA (e.g., hip or knee) were required, the sensitivities of the physician claims were ≤0.50. However, when either the 5-digit site-specific codes or the less specific 3- and 4-digit codes (indicating OA without specification of anatomic site) were accepted, the proportion of medical record diagnoses identified in claims improved to at least 0.90 for the hip and knee." p.1597
Harrold et al. (2000)	Country: USA Source: Administrative data from one health maintenance organization (HMO) Years: 1994-1996	Codes: ICD-9 OA 715.00-715.99 Algorithm: continuously enrolled in the health care plan and at least one health care encounter with an OA diagnosis.	18 years and older	OA related clinical, laboratory, and radiology data from medical records reviewed by two trained nurse reviewers and four clinicians. Max PPV = 83% (for persons with a consultation with a rheumatologist and/or orthopedic surgeon).	"The positive predictive value of an administrative OA diagnosis was 62%...When our sample was limited to those members who had a consultation with a rheumatologist and/or orthopedic surgeon, the positive predictive value increased." p.1884

Appendix Table A.1 Continued

Powell et el. (2003)	Country: USA Source: Kaiser Permanente Georgia Region (KPGR) clinical information (including primary care encounters, hospital and emergency department use, pharmacy data and lab tests). Georgia Medicaid (GAMC) data, including outpatient and hospital claims Years: 1995-1999	Codes: ICD-9-CM (the complete list included 616 diagnostic codes; these are not provided in the article) CPT codes (not specified) Algorithm: Arthritis had to be listed as one of the first 3 diagnoses	KPGR- all ages GAMC patients less than 64 year were excluded	The algorithms were not validated.	"Arthritis was defined as any diagnosis on the National Arthritis Data Workgroup (NADW) list of diagnoses and ICD-9-CM codes." p.292 "A single year's worth of data estimates the prevalence of persons with arthritis seeking medical care, but underestimates the prevalence of arthritis within the population by twofold to threefold." p.298
Losina et al. (2003)	Country: USA Source: Medicare claims, including inpatient and surgical data. Years: 1995	Codes: ICD-9-CM RA: 714	Age of cohort not specified.	Medical records and patient survey data Max Sens = 65% for medical records Max PPV = 98% for medical records. Max κ = .73	
Ruof et al. (2004)	Country: Germany Source: Administrative claims data, including prescription drug data	Codes: ICD-10 RA: M05 and M06 Algorithms: Group A - patients enrolled in a randomized clinical trial about the introduction of a disease management plan	All ages	For group A, clinical data were available for validation, while for groups B and C no clinical data were available for validation.	One of the major outcomes was, "administrative claims data in conjunction with treatment based algorithms (i.e., DMARD therapy yes/no) provide an estimate of disease related medical care costs in RA which is closely related to

Appendix Table A.1 Continued

	Years: July 2000 to June 2001	for. Group B - patients for whom administrative claims data reported the diagnosis of RA, and who had at least one prescription of a disease-modifying anti-rheumatic drug (DMARD) in the year under observation. Group C - patients for whom administrative data had a RA diagnosis, but who had no DMARD prescription in the one-year observation period.		Specificity, sensitivity and predictive values were not reported.	that of a clinically defined cohort of RA patients." p.67
Singh et al. (2004)	Country: USA Source: Minneapolis VA administrative database, including pharmacy and laboratory data. Years: Jan. 2001 to July 2002	Codes: ICD-9 RA: 714 Algorithms: For patients seen at the VA rheumatology clinic. Evaluated five algorithms: 1) ICD code 714; 2) ICD code 714 plus ≥3 month prescription of a DMARD; 3) ICD code 714 plus a positive RF titer; 4) ≥3 month prescription of a DMARD plus positive RF-; and 5) ICD code 714, a DMARD prescription, and a positive RF titer.	Age range of cohort was not specified, but mean age was 64.4 years.	Chart review. A chart diagnosis of RA by a rheumatologist on 2 separate occasions > 6 weeks apart was the gold standard. Max Sens = 88.2 % Max Spec = 97.1% Max PPV = 97.0%	
Rector et al. (2004)	Country: USA Source: Medicare and HMO data: Physician, facility (i.e., hospital), and	Codes: ICD-9-CM codes 714.xx, 715.xx, 716.xx, 720.xx, 721.1-721.4x, 721.9x in physician claims in one of up to four diagnosis fields and hospital	Age of cohort was not specified.	Survey data was the validation source. Max Sens = 70% Max Spec = 99%	"The algorithm that required at least two face-to-face claims with a first-listed diagnosis and at least one prescription for a medication commonly used to treat the condition produced

Appendix Table A.1 Continued

	pharmacy claims Years: 1999 to 2000	claims in one of up to nine diagnosis fields. CPT codes and National Drug Codes (not listed) Algorithms: 38 validated			the highest specificity" p.1849

Note: RA = rheumatoid arthritis; OA = osteoarthritis

Appendix Table A.2: Summary of previous research on methods to identify asthma cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Erzen et al. (1995)	Country: Canada Source: Physician claims from Manitoba Health Years: 1983 and 1988	Codes: ICD-9-CM 493 Algorithm: One or more contacts in physician billing claims.	All ages were examined and then grouped as follows: 0-14 years, 15-34 years, 35+ years	There was no validation data source. Specificity, sensitivity and predictive values are not reported.	
Morrison et al. (2001)	Country: Scotland, UK Source: Scottish Morbidity Records Years: 1981-1997	Codes: ICD-9 493 ICD-10 J45, J46 Algorithm: A patient's first admission for asthma was labelled as the first admission; all subsequent admissions were labelled as readmissions.	Age of cohort was not specified.	There was no validation data source. Specificity, sensitivity and predictive values are not reported.	
Huzel et al. (2002)	Country: Canada Source: Physician billing claims from Manitoba Health Years: 1993-1994	Codes: ICD-9 493 Algorithm: One or more contacts in physician billing claims	20 to 44 years of age	Survey data Max Sens = 70.1% Max Spec = 99.8% Max κ = 0.62	"Even if five years of physician contacts were studied, the claims identified only 63.3% of people reporting an asthma attack in the 12 months before the survey, 65.1% of people reporting being on medication, 62.0% of those reporting either an attack or using medication , and 70.1% of those reporting both an attack and using medication." p.410

Appendix Table A.2 Continued

Hansell et al. (2003)	Country: England Source: Mortality data from the Office for National Statistics (ONS; 1991-1995). Hospital Episode Statistics (HES) from the Dept. of Health (1990/91 to 1993/94) and from the Data Science UK (1994/95). General Practice Research Database (GPRD; 1991-1995)	Codes: ICD-9 493 Algorithms: From HES, emergency hospital admissions with the ICD-9 code as a primary diagnosis were abstracted. From GPRD, patients who consulted in primary care with an inhaler prescription during the current year plus a current or prior diagnosis of asthma in each year from 1991-1995 were identified.	Survey included ages 2 years and above.	Health Survey of England, 1995. Specificity, sensitivity and predictive values are not reported.	"There was little consistency in patterns for asthma from the different data source." p. 281
Kozyrskyj et al. (2004)	Country: Canada Source: Hospital separations, physician billing claims, prescription drug records from Manitoba Health. Years: 1995-1998	Codes: ICD-9 493 Algorithms: At least one physician claim or hospitalization for asthma-like diagnoses in one year. In the absence of these diagnoses, at least one prescription for an inhaled corticosteroid or cromone, or ketotifen concomitant with an inhaled or oral beta-agonist, or two or more prescription for an inhaled or oral beta-agonist.	Ages 5-15 years	A prescription for an asthma drug was used as the gold standard. Max Sens = 73.9% Max Spec = 91.4%	

Appendix Table A.2 Continued

Borzecki et al. (2004)	Country: USA Source: Out-patient clinic (OPC) file for 1998 and 1999, and patient treatment file (1999), from a National Department of Veterans Affairs electronic database.	Codes: ICD-9 chronic obstructive lung disease/asthma 490.x, 491.x (all except for 491.8), 492.0, 493.x, 496.x Algorithms: Varied the minimum required number of claims with a given diagnosis from 1 to 2 and varied the number of years of data from 1 to 2.	Outpatients receiving primary care at 10 different sites across the country.	Electronic clinicians' notes (i.e., medical outpatient charts). Max. Spec = 92% Max. Sens = 81% Max κ = 0.68	
Wilchesky et al. (2004)	Country: Canada Source: Medical claims from Quebec Year: 1995-1996	Codes: ICD-9-CM 493.0-493.9 Algorithm: One or more medical claims with the relevant ICD-9-CM code	66 years of age and older	Physicians' charts Max Sens = 43% Max Spec = 99%	
Macy et al. (2005)	Country: USA Source: Southern California Kaiser Permante hospital, outpatient, and prescription data Year: 2001	Codes: ICD-9-CM 493.xx Drug codes were not specified, but included β-agonists (excluding oral terbutaline), orally inhaled corticosteroids, other inhaled anti-inflammatory drugs, and oral leukotriene modifiers. Algorithm: Any discharge diagnosis of asthma in the hospitalization database; or two or more asthma-related medication dispensations in the prescription database; or any ED or regular clinic asthma-related claim in the	18-64 years	Medical chart review and questionnaire/ examinations. Specificity, sensitivity and predictive values are not reported.	

Appendix Table A.2 Continued

Schatz et al. (2005)	Country: USA Source: Southern California Kaiser-Permanente asthma database (linked hospital discharge database, outpatient diagnosis and procedures database, a membership database, and a prescription database) Year: 2002-2003	outpatient diagnosis and procedures database. Codes: ICD-9 493.xx Algorithms: One or more of the following criteria in the prior 12 months: 1) Any hospital discharge diagnosis of asthma in the hospitalization database (ICD code), 2) the dispensing of two or more asthma-related dispensings (excluding oral steroids), including β-agonists, inhaled steroids, other inhaled anti-inflammatory drugs, and oral leukotriene modifiers, 3) an asthma related visit to the ED or regular clinic listed in the diagnosis and procedures database	Ages 5-56	No validation source was used Specificity, sensitivity and PPV are not calculated.		

Appendix Table A.3: Summary of previous research on methods to identify coronary heart disease cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Rawson et al. (1995)	Country: Canada Source: Hospital discharge abstracts Year: 1986	Codes: ICD-9 410 – 414 Algorithms: Hospital discharge data with a relevant diagnosis code in the primary or secondary discharge diagnosis field.	Age of cohort was not specified.	Validation was performed with medical chart review for AMI only. Physician billing claims were used to validate AMI, angina, and chronic IHD. Concordance between hospital discharge abstracts and medical chart review for AMI was 96.9%. Concordance between hospital discharge abstracts and physician claims was 69.3% for AMI, 70.0% for angina, and 55.6% for chronic IHD when the primary diagnosis was used.	"When diagnostic agreement was broadened to include any physician diagnosis in the same ICD sub-chapter, the concordances all increased ranging from 79% to 94% in the IHD groups…" p.2639
Mahonen et al. (1997	Country: Finland Source: Finnish national hospital discharge register	Codes: ICD-9 410 – 414 Algorithm: Diagnosis of CHD in hospital discharge abstract	Ages 25-65 years	FINMONICA acute myocardial infarction register Max. Sens: 87.1%	"There are no medical record abstractors in Finland; clinicians taking care of the patients assign both the diagnoses and ICD codes for hospitalizations."

Appendix Table A.3 Continued

	Year: 1983-1990			Max. PPV: 95.7%	p.405
O'Connor (1998)	Country: USA Source: Clinic (i.e., physician) data Year: 1992-1994	Codes: ICD-9 428.0 (congestive heart failure), 412 (history of myocardial infarction), 429.2 (arteriosclerotic cardiovascular disease), 413.9 (angina) Algorithm: At least one visit with at least on of the codes in a defined 1-year period	HMO members. Age is not specified	Survey data. The question was: "Have you ever been told by a doctor that you had a heart attack?" When survey data and database data conflicted, a chart review was done. Max. Sens: 89% Max. Spec: 99% Max. PPV: 85%	
Shah et al. (2000)	Country: Canada Source: CIHI hospital discharge abstracts for Ontario Years: 1981-1997	ICD-9 410, 411, 413, 414. CCP codes for angioplasty and bypass: 48.0, 48.1 Algorithm: An event of heart disease was defined as a hospital discharge that has a selected ICD-9 code as one of the first two discharge diagnoses, or a selected CCP code as one of the first 2 procedures listed.	Residents of Ontario communities that had regular census participation and at least 95% of their population claiming Native origins (N=16875 in 1991)	No validation was conducted.	

Appendix Table A.3 Continued

Pajunen et al. (2005)	Country: Finland Source: Finnish national hospital discharge register Years: 1988 – 2002	ICD-9 410 (AMI) ICD-10 I21-I22 (AMI) ICD-9 411.0 (unstable angina) ICD-10 I20.0 (unstable angina) – these were investigated in a secondary component of the analysis Non-fatal events were identified where one of the selected codes was either the main diagnosis or an additional diagnosis.	35 years of age and older.	Validation was conducted using the myocardial infarction (MI) register. For both fatal and non-fatal events: Max Sens = 85% Max PPV = 90% For non-fatal events: Max Sens = 94% However, there were wide geographic variations in sensitivity, from 52% to 94%.	

Appendix Table A.4: Summary of previous research on methods to identify diabetes cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Blanchard et al. (1996)	Country: Canada Source: Physician billing claims and hospital separation records from Manitoba Health Fiscal years: 1986-1991	Codes: ICD-9-CM 250 Algorithms: at least two separate physician claims for diabetes within 2 years of each other or at least one hospital separation record with a diagnosis of diabetes.	25 years or older	Diabetes Education Resource (DER) database which includes all contacts with DER program clients, including clinical, service-related, and demographic information. Specificity, sensitivity and predictive value are not reported. Ascert. rate > 95%.	
Morgan et al. (2000)	Country: Wales, UK Source: Inpatient dataset (1991-1997), outpatient dataset (1991-1996), diabetes clinic dataset (1993 to present), Office of National Statistics mortality dataset (1993-1997)	Codes: ICD-9 250 ICD-10 E10-E14 Algorithms: An inpatient diagnosis of diabetes, attendance at an outpatient clinic coded as diabetic, inclusion on the diabetic clinic dataset or cause of death coded as diabetes on the ONS mortality dataset.	Age of cohort was not specified.	A general practice audit database. Specificity, sensitivity and predictive value are not calculated.	"This study combines primary and secondary care data sources to estimate the prevalence of diagnosed diabetes…" p.143
Martens et al. (2003)	Country: Canada Source: Physician billing claims and hospital separation records from Manitoba Health Years: 1997	Codes: CDI-9 250 Algorithms: One or more hospitalizations or two or more physician claims with a diabetes diagnosis in a three-year period	20 to 79 years of age	Manitoba First Nations Regional Health Survey (1998) Sens = 76.0%	

Appendix Table A.4 Continued

Wilson et al. (2001)	Country: USA Sources: Indian Health Service patient registration databases Years: Does not provide	Codes: ICD-9 250.00-250.93 Algorithm: Four sets of criteria: 1) at least one ICD-9 code, 2) at least two separate ICD-9 codes, 3) a pharmacy prescription entry for sulfonylurea, metformin, acarose, thiazolidindione, or insulin, 4) at least two separate glucose values ≥200 mg/dl.	15 years of age and older American Indian and Alaskan Native people	Medical chart review Max. Sen = 92% Max. Spec = 99 PPV = 95%	"The specificity of a single 250.00 to 250.93 ICD-9 code was nearly the same as the use of two 250.00 to 250.93 ICD-9 codes. The use of two 250.00 to 250.93 ICD-9 codes resulted in significant loss of sensitivity." p.48-49
Hux et al. (2002)	Country: Canada Source: CIHI discharge abstracts and Ontario Health Insurance Plan for physician service claims Fiscal years: 1991-1999	Codes: ICD-8 or ICD-9-CM 250.x Algorithm: One hospital separation or 1 or 2 physician service claims in a two-year period. Examined all 16 diagnosis fields in hospital data.	Age of cohort was not specified.	Ontario Drug Benefit Program database (for individuals 65+ years). National Population Health survey (National Population Health Survey; NPHS) Physician office charts Max Sens = 94% Max PPV = 98%	
Wilchesky M et al. (2004)	Country: Canada Source: Medical claims from Quebec Year: 1995-1996	Codes: ICD-9-CM 250.0-250.9 Definitions: All medical claims with a relevant diagnosis	66 years of age and older	Medical charts Max Sens = 64% Max Spec = 98%	

Appendix Table A.4 Continued

Borzecki et al. (2004)	Country: USA Source: Out-patient clinic (OPC) file for 1998 and 1999, and patient treatment file (1999), from a National Department of Veterans Affairs electronic database.	Codes: ICD-9 250 Algorithms: Varied the minimum required number of claims with a given OPC diagnosis from 1 to 2 and varied the number of years of data from 1 to 2.	Outpatients receiving primary care at 10 different sites across the country. Age of cohort was not specified.	Electronic clinicians' notes (i.e., medical outpatient charts) Max Sens = 97% Max Spec = 96% Max κ = 0.92	
Rector et al. (2004)	Country: USA Source: Medicare+Choice health plan claims data: Physician, hospital, and pharmaceutical claims Years: 1999, 2000	Codes: ICD-9-CM 250.xx, 357.1x, 352.0x, 355.41 in physician claims in one of up to four diagnosis fields and hospital claims in one of up to nine diagnosis fields. National Drug Codes were not specified for the pharmaceutical claims. Algorithms: 38 different algorithms were examined.	Age of cohort was not specified.	Survey data collected from health plan members Max. Sens = 95% Max. Spec = 100%	"Diabetes was the only condition where an algorithm had a specificity and sensitivity greater than 0.90." p.1852

Appendix Table A.5: Summary of previous research on methods to identify hypertension cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Quam et al. (1993)	Country: USA Source: ambulatory physician, hospital, and pharmacy claims from two different medical plans Years: Jan.1988 - Dec.1989	Codes: Drug codes not specified (anti-hypertensive) Algorithms: Used three mutually exclusive algorithms Dx: Has at least one claim indicating a diagnosis of essential hypertension (ICD-9-CM codes) and no prescriptions for the most common antihypertensive medications. Rx: Has at least one prescription for an antihypertensive medication, and no claims with a diagnosis code for hypertension. B: Has at least one hypertension diagnosis code, or at least one prescription for an antihypertensive medication.	Between 18 and 65 years of age	Medical charts Patient survey Sensitivity, specificity and predictive values are not reported.	"We found that the submission of a diagnosis of hypertension on a single claim form is not a valid indicator of the presence of hypertension." p.504
Robinson et al. (1997)	Country: Canada Source: Hospital separations and physician billing claims from Manitoba Health	Codes: ICD-9 401-405, 642, 362.11, 416.0, 437.2, 796.2 Algorithms: Several algorithms were evaluated that varied in the number of years of administrative data and the number of		Survey data (Manitoba Heart Health Survey) Max. Spec = 0.86 Max. Sens = 0.78 Max. PPV= 0.63	

Appendix Table A.5 Continued

				Max. κ = 0.59	
	Years: 1986-1989	occurrences of a relevant diagnostic code in one of the 16 diagnostic fields for hospital discharge abstracts and the single diagnostic code in physician claims.			
Muhajarine et al. (1997)	Country: Canada Source: Physician claims data from Manitoba Health Years: Oct. 1987-Feb 1990	Codes: ICD-9 401, 402 Algorithm: Any claim filed for services during the 2 years prior to the survey with a diagnostic code of 401 or 402.	Ages 18 to 74 years	Survey data (Manitoba Heart Health Survey) Max κ = 0.65	"The overall proportion agreement between self-reported and physician claims hypertension was 81.7%." p.714
Borzecki et al. (2004)	Country: USA Sources: Out-patient clinic (OPC) file (1998 and 1999) and Patient Treatment file (PTF) (1999), from a National Department of Veterans Affairs database Years: 1998, 1999	Codes: ICD-9-CM 401, 402, or 405 Algorithms: Varied the minimum required number of OPC records with a relevant diagnosis from 1 to 2 and varied the number of years of data from 1 to 2	Outpatient s receiving care at 10 different sites across the country	Electronic clinicians' notes (medical outpatient charts), from the Veterans Health Information System and Technology Architecture (VISTA) for 1999 Specificity, sensitivity, and predictive values are not reported.	
Wilchesky et al. (2004)	Country: Canada Source: Physician billing claims from Quebec Year: 1995-1996	Codes: ICD-9-CM 401.0-401.9 Algorithm: a single occurrence of the diagnostic code in one year of data	66 years of age and older	Physician charts Max Sens = 69% Max Spec = 88%	

Appendix Table A.5 Continued

Rector et al. (2004)	Country: USA Source: Physician, facility (i.e., hospital), and pharmacy claims. Years: 1999 and 2000	Codes: ICD-9-CM 401.0, 401.1, 401.9, 402.xx, 403.xx, 404.xx in physician claims in one of up to four diagnosis fields and hospital claims in one of up to nine diagnosis fields. Current Procedural Terminology (CPT) codes are not specified. National Drug Codes are not specified. Algorithms: 38 different algorithms were examined	Age of cohort was not specified.	Survey data Max Sens = 95% Max Spec = 96%	

Appendix Table A.6: Summary of previous research on methods to identify stroke cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Leibson et al. (1994)	Country: USA Source: Hospital abstracts Years: 1970, 1980, 1984, 1989	Codes: ICD-8, ICD-9 codes 430 – 438.9 in either the first/primary diagnosis field, or in any of up to 5 diagnosis fields	Not stated	Stroke registry – developed by chart validation – for first strokes only Max Sens = 88% Max PPV = 79%	Individuals were classified as incident stroke, recurrent stroke, nonstroke event, sequelae. Positive predictive value of the algorithm increased when 432, 435, and 438 were excluded. Fatal events, and events that occur in hospital may be missed
Mayo et al. (1994)	Country: Canada Source: Hospital abstracts Year(s) not specified	Codes: ICD-9 Algorithm #1: 430 – 437, excluding 435 in the first/primary diagnosis field Algorithm #2: 431, 434, 436 – used to examine the hospitalization rate for stroke (not clear which diagnosis field)	15+ years	Chart abstraction – reviewed by neurologists Sensitivity, specificity, and PPV were not calculated.	Authors note that it is difficult to distinguish first strokes from recurrent strokes. Not all patients with stroke are admitted to hospital; some will be seen only in physician office or ERs; some will die before reaching hospital.
Benesch et al. (1997)	Country: USA Source: Hospital abstracts Year: 1992	Codes: ICD-9 codes 433 – 436 in either the primary diagnosis field or in up to 14 other diagnosis fields	Not stated	Medical chart review Max Sens = 95% (for both Transient ischemic attach (TIA) and Stroke) Max κ = .86 (stroke),	Individuals were classified as (1) stroke, (2) transient ischemic attach (TIA) but not stroke, (3) asymptomatic for cerebrovascular disease. The authors

Appendix Table A.6 Continued

					.83 (TIA)	identified that 77% of individuals with a primary diagnosis of 433 were asymptomatic and 85% of individuals with a primary or secondary diagnosis of 433 were asymptomatic. They concluded that 433 is a "non-stroke" code. The authors identified that 89% of individuals with a primary diagnosis of 435 and 77% of those with a primary or secondary diagnosis of 435 were TIA – again, could exclude this code
	Ellekjaer et al. (1999)	Country: Norway Source: Hospital abstracts Years: 1994 - 1996	Codes: ICD-9 430 – 438.9 in any diagnosis field. Also defined acute stroke as 430, 431, 434, 436 in any diagnosis field	15+ years	Stroke register Max Sens = 95% (Hospitalized only with discharge diagnosis ICD-9 codes 430, 431, 434, and 436, first admission) Max PPV = 68% (discharge diagnosis ICD-9 codes 430, 431, 434, and 436, first admission)	Using codes for acute stroke gave an incidence estimate close to the true incidence identified from the stroke register. Only a few cases were classified as non-stroke events.
	Leppala et al. (1999)	Country: Finland	Codes: ICD-8, ICD-9 codes: 430 (SAH); 431 (ICH); 433, 434 (CI	50-69 years,	Chart review	

Appendix Table A.6 Continued

	Source: National Hospital discharges register and the Register of Causes of Death Years: 1985 - 1993	(cerebral infarction)); 436 (unspecified stroke); 437 (other cerebrovascular disease); 438 (sequels); Excluded 432, 435, 4330x, 4331x, 4339x, 4349x Definition: Codes were identified in up to 4 diagnosis fields; or as the underlying cause of death.	males Note, that the data were collected for another study, but the validation was part of the preliminary analysis	Max Sens = 100% (SAH between reviewer's diagnoses and the HDR diagnoses) 100% (for both SAH and ICH between reviewer's diagnoses and underlying cause of death) 100% (SAH between HDR diagnoses and underlying cause of death)	
Tirschwell & Longstreth (2002)	Country: USA Source: Hospital abstracts Years: 1990 - 1996	Codes: ICD-9-CM 430 – 438 Ischemic stroke - 433x1, 434 (excluding 434x0), 436 TIA – 435 SAH – 430 ICH – 431 Charts with only 432, 437, 438 were classified as not a stroke Cases were excluded (i.e., not a stroke) if "traumatic brain injury" – 800-804 or 850-854, or "rehab care" – V57 was present	20+ years	Chart abstraction by stroke neurologist Max Spec = 97% (SAH using only primary discharge diagnosis) Max Sens = 98% (SAH using all discharge diagnoses) Max κ = 0.88 (SAH using only primary discharge diagnosis) Max PPV = 94% (SAH using only primary discharge diagnosis)	Algorithm #1 – all diagnosis fields; hierarchical assignment to one of the following categories: not a stroke, TIA, ICH (intracerebral hemorrhage), SAH (subarachnoid hemorrhage); cases with both ICH and SAH were assigned to SAH Algorithm #2 – same as #1, but search only the first two diagnosis fields Algorithm #3 – search only the first/primary diagnosis field Algorithm #1 maximized sensitivity and kappa; Algorithm

Appendix Table A.6 Continued

	Reker et al. (2002)	Country: USA Source: Hospital abstracts Years: 1996 - 1998	Codes: High sensitivity algorithm – used any of the following three criteria for stroke identification in ICD-9-CM: 430xx, 431xx, 434xx, 436xx , 433.01, 433.11, 433.21, 433.31, 433.81, 433.91; Rehabilitation (V57) as primary diagnosis and a secondary diagnosis of 342xx, 430xx, 431xx, 433xx, 434xx, 435xx, 436xx, 437xx, 438xx; Primary diagnoses of occlusion and stenosis of precerebral arteries (433xx) or TIA (435) and secondary diagnoses with any of the following: 342xx, 430xx, 431xx, 432xx, 434xx, 436xx High specificity algorithm: 431xx, 433x1, 434x1	Not stated	Survey data	#3 maximized specificity and positive predictive value. Authors limited attention to first hospitalization in the study interval	Goal was to examine 30-day mortality following stroke as a patient outcome; different algorithms give different perspectives on this quality indicator "Many algorithms have been used by researchers and reporting agencies to identify samples or populations of stroke patients"

Appendix Table A.7: Summary of previous research on methods to identify congestive heart failure cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Jollis et al. (1993)	Country: USA Source: Hospital discharge abstracts from Duke University Medical Center Years: July 1985 and May 1990	Codes: ICD-9-CM 428.0, 428.1, 428.9, 398.91, 402.01, 402.11, 402.91 Algorithms: All discharge abstracts with a diagnosis code for CHF	12,937 patients having an inpatient cardiac catheterization i.e., procedure code for coronary arteriography	Duke Databank for Cardiovascular Disease (clinical data) Max Spec = 96% Max Sens = 36% κ = 0.39	"Claims data identified 31% of clinically identified congestive heart failure that was New York Health Association class 1 and 11 and identified 45% of class III and IV heart failure (P< 0.0001)." p.846
Philbin et al. (1999)	Country: USA Source: New York State hospital discharge abstracts Year: 1995	Codes: ICD-9-CM 428.0, 402.91, 404.93, 428.1, 402.11, 398.91, 404.91, 404.13, 402.01, 404.03, 404.11, 404.01, 428.9 Definitions: ICD-9-CM code in principle diagnosis position	Age of cohort was not specified.	Chart review Sensitivity, specificity and PPV were not calculated.	"Coexistent illnesses were determined by searching the principal ICD-9-CM diagnosis code and up to 14 secondary diagnosis codes for each patient" p.1561
Goff et al. (2000)	Country: USA Source: Discharge diagnoses (including notations from chest radiograph) Years: May 1, 1988-April 30, 1994	Codes: ICD-9-CM 402, 410, 411, 412, 413, 414, 127, 428, 429, 440, 786.5 Definitions: Eligible cases were determined by monitoring admissions to special care units and obtaining discharge lists with selected CHD-related diagnoses. Only codes 402 and 428 were	Age distribution of cohort was not specified. Mean age: 60.4	Medical records Max Sens = 73% (any CHF code for ages ≥63 years) Max Spec = 96.5% (for ICD-9-CM code 428.0 ages < 63) Max PPV = 86.3% (for ICD-9-CM code	"Reliance on ICD codes results in the exclusion of one third of the patients with clinical evidence of acute CHF" p.197

Appendix Table A.7 Continued

		assigned to more than 1%		428.0)	
Feldman et al. (2001)	Country: Canada Sources: Quebec death certificates and hospital discharge abstracts (Med-Echo data) Years: 1990-1997	Codes: ICD-9 428 Definitions: All hospital abstracts with a primary diagnosis of 428	Individuals of CHD related hospitalizations 65 years and over	No validation source was used Sensitivity, specificity and predictive values were not calculated.	
Martinez-Selles et al. (2002)	Country: Spain Source: hospital admission records Year: 1996	Codes: ICD-9-CM 402.9, 428.0, 428.1, 428.9, 425.4, 425.5, 425.9 Algorithms: Any admission with a relevant ICD-9-CM code as either the principle or any other diagnosis.	15 years or older	All data collected was reviewed and agreed upon by two cardiologists Sensitivity, specificity and PPV were not calculated.	
Clarke et al. (2002)	Country: USA Source: professional, facility, and ancillary claims (inpatient, outpatient, home health, pharmacy, durable medical equipment, and any other ancillary) Year: 2000-2001	Codes: ICD-9 398.91, 402.01, 402.11, 402.91, 404.01, 404.02, 404.11, 404.12, 404.91, 404.93, 425.0, 428.0, 428.9 Algorithm: A minimum of one HF inpatient claims/encounters or two CHF outpatient claims/encounters within a 12-month period with matching primary or secondary ICD-9 diagnosis codes.	Health plan members identified as having CHF were automatically included unless they choose not to participate	No validation source used Sensitivity, specificity and PPV were not calculated.	

Appendix Table A.7 Continued

Austin et al. (2002)	Country: Canada Source: Canadian Institute for Health (CIHI) Information discharge abstract database Years: 1996-2000	Codes: ICD-9 428 Definitions: The diagnosis was deemed as present if it was present either as the most responsible diagnosis or as secondary diagnosis and if it did not arise after hospital admission	Ages 20 and older	The Fastrak II acute coronary syndromes database (clinical registry) Max Spec = 96.8 (including only the most responsible diagnosis) Max Sens = 85.4 (including both the most responsible diagnosis and the secondary diagnosis) PPV = 65.1 (including only the most responsible diagnosis) κ = 0.58 (including only the most responsible diagnosis)	
Cujec et al. (2004)	Country: Canada Source: CIHI hospital discharge abstracts and Alberta physican claims (1995/96-1999/00) Fiscal years: 1992/93-1999/00	Codes: ICD-9-CM 428.x, 398.91, 402.x1, 404.x greater than 0, 514.x,518.4, 425.9 in the most responsible diagnosis field of the hospital abstract Algorithm: ICD-9-CM code in the most responsible diagnosis field of the hospital	Aged 20 or over Patients diagnosed at the time of hospitalization and patients diagnosed in specialists offices without a prior hospitalization for CHF	No validation source used Sensitivity, specificity and PPV were not calculated.	

Appendix Table A.7 Continued

		abstract, and office or ambulatory care internal medicine specialist or sub-specialist consultation claim for CHF as defined by ICD-9-CM			
Borzecki et al. (2004)	Country: USA Source: Out-Patient clinic (OPC) file (1998 and 1999) and Patient Treatment file (PTF) (1999), from a National Department of Veterans Affairs database Years: 1998, 1999	Codes: ICD-9-CM 398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 414.8, 428.x Algorithm: Varied the minimum required number of OPC claims with a relevant diagnosis from 1 to 2 and varied the number of data years from 1 to 2.	Outpatients receiving care at 10 different sites across the country	Electronic clinicians' notes (medical outpatient charts), from the Veterans Health Information System and Technology Architecture (VISTA) (1999) Max Sens = 77% Max Spec = 99% Max κ = 0.74	
Lee et al. (2005)	Country: Canada Source: CIHI discharge abstract database Years: 1997-1999	Codes: ICD-9 428 Definitions: A primary diagnosis of CHF	Excluded patients ≥105 years of age	Clinical chart data PPV: 94.3% (primary diagnosis using the Framingham criteria)	

Appendix Table A.8: Summary of previous research on methods to identify renal disease cases from administrative data

Author	Data Source	Diagnosis/Treatment Codes and Algorithms	Study Cohort	Validation Methodology	Comments
Brameld et al. (1999)	Country: Australia Source: Hospital discharge abstracts and death records from the Western Australia Health Services Research Linked Database Year: 1980-1994	Codes: ICD-9-CM or ICD-9 586-6, V560, V568 Procedure code ICD-9-CM 39.95, 54.98 or International Classification of Procedures in Medicine (ICPM) 8-853, 8-860 Renal transplant: ICD-9-CM 55.69 or ICPM 5-555 Algorithms: Hospital morbidity and death records were extracted by the following criteria: a principal condition of chronic or unspecified renal failure, haemodialysis or peritoneal dialysis; or a principal procedure coded to haemodialysis or peritoneal dialysis; or any mention of renal transplant. Patients with end-stage renal failure were defined as those who had more than 10 hospital admissions for dialysis in a period greater than 28 days.	Age of cohort was not specified.	The Australia and New Zealand Dialysis and Transplant (ANZDATA) Registry Sensitivity, specificity, and PPV were not calculated.	Renal transplants were assumed to be failed if the procedure code ICD-9 55.53 or ICPM 5-554 (removal of the kidney) was encountered.
Haroun et al. (2003)	Country: USA Sources: CLUE Records (a health interview study) linked with Health Care Financing	Codes: ICD-9 250.4, 274.1, 275.4, 403, 404, 580-589, 593.9 Algorithm: Diagnostic codes were used in a comprehensive medical record review of all deaths by Feb. 1994 with renal disease listed on the death certificate.	Washington County, MD Participants were entered at the age of participation in CLUE	Medical chart review Sensitivity, specificity, and PPV were not calculated.	"CKD was defined as the presence of any of the following: 1) CKD in a discharge summary and/or renal consultation note, or 2) at least two serum creatinine values > 2.0 mg/dl over the period

Appendix Table A.8 Continued

	Administration (HCFA) database; Death certificates Year: 1994		and were withdraw at the time of onset of Chronic kidney disease (CKD) or Feb. 19,1994		of follow-up." p.2935
Li et al. (2004)	Country: Australia Source: death certificates Year: 1997-1999	Codes: ICD-10, renal causes of death were: 1) Disease of the kidney and ureter: N00-N29; 2) Diabetic renal disease: E10.2, E11.2, E13.2 and E14.2; 3) Hypertensive renal disease: I12, I13, I15.0, I15.1; 4) Congenital malformations of the kidney and ureter: Q60-A63 Definitions: Extracted all data on deaths between 1997-99, analysed for renal disease as both underlying cause and associated cause.	All deaths including one or more with the listed ICD-10 codes between 1997-1999	No validation source was used. Sensitivity, specificity, and PPV were not calculated.	9.5% of death certificates named renal disease as a cause of death. p.260 "Renal disease is 3.6-fold more likely to be reported as an associated cause of death than as the underlying cause of death." p.260
Borzecki et al. (2004)	Country: USA Sources: Out-Patient clinic (OPC) file (1998 and 1999) and Patient Treatment file (PTF) (1999), from a National Department of	Codes: ICD-9-CM 403, 405.01, 405.11, 405.91, 582, 583, 585, 586, 593.9 Algorithms: Varied the minimum required number of claims with a given OPC diagnosis from 1 to 2 and varied the number of years of data from 1 to 2.	Outpatients receiving care at 10 different sites across the country	Electronic clinicians' notes (medical outpatient charts), from the Veterans Health Information System and Technology Architecture (VISTA) (1999) Max Spec = 98%	

Appendix Table A.8 Continued

	Veterans Affairs database Years: 1998, 1999			Max Sens = 62% Max Kappa = 0.62	

APPENDIX B: SUPPLEMENTARY DATA FOR ARTHRITIS ALGORITHMS

Appendix Table B.1: Supplementary data for arthritis algorithms

	ATC Code	Generic Name
Disease-Modifying and Anti-Rheumatic Drugs (DMARDS): Xenobiotic Agents	A07EC01	sulfasalazine
	J01AA08	minocycline
	L01AA01	cyclophosphamide
	L01BA01	methotrexate
	L04AA01	cyclosporine
	L04AA13	leflunomide
	L04AX01	azathioprine
	L04AX03	methotrexate
	M01CB01	sodium aurothiomalate
	M01CB03	auranofin
	M01CB04	aurothioglucose
	M01CC01	penicillamine
	P01BA02	hydroxychloroquine
DMARDS: Biological Agents	L04AA11	etanercept
	L04AA12	infliximab
	L04AA14	anakinra
	L04AA17	adalimumab
Analgesics	N02AA05	oxycodone
	N02AD01	pentazocine
	N02BA51	codeine in combination
	N02BE01	acetaminophen
	N02BE51	paracetamol, combinations excluding psycholeptics
	R05DA03	hydrocodone
	R05DA04	codeine
	R05DA05	opium alkaloids with morphine
Glucocorticosteroids	H02AB04	methylprednisolone
	H02AB06	prednisolone
	H02AB07	prednisone
	H02AB08	triamcinolone
	H02AB10	cortisone
Non-Steroidal Anti-Inflammatory Drugs	M01AH03	valdecoxib
	M01AA01	phenylbutazone

Appendix Table B.1 Continued

(NSAIDs)	M01AB01	indometacin
	M01AB02	sulindac
	M01AB03	tolmetin
	M01AB05	diclofenac
	M01AB08	etodolac
	M01AB15	ketorolac
	M01AB55	diclofenac in combination
	M01AC01	piroxicam
	M01AC02	tenoxicam
	M01AC06	meloxicam
	M01AE01	ibuprofen
	M01AE02	naproxen
	M01AE03	ketoprofen
	M01AE04	fenoprofen
	M01AE09	flurbiprofen
	M01AE11	tiaprofenic acid
	M01AE12	oxaprozin
	M01AG01	mefenamic acid
	M01AH01	celecoxib
	M01AH02	rofecoxib
	M01AX01	nabumetone
	M02AA	anti-inflammatory preparations, non-steroids for topical use
	M02AB01	capsicum
	M02AC	preparation with salicylic acid derivations
	M02AX03	dimethyl sulfoxide
Other	M04AA	preparation inhibiting uric acid production
	N02BA11	diflunisal
	N02BA01	acetylsalicylic acid
	N02BA03	choline salicylate

APPENDIX C: ADDITIONAL VALIDATION RESULTS FOR ARTHRITIS ALGORITHMS

Appendix Table C.1: Estimates of agreement, sensitivity, specificity, and predictive values for additional arthritis algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ H or 1+ P or 1+ Rx	0.28	63.5	70.0	0.34	40.1	85.8
	2	1+ H or 2+ P or 1+ Rx	0.33	57.8	77.9	0.36	45.3	85.4
	3	1+ H or 2+ P or 2+ Rx	0.35	48.5	85.5	0.34	51.5	84.0
2	4	1+ H or 1+ P or 1+ Rx	0.24	75.1	57.6	0.33	35.9	87.9
	5	1+ H or 2+ P or 1+ Rx	0.30	70.5	67.1	0.38	40.4	87.8
	6	1+ H or 2+ P or 2+ Rx	0.33	61.8	76.1	0.38	45.0	86.3
3	7	1+ H or 1+ P or 1+ Rx	0.21	82.4	48.6	0.31	33.7	89.7
	8	1+ H or 2+ P or 1+ Rx	0.26	77.7	58.5	0.36	37.2	89.2
	9	1+ H or 2+ P or 2+ Rx	0.30	68.5	68.8	0.37	41.1	87.3
5	10	1+ H or 1+ P or 1+ Rx	0.16	88.4	36.8	0.25	30.8	91.2
	11	1+ H or 2+ P or 1+ Rx	0.21	85.3	46.8	0.32	33.7	91.0
	12	1+ H or 2+ P or 2+ Rx	0.27	78.1	58.5	0.37	37.4	89.4

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record

Appendix Table C.2: Estimates of agreement, sensitivity, specificity, and predictive values for additional rheumatoid arthritis algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ H or 1+ P or 1+ Rx	0.11	7.8	99.1	0.07	43.9	92.3
	2	1+ H or 2+ P or 1+ Rx	0.12	7.8	99.3	0.07	50.7	92.3
	3	1+ H or 2+ P or 2+ Rx	0.10	6.5	99.4	0.06	50.0	92.2
2	4	1+ H or 1+ P or 1+ Rx	0.12	9.8	99.4	0.09	35.2	92.4
	5	1+ H or 2+ P or 1+ Rx	0.13	9.6	98.8	0.08	42.7	92.4
	6	1+ H or 2+ P or 2+ Rx	0.11	7.8	99.0	0.07	41.9	92.3
3	7	1+ H or 1+ P or 1+ Rx	0.14	11.8	97.8	0.10	32.9	92.5
	8	1+ H or 2+ P or 1+ Rx	0.15	11.8	98.5	0.10	41.2	92.6
	9	1+ H or 2+ P or 2+ Rx	0.12	9.4	98.7	0.08	39.8	92.4
5	10	1+ H or 1+ P or 1+ Rx	0.12	12.4	96.9	0.09	26.2	92.5
	11	1+ H or 2+ P or 1+ Rx	0.10	10.0	97.3	0.07	24.7	92.3
	12	1+ H or 2+ P or 2+ Rx	0.14	12.4	97.9	0.10	34.6	92.6

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record

Appendix Table C.3: Estimates of agreement, sensitivity, specificity, and predictive values for additional osteoarthritis algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ H or 1+ P or 1+ Rx	0.19	55.2	76.5	0.32	22.1	93.4
	2	1+ H or 2+ P or 1+ Rx	0.26	46.3	86.6	0.33	29.4	93.0
	3	1+ H or 2+ P or 2+ Rx	0.25	41.3	88.1	0.29	29.4	92.6
2	4	1+ H or 1+ P or 1+ Rx	0.17	68.6	66.4	0.35	19.7	94.6
	5	1+ H or 2+ P or 1+ Rx	0.24	58.7	79.1	0.38	25.3	94.1
	6	1+ H or 2+ P or 2+ Rx	0.22	52.9	80.6	0.34	24.8	93.4
3	7	1+ H or 1+ P or 1+ Rx	0.14	75.0	59.0	0.34	18.1	95.2
	8	1+ H or 2+ P or 1+ Rx	0.22	66.1	73.5	0.40	23.1	94.7
	9	1+ H or 2+ P or 2+ Rx	0.20	59.2	75.4	0.35	22.5	93.9
5	10	1+ H or 1+ P or 1+ Rx	0.12	83.4	49.7	0.33	16.7	96.1
	11	1+ H or 2+ P or 1+ Rx	0.19	75.0	65.7	0.41	20.9	95.6
	12	1+ H or 2+ P or 2+ Rx	0.18	68.0	67.8	0.36	20.3	94.6

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record

Appendix Table C.4: Crude provincial prevalence estimates for additional arthritis algorithms

# Years		Algorithm	Arthritis (%)	RA (%)	OA (%)
1	1	1+ H or 1+ P or 1+ Rx	33.3	1.6	22.7
	2	1+ H or 2+ P or 1+ Rx	26.5	1.3	14.2
	3	1+ H or 2+ P or 2+ Rx	19.6	1.1	12.6
2	4	1+ H or 1+ P or 1+ Rx	46.0	2.3	32.8
	5	1+ H or 2+ P or 1+ Rx	37.9	1.8	21.5
	6	1+ H or 2+ P or 2+ Rx	29.2	1.5	19.5
3	7	1+ H or 1+ P or 1+ Rx	55.0	2.8	40.4
	8	1+ H or 2+ P or 1+ Rx	46.3	2.3	27.3
	9	1+ H or 2+ P or 2+ Rx	36.7	1.9	25.1
5	10	1+ H or 1+ P or 1+ Rx	66.5	3.9	50.9
	11	1+ H or 2+ P or 1+ Rx	57.9	3.1	35.9
	12	1+ H or 2+ P or 2+ Rx	47.9	2.5	33.6

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record; 1-year estimates are for 2002/03, 2-year estimates are for 2001/02 – 2002/03, 3-year estimates are for 2000/01 – 2002/03, 5-year estimates are for 1998/99 – 2002/03.

APPENDIX D: POINT ESTIMATES AND CONFIDENCE INTERVALS FOR VALIDATION INDICES

Appendix Table D.1: 95% confidence intervals for validation indices for arthritis algorithms

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ P	.28	.25	.31	43.2	40.5	45.8	84.3	83.2	85.4	46.6	43.8	49.4	82.4	81.3	83.5
	2	2+ P	.24	.21	.26	25.3	23.0	27.6	93.8	93.1	94.5	56.4	52.4	60.3	79.8	78.7	81.0
	3	1+ H or 2+ P	.24	.21	.27	26.3	23.9	28.6	93.6	92.8	94.3	56.4	52.5	60.3	80.0	78.9	81.1
	4	1+ H or 2+ P or (1 P & 2+ Rx)	.30	.28	.33	34.5	32.0	37.1	91.9	91.0	92.7	57.4	53.9	60.8	81.6	80.5	82.7
2	5	1+ P	.32	.29	.35	60.0	57.4	62.6	75.9	74.6	77.2	44.1	41.8	46.4	85.7	84.6	86.8
	6	2+ P	.32	.30	.35	41.7	39.0	44.3	88.4	87.4	89.4	53.2	50.2	56.2	82.7	81.6	83.8
	7	1+ H or 2+ P	.33	.30	.36	42.6	40.0	45.3	88.1	87.1	89.1	53.2	50.2	56.1	82.9	81.8	84.0
	8	1+ H or 2+ P or (1 P & 2+ Rx)	.37	.34	.39	51.7	49.0	54.4	84.9	83.8	85.9	52.0	49.3	54.7	84.7	83.6	85.8
3	9	1+ P	.31	.28	.33	69.0	66.5	71.4	68.7	67.3	70.1	41.1	39.1	43.2	87.5	86.4	88.6
	10	2+ P	.34	.31	.37	50.9	48.2	53.6	83.7	82.5	84.8	49.7	47.0	52.3	84.3	83.2	85.4
	11	1+ H or 2+ P	.35	.32	.37	51.8	49.1	54.5	83.4	82.3	84.5	49.7	47.1	52.3	84.5	83.4	85.6
	12	1+ H or 2+ P or (1 P & 2+ Rx)	.36	.33	.39	60.3	57.7	63.0	78.9	77.7	80.1	47.6	45.2	49.9	86.3	85.2	87.3
5	13	1+ P	.27	.24	.29	78.1	75.8	80.3	58.6	57.1	60.1	37.4	35.6	39.2	89.4	88.2	90.5
	14	2+ P	.35	.32	.37	63.1	60.5	65.7	76.2	75.0	77.5	45.7	43.4	48.0	86.7	85.6	87.8
	15	1+ H or 2+ P	.35	.32	.37	63.7	61.1	66.3	75.9	74.6	77.2	45.6	43.4	47.9	86.8	85.8	87.9
	16	1+ H or 2+ P or (1 P & 2+ Rx)	.34	.31	.36	71.1	68.6	73.5	70.1	68.7	71.4	42.9	40.9	45.0	88.4	87.3	89.5

Note: Est. = validation index estimate; LCL = Lower confidence limit; UCL = upper confidence limit

Appendix Table D.2: 95% confidence intervals for validation indices for rheumatoid arthritis algorithms

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ P	.12	.08	.16	7.4	5.0	9.8	99.8	99.6	99.9	73.9	61.2	86.6	92.3	91.6	93.0
	2	2+ P	.08	.05	.12	5.0	3.0	7.0	99.9	99.8	100.0	76.7	61.5	91.8	92.2	91.4	92.9
	3	1+ H or 2+ P	.09	.06	.13	5.4	3.4	7.5	99.9	99.8	100.0	78.1	63.8	92.4	92.2	91.5	92.9
	4	1+ H or 2+ P or (1 P & 2+ Rx)	.12	.08	.16	6.3	4.1	8.5	99.9	99.8	100.0	80.6	67.6	93.5	92.3	91.5	93.0
2	5	1+ P	.14	.10	.18	8.9	6.3	11.5	99.6	99.4	99.7	65.1	53.3	76.9	92.4	91.7	93.1
	6	2+ P	.11	.07	.14	6.5	4.3	8.8	99.8	99.6	99.9	71.4	57.8	85.1	92.3	91.6	93.0
	7	1+ H or 2+ P	.11	.08	.15	7.0	4.6	9.3	99.8	99.6	99.9	72.7	59.6	85.9	92.3	91.6	93.0
	8	1+ H or 2+ P or (1 P & 2+ Rx)	.14	.10	.18	7.6	5.2	10.1	99.7	99.6	99.9	72.9	60.3	85.5	92.3	91.6	93.0
3	9	1+ P	.16	.12	.20	10.7	7.9	13.5	99.4	99.2	99.6	62.8	52.1	73.5	92.6	91.9	93.2
	10	2+ P	.13	.09	.17	7.8	5.4	10.3	99.7	99.6	99.9	70.6	58.1	83.1	92.4	91.7	93.1
	11	1+ H or 2+ P	.14	.10	.18	8.5	5.9	11.0	99.7	99.6	99.9	72.2	60.3	84.2	92.4	91.7	93.1
	12	1+ H or 2+ P or (1 P & 2+ Rx)	.14	.10	.18	8.9	6.3	11.5	99.7	99.5	99.8	70.7	59.0	82.4	92.4	91.7	93.1
5	13	1+ P	.17	.12	.21	11.3	8.4	14.2	99.2	99.0	99.4	55.9	45.8	66.0	92.6	91.9	93.3
	14	2+ P	.13	.09	.17	8.3	5.8	10.8	99.7	99.5	99.8	69.1	56.9	81.3	92.4	91.7	93.1
	15	1+ H or 2+ P	.14	.10	.18	8.9	6.3	11.5	99.7	99.5	99.8	70.7	59.0	82.4	92.4	91.7	93.1
	16	1+ H or 2+ P or (1 P & 2+ Rx)	.17	.13	.21	9.4	6.7	12.0	99.6	99.4	99.8	68.3	56.8	79.7	92.5	91.8	93.2

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.3: 95% confidence intervals for validation indices for osteoarthritis algorithms

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ P	.24	.20	.28	23.5	20.1	26.8	95.9	95.3	96.4	40.8	35.6	45.9	91.2	90.5	92.0
	2	2+ P	.16	.12	.20	12.3	9.7	14.9	98.5	98.1	98.8	49.3	41.3	57.3	90.3	89.5	91.1
	3	1+ H or 2+ P	.17	.13	.21	13.1	10.4	15.8	98.3	98.0	98.7	48.5	40.8	56.1	90.4	89.6	91.2
	4	1+ H or 2+ P or (1 P & 2+ Rx)	.22	.18	.26	19.0	15.8	22.1	97.2	96.8	97.7	45.2	39.1	51.4	90.9	90.1	91.6
2	5	1+ P	.30	.26	.34	35.1	31.3	38.9	93.3	92.6	94.0	38.9	34.8	43.0	92.3	91.5	93.0
	6	2+ P	.23	.19	.27	19.8	16.6	23.0	97.2	96.7	97.6	45.8	39.7	51.8	90.9	90.2	91.7
	7	1+ H or 2+ P	.23	.19	.27	20.8	17.6	24.0	97.0	96.5	97.4	45.3	39.4	51.2	91.0	90.3	91.8
	8	1+ H or 2+ P or (1 P & 2+ Rx)	.28	.24	.31	29.8	26.1	33.4	94.7	94.1	95.3	40.5	35.9	45.1	91.8	91.0	92.5
3	9	1+ P	.31	.27	.34	41.4	37.5	45.4	91.2	90.4	92.0	36.3	32.7	39.9	92.8	92.1	93.5
	10	2+ P	.25	.21	.29	24.5	21.0	27.9	95.9	95.3	96.4	41.8	36.6	46.9	91.3	90.6	92.1
	11	1+ H or 2+ P	.25	.21	.29	25.3	21.8	28.8	95.7	95.1	96.2	41.3	36.3	46.3	91.4	90.6	92.2
	12	1+ H or 2+ P or (1 P & 2+ Rx)	.25	.21	.29	34.9	31.1	38.8	92.9	92.2	93.6	37.2	33.2	41.2	92.2	91.5	93.0
5	13	1+ P	.32	.29	.36	49.9	45.9	53.9	88.7	87.8	89.6	34.8	31.6	38.0	93.6	92.9	94.3
	14	2+ P	.29	.25	.33	31.6	27.9	35.3	94.3	93.7	95.0	40.2	35.8	44.6	92.0	91.2	92.7
	15	1+ H or 2+ P	.29	.25	.33	32.8	29.0	36.5	94.0	93.4	94.7	39.8	35.5	44.1	92.1	91.3	92.8
	16	1+ H or 2+ P or (1 P & 2+ Rx)	.31	.28	.35	43.1	39.1	47.1	90.7	89.9	91.5	35.8	32.3	39.3	93.0	92.2	93.7

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.4: 95% confidence intervals for validation indices for asthma algorithms, all ages

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+P	.40	.35	.44	30.8	26.9	34.7	98.8	98.6	99.1	70.3	64.4	76.1	94.1	93.5	94.7
	2	2+ P	.27	.23	.32	18.1	14.9	21.4	99.5	99.3	99.7	77.4	70.1	84.8	93.1	92.5	93.7
	3	1+ Rx	.51	.47	.55	55.4	51.2	59.6	95.9	95.4	96.4	54.7	50.4	58.9	96.0	95.5	96.5
	4	1+H or 1+P	.40	.36	.43	31.4	27.4	35.3	98.8	98.5	99.1	70.3	64.5	76.2	94.1	93.5	94.7
	5	1+ H or 2+ P	.28	.23	.32	18.9	15.6	22.2	99.5	99.3	99.7	77.5	70.3	84.7	93.2	92.6	93.8
	6	1+H or 1+P or 1+Rx	.52	.48	.56	58.4	54.2	62.6	95.5	95.0	96.0	53.7	49.7	57.8	96.2	95.7	96.7
	7	1+ H or 2+ P or 2+ Rx	.51	.47	.55	46.9	42.6	51.1	97.7	97.3	98.1	64.4	59.6	69.2	95.3	94.8	95.9
2	8	1+P	.37	.32	.41	43.9	39.6	48.1	98.0	97.6	98.3	66.1	61.1	71.0	95.1	94.6	95.6
	9	2+ P	.26	.22	.30	30.1	26.1	34.0	99.1	98.8	99.3	74.3	68.4	80.2	94.0	93.4	94.6
	10	1+ Rx	.54	.50	.57	69.8	65.8	73.7	93.7	93.1	94.3	49.9	46.3	53.5	97.2	96.8	97.6
	11	1+H or 1+ P	.38	.34	.43	44.8	40.6	49.0	97.9	97.6	98.3	66.2	61.3	71.1	95.2	94.6	95.7
	12	1+ H or 2+ P	.28	.24	.33	31.2	27.2	35.1	99.0	98.8	99.3	74.3	68.6	80.1	94.1	93.5	94.7
	13	1+H or 1+P or 1+Rx	.54	.50	.57	72.6	68.8	76.4	93.1	92.4	93.7	48.4	44.9	51.9	97.4	97.0	97.8
	14	1+ H or 2+ P or 2+ Rx	.55	.52	.59	57.5	53.3	61.7	96.7	96.2	97.1	60.9	56.6	65.2	96.2	95.7	96.7
3	15	1+P	.36	.31	.40	52.9	48.7	57.2	96.9	96.5	97.4	60.7	56.3	65.2	95.8	95.3	96.3
	16	2+ P	.24	.20	.29	37.6	33.5	41.7	98.6	98.3	98.9	71.3	66.0	76.6	94.6	94.1	95.2
	17	1+ Rx	.52	.49	.56	74.7	71.0	78.4	92.3	91.6	93.0	46.5	43.2	49.9	97.6	97.2	98.0
	18	1+H or 1+P	.37	.42	.33	53.7	49.4	57.9	96.9	96.4	97.3	60.8	56.4	65.2	95.9	95.4	96.4
	19	1+ H or 2+ P	.28	.23	.32	38.6	34.4	42.7	98.6	98.3	98.9	71.1	65.8	76.3	94.7	94.1	95.3
	20	1+H or 1+P or 1+Rx	.52	.48	.55	77.7	74.1	81.2	91.2	90.5	91.9	44.2	41.0	47.4	97.9	97.5	98.2
	21	1+ H or 2+ P or 2+ Rx	.58	.54	.61	66.2	62.1	70.2	95.7	95.1	96.2	57.8	53.8	61.7	96.9	96.5	97.4
5	22	1+P	.45	.41	.49	63.5	59.4	67.6	95.3	94.8	95.9	55.0	51.0	58.9	96.7	96.2	97.1
	23	2+ P	.35	.30	.39	50.3	46.0	54.5	97.7	97.3	98.1	66.2	61.5	70.8	95.6	95.1	96.1
	24	1+ Rx	.51	.47	.54	81.5	78.2	84.8	90.2	89.4	90.9	42.7	39.6	45.7	98.2	97.8	98.5
	25	1+H or 1+P	.48	.46	.50	63.7	59.6	67.8	95.3	94.8	95.8	54.9	51.0	58.8	96.7	96.2	97.2
	26	1+ H or 2+ P	.37	.32	.41	50.7	46.4	54.9	97.6	97.2	98.0	65.7	61.1	70.3	95.7	95.1	96.2
	27	1+H or 1+P or 1+Rx	.50	.47	.53	84.3	81.2	87.4	88.6	87.8	89.4	39.9	37.0	42.7	98.4	98.1	98.8
	28	1+ H or 2+ P or 2+ Rx	.59	.56	.62	75.4	71.8	79.1	94.2	93.6	94.8	53.7	50.1	57.3	97.7	97.3	98.1

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.5: 95% confidence intervals for validation indices for asthma algorithms, 12-18 years

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+P	.37	.27	.47	29.7	21.2	38.2	98.1	97.1	99.1	70.2	57.1	83.3	90.1	88.0	92.2
	2	2+ P	.24	.14	.33	16.2	9.4	23.1	99.4	98.9	100.0	81.8	65.7	97.9	88.5	86.3	90.7
	3	1+ Rx	.48	.38	.58	43.2	34.0	52.5	97.2	96.0	98.4	70.6	59.8	81.4	91.8	89.8	93.7
	4	1+H or 1+P	.37	.27	.47	29.7	21.2	38.2	98.1	97.1	99.1	70.2	57.1	83.3	90.1	88.0	92.2
	5	1+ H or 2+ P	.24	.14	.33	16.2	9.4	23.1	99.4	98.9	100.0	81.8	65.7	97.9	88.5	86.3	90.7
	6	1+ H or 1+ P or 1+ Rx	.52	.43	.61	48.6	39.4	57.9	96.8	95.5	98.1	70.1	59.9	80.4	92.5	90.6	94.3
	7	1+ H or 2+ P or 2+ Rx	.42	.32	.52	33.3	24.6	42.1	98.6	97.8	99.5	78.7	67.0	90.4	90.6	88.5	92.6
2	8	1+P	.43	.33	.53	45.9	36.7	55.2	96.5	95.2	97.9	67.1	56.5	77.7	92.1	90.1	94.0
	9	2+ P	.30	.20	.40	33.3	24.6	42.1	99.0	98.3	99.7	84.1	73.3	94.9	90.6	88.6	92.7
	10	1+ Rx	.64	.56	.72	67.6	58.9	76.3	95.4	93.9	97.0	69.4	60.8	78.1	95.0	93.5	96.6
	11	1+H or 1+ P	.44	.34	.54	45.9	36.7	55.2	96.5	95.2	97.9	67.1	56.5	77.7	92.1	90.1	94.0
	12	1+ H or 2+ P	.31	.21	.41	33.3	24.6	42.1	99.0	98.3	99.7	84.1	73.3	94.9	90.6	88.6	92.7
	13	1+ H or 1+ P or 1+ Rx	.63	.55	.71	69.4	60.8	77.9	94.5	92.8	96.1	65.8	57.2	74.4	95.3	93.7	96.8
	14	1+ H or 2+ P or 2+ Rx	.55	.46	.64	48.6	39.4	57.9	98.2	97.2	99.2	80.6	71.1	90.1	92.6	90.7	94.4
3	15	1+P	.46	.37	.56	62.2	53.1	71.2	95.2	93.6	96.7	66.3	57.3	75.4	94.2	92.5	95.9
	16	2+ P	.30	.21	.40	47.7	38.5	57.0	98.3	97.4	99.3	81.5	72.1	91.0	92.4	90.6	94.3
	17	1+ Rx	.66	.58	.73	76.6	68.7	84.5	93.9	92.2	95.7	65.9	57.7	74.1	96.3	94.9	97.7
	18	1+H or 1+P	.48	.39	.58	62.2	53.1	71.2	95.2	93.6	96.7	66.3	57.3	75.4	94.2	92.5	95.9
	19	1+ H or 2+ P	.34	.24	.44	47.7	38.5	57.0	98.3	97.4	99.3	81.5	72.1	91.0	92.4	90.6	94.3
	20	1+ H or 1+ P or 1+ Rx	.66	.58	.73	78.4	70.7	86.0	92.4	90.4	94.3	61.3	53.3	69.3	96.5	95.2	97.9
	21	1+ H or 2+ P or 2+ Rx	.68	.61	.76	69.4	60.8	77.9	96.8	95.5	98.1	77.0	68.8	85.2	95.4	93.8	96.9
5	22	1+P	.49	.40	.58	74.8	66.7	82.9	92.0	90.0	93.9	58.9	50.7	67.0	96.0	94.5	97.4
	23	2+ P	.36	.26	.46	60.4	51.3	69.5	96.7	95.4	98.0	73.6	64.6	82.7	94.1	92.4	95.8
	24	1+ Rx	.64	.57	.71	86.5	80.1	92.8	90.7	88.6	92.8	58.9	51.3	66.4	97.8	96.6	98.9
	25	1+H or 1+P	.51	.42	.60	74.8	66.7	82.9	92.0	90.0	93.9	58.9	50.7	67.0	96.0	94.5	97.4
	26	1+ H or 2+ P	.38	.28	.48	60.4	51.3	69.5	96.7	95.4	98.0	73.6	64.6	82.7	94.1	92.4	95.8
	27	1+ H or 1+ P or 1+ Rx	.63	.56	.70	87.4	81.2	93.6	88.5	86.2	90.8	53.9	46.6	61.2	97.9	96.7	99.0
	28	1+ H or 2+ P or 2+ Rx	.70	.62	.77	80.2	72.8	87.6	94.6	92.9	96.2	69.5	61.6	77.5	96.9	95.6	98.2

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.6: 95% confidence intervals for validation indices for asthma algorithms, 19-49 years

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+P	.39	.32	.46	29.8	23.9	35.8	99.0	98.7	99.4	71.6	62.5	80.6	94.5	93.7	95.3
	2	2+ P	.26	.19	.32	17.1	12.2	22.0	99.5	99.2	99.8	73.6	61.7	85.5	93.6	92.7	94.5
	3	1+ Rx	.52	.46	.58	50.0	43.5	56.5	97.5	96.9	98.1	62.3	55.3	69.3	96.0	95.2	96.7
	4	1+H or 1+P	.40	.33	.47	30.3	24.3	36.2	99.0	98.7	99.4	71.9	62.9	80.9	94.5	93.7	95.4
	5	1+ H or 2+ P	.27	.20	.34	18.0	13.0	23.0	99.5	99.2	99.8	74.5	63.0	86.1	93.7	92.8	94.5
	6	1+H or 1+P or 1+Rx	.53	.47	.59	53.5	47.0	60.0	97.1	96.5	97.7	60.1	53.4	66.8	96.2	95.5	96.9
	7	1+ H or 2+ P or 2+ Rx	.52	.45	.58	41.7	35.3	48.1	99.0	98.7	99.4	77.9	70.5	85.2	95.4	94.6	96.2
2	8	1+P	.37	.30	.43	42.5	36.1	49.0	98.5	98.0	98.9	69.3	61.6	76.9	95.4	94.7	96.2
	9	2+ P	.24	.18	.31	28.9	23.1	34.8	99.1	98.8	99.5	73.3	64.2	82.5	94.4	93.6	95.3
	10	1+ Rx	.56	.51	.62	65.4	59.2	71.5	95.7	95.0	96.5	55.6	49.6	61.5	97.1	96.5	97.7
	11	1+H or 1+ P	.38	.31	.45	43.0	36.6	49.4	98.4	98.0	98.9	69.0	61.4	76.6	95.5	94.7	96.2
	12	1+ H or 2+ P	.26	.18	.33	29.8	23.9	35.8	99.1	98.7	99.5	73.1	64.1	82.1	94.5	93.7	95.3
	13	1+ H or 1+ P or 1+ Rx	.57	.52	.63	70.2	64.2	76.1	95.1	94.3	95.9	53.9	48.2	59.5	97.5	96.9	98.1
	14	1+ H or 2+ P or 2+ Rx	.59	.53	.64	53.5	47.0	60.0	98.4	97.9	98.9	73.1	66.3	79.8	96.3	95.6	97.0
3	15	1+P	.35	.28	.42	48.7	42.2	55.2	97.6	97.0	98.1	62.0	54.9	69.1	95.9	95.1	96.6
	16	2+ P	.21	.15	.28	34.2	28.1	40.4	98.8	98.5	99.2	70.9	62.4	79.4	94.8	94.0	95.6
	17	1+ Rx	.55	.50	.60	70.2	64.2	76.1	94.4	93.5	95.2	50.6	45.1	56.1	97.5	96.9	98.1
	18	1+H or 1+P	.35	.28	.42	49.1	42.6	55.6	97.5	96.9	98.1	61.9	54.8	69.0	95.9	95.2	96.6
	19	1+ H or 2+ P	.24	.17	.30	35.1	28.9	41.3	98.8	98.4	99.2	70.8	62.4	79.2	94.9	94.1	95.7
	20	1+ H or 1+ P or 1+ Rx	.54	.48	.59	74.6	68.9	80.2	93.3	92.4	94.2	47.8	42.6	52.9	97.8	97.3	98.4
	21	1+ H or 2+ P or 2+ Rx	.59	.53	.65	60.1	53.7	66.4	97.6	97.1	98.2	67.5	61.0	73.9	96.8	96.1	97.4
5	22	1+P	.47	.41	.54	60.5	54.2	66.9	95.9	95.2	96.6	54.8	48.6	60.9	96.7	96.1	97.4
	23	2+ P	.35	.28	.42	50.4	43.9	56.9	97.8	97.2	98.3	65.0	57.9	72.0	96.0	95.3	96.7
	24	1+ Rx	.53	.48	.57	78.1	72.7	83.4	92.2	91.2	93.2	45.1	40.2	50.0	98.1	97.6	98.6
	25	1+H or 1+P	.47	.41	.54	60.5	54.2	66.9	95.9	95.1	96.6	54.5	48.4	60.7	96.7	96.1	97.4
	26	1+ H or 2+ P	.36	.29	.42	50.4	43.9	56.9	97.7	97.2	98.3	64.6	57.6	71.6	96.0	95.3	96.7
	27	1+ H or 1+ P or 1+ Rx	.51	.46	.56	82.0	77.0	87.0	90.4	89.3	91.5	41.3	36.7	45.8	98.4	97.9	98.9
	28	1+ H or 2+ P or 2+ Rx	.61	.55	.66	71.1	65.2	76.9	96.0	95.2	96.7	59.1	53.3	54.9	97.6	97.0	98.2

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.7: 95% confidence intervals for validation indices for asthma algorithms, 50+ years

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+P	.42	.34	.49	32.6	26.0	39.3	98.8	98.4	99.3	68.9	59.3	78.5	94.9	94.0	95.7
	2	2+ P	.30	.23	.38	20.5	14.8	26.3	99.6	99.3	99.8	79.6	68.3	90.9	94.0	93.1	95.0
	3	1+ Rx	.51	.45	.56	68.9	62.4	75.5	93.6	92.6	94.5	46.0	40.2	51.8	97.4	96.8	98.1
	4	1+H or 1+P	.42	.35	.50	33.7	27.0	40.4	98.8	98.3	99.2	68.8	59.4	78.2	94.9	94.1	95.8
	5	1+ H or 2+ P	.32	.24	.39	21.6	15.7	27.4	99.5	99.3	99.8	78.8	67.7	89.9	94.1	93.2	95.0
	6	1+H or 1+P or 1+Rx	.50	.45	.56	70.0	63.5	76.5	93.2	92.2	94.2	45.1	39.4	50.8	97.5	96.9	98.1
	7	1+ H or 2+ P or 2+ Rx	.54	.47	.60	61.1	54.1	68.0	95.8	95.0	96.6	53.7	47.1	60.4	96.9	96.2	97.6
2	8	1+P	.33	.26	.40	44.2	37.1	51.3	97.9	97.3	98.4	62.2	54.0	70.4	95.7	94.9	96.5
	9	2+ P	.25	.18	.33	29.5	23.0	36.0	99.0	98.6	99.4	70.0	60.0	80.0	94.6	93.8	95.5
	10	1+ Rx	.47	.42	.53	76.3	70.3	82.4	90.9	89.7	92.0	39.9	34.9	45.0	98.0	97.4	98.6
	11	1+H or 1+ P	.36	.28	.43	46.3	39.2	53.4	97.8	97.2	98.4	62.9	54.9	70.9	95.8	95.0	96.6
	12	1+ H or 2+ P	.28	.21	.36	31.6	25.0	38.2	99.0	98.5	99.4	70.6	60.9	80.3	94.8	93.9	95.7
	13	1+H or 1+P or 1+Rx	.47	.41	.52	77.4	71.4	83.3	90.3	89.1	91.5	38.8	33.9	43.7	98.0	97.5	98.6
	14	1+ H or 2+ P or 2+ Rx	.52	.47	.58	67.4	60.7	74.0	94.3	93.3	95.2	48.3	42.3	54.3	97.3	96.7	98.0
3	15	1+P	.31	.23	.38	52.6	45.5	59.7	96.7	96.0	97.4	56.2	48.9	63.5	96.3	95.5	97.0
	16	2+ P	.24	.17	.31	35.8	29.0	42.6	98.5	98.0	99.0	65.4	56.2	74.5	95.1	94.2	95.9
	17	1+ Rx	.45	.40	.50	78.9	73.2	84.7	89.4	88.1	90.6	37.1	32.4	41.8	98.2	97.6	98.7
	18	1+H or 1+P	.33	.26	.40	54.2	47.1	61.3	96.7	96.0	97.4	56.6	49.4	63.8	96.4	95.6	97.1
	19	1+ H or 2+ P	.28	.20	.35	37.4	30.5	44.2	98.4	97.9	98.9	65.1	56.2	74.1	95.2	94.3	96.0
	20	1+H or 1+P or 1+Rx	.44	.39	.49	81.1	75.5	86.6	88.4	87.1	89.7	35.7	31.2	40.3	98.3	97.8	98.9
	21	1+ H or 2+ P or 2+ Rx	.52	.46	.58	71.6	65.2	78.0	93.0	92.0	94.0	44.9	39.3	50.5	97.6	97.0	98.3
5	22	1+P	.39	.32	.46	60.5	53.6	67.5	95.7	94.9	96.5	52.8	46.1	59.4	96.8	96.1	97.5
	23	2+ P	.34	.26	.41	44.2	37.1	51.3	97.9	97.3	98.5	62.7	54.5	70.9	95.7	94.9	96.5
	24	1+ Rx	.43	.38	.48	82.6	77.2	88.0	87.7	86.3	89.0	34.7	30.3	39.1	98.4	97.9	99.0
	25	1+H or 1+P	.41	.34	.48	61.1	54.1	68.0	95.6	94.8	96.5	52.7	46.1	59.3	96.9	96.2	97.6
	26	1+ H or 2+ P	.37	.30	.44	45.3	38.2	52.3	97.8	97.2	98.4	61.9	53.8	69.9	95.7	94.9	96.5
	27	1+H or 1+P or 1+Rx	.43	.38	.48	85.3	80.2	90.3	86.4	85.1	87.8	33.3	29.1	37.5	98.7	98.2	99.2
	28	1+ H or 2+ P or 2+ Rx	.56	.47	.60	77.9	72.0	83.8	91.9	90.8	93.0	43.4	38.1	48.7	98.1	97.6	98.7

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.8: 95% confidence intervals for validation indices for coronary heart disease algorithms

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ P	.44	.39	.49	40.2	35.2	45.2	97.9	97.5	98.2	57.1	51.1	63.1	95.8	95.3	96.4
	2	2+ P	.37	.31	.42	28.6	24.0	33.2	98.8	98.5	99.1	62.4	55.1	69.6	95.1	94.5	95.7
	3	1+ H or 1+ P	.46	.41	.51	42.9	37.8	47.9	97.7	97.3	98.1	57.0	51.2	62.8	96.0	95.5	96.5
	4	1+ H or 2+ P	.40	.34	.45	32.3	27.6	37.1	98.6	98.2	98.9	61.5	54.7	68.4	95.3	94.8	95.9
	5	1+ H or 2+ P or (1 P & 2+ Rx)	.46	.41	.51	41.0	36.0	46.0	98.1	97.7	98.5	60.3	54.3	66.4	95.9	95.4	96.4
2	6	1+ P	.51	.46	.55	52.8	47.8	57.9	96.9	96.5	97.4	55.1	49.9	60.2	96.7	96.2	97.1
	7	2+ P	.47	.42	.52	41.8	36.8	46.8	98.1	97.7	98.4	60.5	54.6	66.5	95.9	95.4	96.5
	8	1+ H or 1+ P	.52	.47	.57	55.5	50.5	60.6	96.7	96.3	97.2	54.8	49.8	59.8	96.8	96.4	97.3
	9	1+ H or 2+ P	.49	.44	.54	45.6	40.5	50.6	97.8	97.4	98.2	59.7	54.0	65.4	96.2	95.7	96.7
	10	1+ H or 2+ P or (1 P & 2+ Rx)	.52	.48	.57	53.1	48.0	58.2	97.3	96.8	97.7	57.9	52.7	63.2	96.7	96.2	97.2
3	11	1+ P	.52	.48	.56	58.5	53.5	63.5	96.2	95.7	96.7	52.4	47.6	57.2	97.0	96.6	97.5
	12	2+ P	.50	.46	.55	48.2	43.2	53.3	97.7	97.3	98.1	59.5	53.9	65.0	96.4	95.9	96.9
	13	1+ H or 1+ P	.53	.49	.59	61.5	56.5	66.4	96.0	95.5	96.6	52.4	47.7	57.1	97.2	96.8	97.7
	14	1+ H or 2+ P	.53	.48	.58	53.4	48.3	58.4	97.4	96.9	97.8	59.1	53.8	64.4	96.7	96.2	97.2
	15	1+ H or 2+ P or (1 P & 2+ Rx)	.55	.50	.59	60.1	55.1	65.1	96.6	96.1	97.1	55.9	51.0	60.8	97.1	96.7	97.6
5	16	1+ P	.52	.48	.57	65.2	60.4	70.1	95.2	94.6	95.8	49.2	44.8	53.6	97.5	97.0	97.9
	17	2+ P	.54	.49	.58	56.6	51.6	61.6	96.9	96.5	97.4	56.8	51.7	61.8	96.9	96.4	97.4
	18	1+ H or 1+ P	.53	.49	.57	67.9	63.2	72.7	95.0	94.4	95.6	49.0	44.7	53.3	97.7	97.2	98.1
	19	1+ H or 2+ P	.55	.51	.59	60.4	55.4	65.4	96.6	96.1	97.1	55.9	51.0	60.7	97.2	96.7	97.6
	20	1+ H or 2+ P or (1 P & 2+ Rx)	.55	.51	.60	66.6	61.8	71.4	95.7	95.2	96.3	52.6	48.0	57.1	97.6	97.2	98.0

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.9: 95% confidence intervals for validation indices for diabetes algorithms

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ H or 1+ P	.77	.73	.80	76.9	72.4	81.4	98.7	98.4	99.0	79.2	74.8	83.6	98.5	98.2	98.8
	2	1+ H or 2+ P	.73	.69	.77	63.2	58.1	68.4	99.5	99.3	99.7	89.5	85.6	93.4	97.7	97.3	98.1
	3	1+ H or 1+ P or 1+ Rx	.81	.78	.85	85.8	82.0	89.5	98.6	98.3	98.9	79.4	75.2	83.6	99.1	98.8	99.3
	4	1+ H or 2+ P or 1+ Rx	.84	.81	.87	80.7	76.5	84.9	99.4	99.1	99.6	88.9	85.4	92.4	98.8	98.5	99.1
	5	1+ H or 1+ P or 2+ Rx	.81	.78	.85	85.5	81.7	89.2	98.6	98.3	98.9	79.8	75.6	83.9	99.1	98.8	99.3
	6	1+ H or 2+ P or 2+ Rx	.84	.80	.87	80.1	75.9	84.4	99.4	99.2	99.6	89.4	85.9	92.9	98.7	98.4	99.0
2	7	1+ H or 1+ P	.78	.74	.81	85.2	81.4	89.0	98.1	97.7	98.4	74.0	69.6	78.3	99.0	98.8	99.3
	8	1+ H or 2+ P	.82	.79	.86	79.5	75.2	83.8	99.3	99.1	99.5	87.9	84.2	91.5	98.7	98.4	99.0
	9	1+ H or 1+ P or 1+ Rx	.80	.76	.83	89.6	86.4	92.9	97.9	97.6	98.3	73.7	69.4	77.9	99.3	99.1	99.5
	10	1+ H or 2+ P or 1+ Rx	.86	.83	.88	86.6	83.0	90.3	99.1	98.9	99.4	86.1	82.5	89.8	99.1	98.9	99.4
	11	1+ H or 1+ P or 2+ Rx	.80	.76	.83	89.3	86.0	92.6	98.0	97.6	98.4	74.0	69.7	78.2	99.3	99.1	99.5
	12	1+ H or 2+ P or 2+ Rx	.86	.83	.88	86.1	82.4	89.8	99.2	98.9	99.4	86.8	83.2	90.5	99.1	98.9	99.4
3	13	1+ H or 1+ P	.75	.72	.79	87.8	84.3	91.3	97.4	97.0	97.9	68.7	64.3	73.1	99.2	99.0	99.4
	14	1+ H or 2+ P	.83	.80	.86	84.9	81.0	88.7	99.0	98.7	99.2	83.9	80.0	87.8	99.0	98.8	99.3
	15	1+ H or 1+ P or 1+ Rx	.76	.73	.80	90.5	87.4	93.6	97.3	96.9	97.7	68.2	63.9	72.5	99.4	99.2	99.6
	16	1+ H or 2+ P or 1+ Rx	.84	.81	.87	88.4	85.0	91.8	98.8	98.5	99.1	82.1	78.1	86.0	99.3	99.0	99.5
	17	1+ H or 1+ P or 2+ Rx	.76	.73	.80	90.2	87.0	93.4	97.4	96.9	97.8	68.6	64.3	72.9	99.4	99.1	99.6
	18	1+ H or 2+ P or 2+ Rx	.85	.82	.87	88.1	84.7	91.6	98.8	98.5	99.1	83.0	79.1	86.9	99.2	99.0	99.5

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.10: 95% confidence intervals for validation indices for hypertension algorithms

# of Years		Algorithm	κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ P	.65	.63	.68	68.4	65.6	71.3	94.7	94.1	95.4	74.6	71.8	77.3	93.0	92.2	93.7
	2	2+ P	.54	.51	.57	48.4	45.4	51.5	97.5	97.0	97.9	81.4	78.4	84.5	89.3	88.4	90.1
	3	1+ H or 1+ P	.66	.63	.68	70.1	67.3	72.9	94.4	93.7	95.0	73.9	71.1	76.6	93.3	92.6	94.0
	4	1+ H or 2+ P	.56	.53	.59	51.1	48.1	54.2	97.1	96.6	97.6	79.9	76.8	82.9	89.7	88.9	90.6
	5	1+ H or 1+ P or 1+ Rx	.68	.66	.70	89.9	88.1	91.8	88.7	87.8	89.6	64.4	61.9	66.9	97.5	97.0	98.0
	6	1+ H or 1+ P or 2+ Rx	.70	.68	.72	89.0	87.1	90.9	89.9	89.0	90.8	66.6	64.2	69.1	97.3	96.8	97.8
2	7	1+ P	.67	.65	.70	79.4	76.9	81.8	91.9	91.1	92.7	69.1	66.5	71.7	95.2	94.5	95.8
	8	2+ P	.66	.63	.68	66.3	63.4	69.2	95.6	95.0	96.2	77.5	74.7	80.2	92.6	91.8	93.3
	9	1+ H or 1+ P	.68	.66	.71	81.2	78.8	83.6	91.6	90.8	92.5	68.8	66.2	71.4	95.6	94.9	96.2
	10	1+ H or 2+ P	.67	.65	.70	69.4	66.6	72.2	95.2	94.6	95.9	76.8	74.1	79.5	93.2	92.5	93.9
	11	1+ H or 1+ P or 1+ Rx	.64	.62	.67	91.9	90.2	93.5	86.0	84.9	87.0	59.8	57.3	62.2	97.9	97.5	98.3
	12	1+H or 1+ P or 2+ Rx	.66	.64	.69	91.2	89.5	92.9	87.3	86.4	88.3	62.0	59.6	64.5	97.8	97.3	98.2
3	13	1+ P	.67	.64	.69	83.2	80.9	85.4	90.3	89.4	91.1	66.0	63.5	68.6	95.9	95.3	96.5
	14	2+ P	.68	.66	.71	72.4	69.7	75.1	94.8	94.2	95.5	76.0	73.3	78.7	93.8	93.1	94.5
	15	1+ H or 1+ P	.67	.65	.70	84.9	82.7	87.1	89.9	89.1	90.8	65.7	63.1	68.2	96.3	95.8	96.9
	16	1+ H or 2+ P	.70	.67	.72	75.6	73.0	78.2	94.4	93.7	95.0	75.2	72.6	77.9	94.5	93.8	95.1
	17	1+ H or 1+ P or 1+ Rx	.61	.59	.64	92.8	91.3	94.4	84.0	82.9	85.0	56.8	54.4	59.1	98.1	97.7	98.5
	18	1+H or 1+ P or 2+ Rx	.64	.62	.66	92.2	90.5	93.8	85.7	84.7	86.7	59.4	56.9	61.8	98.0	97.5	98.4

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

Appendix Table D.11: 95% confidence intervals for validation indices for stroke algorithms

# of Years	Algorithm		κ			Sensitivity			Specificity			PPV			NPV		
			Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL	Est.	LCL	UCL
1	1	1+ H	.12	.04	.20	7.3	2.4	12.2	99.8	99.7	99.9	42.1	19.9	64.3	98.3	98.0	98.6
	2	1+ H or 1+ P	.35	.26	.44	28.4	20.0	36.9	99.5	99.3	99.6	49.2	36.9	61.6	98.7	98.4	99.0
	3	1+ H or 2+ P or (1P & 2+ Rx)	.35	.25	.44	26.6	18.3	34.9	99.6	99.4	99.7	52.7	39.5	65.9	98.6	98.3	98.9
	1f	1+ H	.22	.13	.31	14.7	8.0	21.3	99.7	99.6	99.9	51.6	34.0	69.2	98.4	98.1	98.7
	2f	1+ H or 1+ P	.39	.30	.48	34.9	25.9	43.8	99.2	99.0	99.5	46.3	35.5	57.1	98.8	98.5	99.1
	3f	1+ H or 2+ P or (1P & 2+ Rx)	.40	.31	.50	33.9	25.1	42.8	99.4	99.2	99.6	52.9	41.2	64.6	98.8	98.5	99.1
2	4	1+ H	.24	.14	.33	15.6	8.8	22.4	99.8	99.7	99.9	56.7	38.9	74.4	98.4	98.1	98.8
	5	1+ H or 1+ P	.44	.35	.52	43.1	33.8	52.4	99.1	98.8	99.3	46.5	36.8	56.3	98.9	98.7	99.2
	6	1+ H or 2+ P or (1P & 2+ Rx)	.43	.34	.52	38.5	29.4	47.7	99.3	99.1	99.5	51.2	40.4	62.0	98.9	98.6	99.1
	4f	1+ H	.33	.23	.42	23.9	15.9	31.9	99.7	99.5	99.8	56.5	42.2	70.8	98.6	98.3	98.9
	5f	1+ H or 1+ P	.42	.34	.50	46.8	37.4	56.2	98.7	98.4	99.0	40.5	31.9	49.0	99.0	98.7	99.3
	6f	1+ H or 2+ P or (1P & 2+ Rx)	.43	.35	.52	42.2	32.9	51.5	99.1	98.9	99.3	46.9	37.1	56.8	98.9	98.7	99.2
3	7	1+ H	.26	.17	.35	18.3	11.1	25.6	99.7	99.5	99.8	50.0	34.5	65.5	98.5	98.2	98.8
	8	1+ H or 1+ P	.42	.35	.50	47.7	38.3	57.1	98.7	98.4	99.0	40.3	31.8	48.8	99.0	98.8	99.3
	9	1+ H or 2+ P or (1P & 2+ Rx)	.44	.35	.52	44.0	34.7	53.4	99.0	98.8	99.3	45.7	36.2	55.2	99.0	98.7	99.2
	7f	1+ H	.38	.28	.47	29.4	20.8	37.9	99.6	99.4	99.7	55.2	42.4	68.0	98.7	98.4	99.0
	8f	1+ H or 1+ P	.42	.35	.50	54.1	44.8	63.5	98.3	97.9	98.6	36.6	29.2	44.1	99.1	98.9	99.4
	9f	1+ H or 2+ P or (1P & 2+ Rx)	.45	.37	.53	49.5	40.2	58.9	98.8	98.5	99.1	43.2	34.5	51.9	99.1	98.8	99.3
5	10	1+ H	.41	.32	.51	33.0	24.2	41.9	99.6	99.4	99.7	58.1	45.8	70.3	98.8	98.5	99.0
	11	1+ H or 1+ P	.45	.37	.52	60.6	51.4	69.7	98.1	97.7	98.4	37.1	30.0	44.2	99.3	99.0	99.5
	12	1+ H or 2+ P or (1P & 2+ Rx)	.46	.38	.54	54.1	44.8	63.5	98.6	98.3	98.9	42.1	34.0	50.3	99.1	98.9	99.4
	10f	1+ H	.45	.36	.54	39.4	30.3	48.6	99.4	99.2	99.6	53.8	42.8	64.7	98.9	98.6	99.1
	11f	1+ H or 1+ P	.43	.36	.49	66.1	57.2	74.9	97.5	97.1	97.9	33.0	26.8	39.3	99.4	99.1	99.6
	12f	1+ H or 2+ P or (1P & 2+ Rx)	.47	.40	.54	61.5	52.3	70.6	98.3	97.9	98.6	39.9	32.5	47.3	99.3	99.1	99.5

Note: Est. = validation index point estimate; LCL = lower confidence limit; UCL = upper confidence limit

APPENDIX E: SUPPLEMENTARY DATA FOR ASTHMA ALGORITHMS

Appendix Table E.1: Supplementary data for asthma algorithms

Group	ATC Code	Generic Name	Route	Notes
R03A ADRENERGICS, INHALANTS	R03AA01	EPINEPHRINE	Inhaled	Include English product names Vaponefrin, Procaterol or DINs '00900400', '00900700'
	R03AB02	ISOPROTERENOL	Inhaled	Include only DINs '00026298', '00026301', '00033227', '00033219', '01923870', '01928449', '02017660'. Exclude all other DINs.
	R03AB03	ORCIPRENALINE	Inhaled	
	R03AC02	SALBUTAMOL	Inhaled	Include only DINs '00444774', '00786616', '00818739', '00980838'. Exclude all other DINs.
	R03AC03	TERBUTALINE	Inhaled	
	R03AC04	FENOTEROL	Inhaled	
	R03AC08	PIRBUTEROL	Inhaled	
	R03AC12	SALMETEROL	Inhaled	
	R03AC13	FORMOTEROL	Inhaled	
	R03AK01	EPINEPHRINE	Inhaled	Exclude only DINs '02163705', '02163713', '00824216'. Exclude all other DINs.
	R03AK03	IPRATROPIUM/FENOTEROL	Inhaled	
	R03AK04	IPRATROPIUM/SALBUTAMOL	Inhaled	
	R03AK06	FLUTICASONE/SALMETEROL	Inhaled	Include English product names Advair only
R03B OTHER DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES, INHALANTS	R03BA01	BECLOMETHASONE	Inhaled	
	R03BA02	BUDESONIDE	Inhaled	
	R03BA03	FLUNISOLIDE	Inhaled	
	R03BA05	FLUTICASONE	Inhaled	
	R03BA06	TRIAMCINOLONE	Inhaled	
	R03BB01	IPRATROPIUM SODIUM	Inhaled	
	R03BC01	CROMOGLYCATE	Inhaled	
	R03BC03	NEDOCROMIL	Oral	

R03C ADRENERGICS FOR SYSTEMIC USE	R03CB01	ISOPROTERENOL	Oral
	R03CB03	ORCIPRENALINE	Oral
	R03CC02	SALBUTAMOL	Oral
	R03CC03	TERBUTALINE	Oral
	R03CC07	PIRBUTEROL	Oral
	R03CC53	TERBUTALINE	Oral
	R03CK	BUDESONIDE/FORMOTEROL	Inhaled
R03D OTHER SYSTEMIC DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	R03DA02	OXTRIPHYLLINE	Oral
	R03DA04	THEOPHYLLINE	Oral
	R03DA05	AMINOPHYLLINE	Oral
	R03DA43	THEOPHYLLINE	Oral
	R03DA53	THEOPHYLLINE	Oral
	R03DA54	THEOPHYLLINE	Oral
	R03DA55	AMINOPHYLLINE	Oral
	R03DA74	THEOPHYLLINE	Oral
	R03DB05	AMINOPHYLLINE	Oral
	R03DC01	ZAFIRLUKAST	Oral
	R03DC03	MONTELUKAST	Oral
R06A ANTIHISTAMINES FOR SYSTEMIC USE	R06AX17	KETOTIFEN	Oral

APPENDIX F: ADDITIONAL VALIDATION RESULTS FOR ASTHMA ALGORITHMS

Appendix Table F.1: Estimates of agreement, sensitivity, specificity, and predictive values for asthma algorithms, all ages (COPD & emphysema removed)

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.39	30.6	98.7	0.29	66.8	94.3
	2	2+ P	0.27	17.9	99.4	0.17	73.4	93.4
	3	1+ Rx	0.48	53.6	95.5	0.49	50.7	96.0
	4	1+ H or 1+ P	0.39	31.0	98.7	0.30	66.5	94.3
	5	1+ H or 2+ P	0.27	18.5	99.4	0.18	72.9	93.4
	6	1+ H or 1+ P or 1+ Rx	0.49	56.8	95.1	0.52	50.1	96.3
	7	1+ H or 2+ P or 2+ Rx	0.47	44.8	97.3	0.42	59.0	95.4
2	8	1+ P	0.49	44.2	97.9	0.42	63.8	95.3
	9	2+ P	0.40	30.2	99.0	0.29	71.5	94.3
	10	1+ Rx	0.51	68.6	93.4	0.62	47.1	97.2
	11	1+ H or 1+ P	0.49	45.0	97.8	0.43	63.7	95.4
	12	1+ H or 2+ P	0.40	31.2	98.9	0.30	71.2	94.4
	13	1+ H or 1+ P or 1+ Rx	0.51	71.6	92.7	0.64	45.8	97.4
	14	1+ H or 2+ P or 2+ Rx	0.53	55.8	96.3	0.52	56.7	96.2
3	15	1+ P	0.52	53.1	96.8	0.50	58.4	96.0
	16	2+ P	0.46	37.9	98.5	0.36	68.8	94.9
	17	1+ Rx	0.50	73.8	92.0	0.66	44.1	97.6
	18	1+ H or 1+ P	0.52	53.6	96.7	0.50	58.2	96.1
	19	1+ H or 2+ P	0.46	38.7	98.5	0.37	68.3	94.9
	20	1+ H or 1+ P or 1+ Rx	0.49	76.9	90.9	0.68	42.0	97.9
	21	1+ H or 2+ P or 2+ Rx	0.55	64.9	95.3	0.60	54.3	96.9
5	22	1+ P	0.54	63.7	95.1	0.59	52.9	96.8
	23	2+ P	0.53	50.5	97.5	0.48	63.7	95.8
	24	1+ Rx	0.49	80.9	89.9	0.71	40.6	98.2
	25	1+ H or 1+ P	0.54	63.7	95.1	0.59	52.6	96.8
	26	1+ H or 2+ P	0.53	50.7	97.4	0.48	63.0	95.8
	27	1+ H or 1+ P or 1+ Rx	0.46	83.8	88.3	0.72	38.0	98.5
	28	1+ H or 2+ P or 2+ Rx	0.56	74.6	93.8	0.68	50.9	97.7

Appendix Table F.2: Estimates of agreement, sensitivity, specificity, and predictive values for asthma algorithms, 50+ years (COPD and emphysema removed)

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ P	0.39	32.0	98.5	0.30	60.0	95.4
	2	2+ P	0.29	20.1	99.4	0.19	69.4	94.7
	3	1+ Rx	0.44	65.7	92.8	0.58	38.9	97.5
	4	1+ H or 1+ P	0.39	32.5	98.4	0.31	59.1	95.4
	5	1+ H or 2+ P	0.30	20.7	99.3	0.20	67.3	94.7
	6	1+ H or 1+ P or 1+ Rx	0.44	66.9	92.4	0.59	38.3	97.5
	7	1+ H or 2+ P or 2+ Rx	0.46	56.8	95.0	0.52	44.4	96.9
2	8	1+ P	0.47	45.0	97.6	0.43	56.3	96.2
	9	2+ P	0.38	29.6	98.8	0.28	62.5	95.2
	10	1+ Rx	0.42	74.0	90.1	0.64	34.4	98.0
	11	1+ H or 1+ P	0.48	46.7	97.5	0.44	56.4	96.3
	12	1+ H or 2+ P	0.39	31.4	98.7	0.30	62.4	95.3
	13	1+ H or 1+ P or 1+ Rx	0.41	75.1	89.5	0.65	33.5	98.1
	14	1+ H or 2+ P or 2+ Rx	0.45	63.9	93.5	0.57	40.8	97.4
3	15	1+ P	0.48	53.3	96.3	0.50	50.6	96.7
	16	2+ P	0.42	36.1	98.2	0.34	58.7	95.6
	17	1+ Rx	0.40	76.9	88.6	0.66	32.2	98.2
	18	1+ H or 1+ P	0.49	54.4	96.3	0.51	50.5	96.8
	19	1+ H or 2+ P	0.42	37.3	98.1	0.35	57.8	95.7
	20	1+ H or 1+ P or 1+ Rx	0.39	79.3	87.7	0.67	31.1	98.4
	21	1+ H or 2+ P or 2+ Rx	0.44	68.6	92.2	0.61	38.3	97.7
5	22	1+ P	0.50	60.9	95.2	0.56	47.2	97.2
	23	2+ P	0.46	44.4	97.6	0.42	56.0	96.2
	24	1+ Rx	0.38	81.1	86.9	0.68	30.3	98.5
	25	1+ H or 1+ P	0.49	60.9	95.1	0.56	46.8	97.2
	26	1+ H or 2+ P	0.46	45.0	97.4	0.42	54.7	96.2
	27	1+ H or 1+ P or 1+ Rx	0.37	84.0	85.7	0.70	29.2	98.7
	28	1+ H or 2+ P or 2+ Rx	0.45	75.7	91.2	0.67	37.5	98.2

APPENDIX G: ADDITIONAL VALIDATION RESULTS FOR CORONARY HEART DISEASE ALGORITHMS

Appendix Table G.1: Estimates of agreement, sensitivity, specificity, and predictive values for additional coronary heart disease algorithms

# Years		Algorithm	κ	Sens. (%)	Spec. (%)	Youden	PPV (%)	NPV (%)
1	1	1+ H or 1+ P or 1+ Rx	0.31	80.3	82.6	0.63	24.8	98.3
	2	1+ H or 1+ P or 2+ Rx	0.32	78.7	84.0	0.63	25.9	98.3
2	3	1+ H or 1+ P or 1+ Rx	0.30	83.8	81.3	0.65	24.2	98.6
	4	1+ H or 1+ P or 2+ Rx	0.32	82.2	82.9	0.65	25.4	98.5
3	5	1+ H or 1+ P or 1+ Rx	0.30	87.1	80.2	0.67	23.9	98.9
	6	1+ H or 1+ P or 2+ Rx	0.32	85.2	82.0	0.67	25.2	98.7

Note: H = Hospital separation; P = Physician claim; Rx = Prescription drug record

Recent MCHP Publications

2006

Application of Patient Safety Indicators in Manitoba: A First Look by Sharon Bruce, Heather Prior, Alan Katz, Mark Taylor, Steven Latosinsky, Patricia Martens, Carolyn De Coster, Marni Brownell, Ruth-Ann Soodeen and Carmen Steinbach

2005

Sex Differences in Health Status, Health Care Use, and Quality of Care: A Population-Based Analysis for Manioba's Regional Health Authorities by Randy Fransoo, Patricia Martens, *The Need to Know* Team (funded through CIHR), Elaine Burland, Heather Prior, Charles Burchill, Dan Chateau, and Randy Walld.

Health and Health Care Use Among Older Adults: Using Population-Based Information Systems to Inform Policy in Manitoba, Canadian Journal on Aging, Volume 24, Supplement 1, 2005

High-Cost Users of Pharmaceuticals: Who Are They? by Anita Kozyrskyj, Lisa Lix, Matthew Dahl and Ruth-Ann Soodeen

Primary Prevention: An Examination of Data Capabilities in Manitoba, by Lisa Lix, Greg Finlayson, Marina Yogendran, Ruth Bond, Jennifer Bodnarchuk, and Ruth-Ann Soodeen

Aboriginal Health Research and Policy: First Nations-University Collaboration in Manitoba, Canadian Journal of Public Health, Volume 96, Supplement 1, January/February 2005

2004

Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study, by Patricia Martens, Randy Fransoo, Nancy McKeen, *The Need To Know Team* (funded through CIHR), Elaine Burland, Laurel Jebamani, Charles Burchill, Carolyn De Coster, Okechukwu Ekuma, Heather Prior, Dan Chateau, Renée Robinson, and Colleen Metge

Diagnostic Imaging Data in Manitoba, Assessment and Applications, by Greg Finlayson, Bill Leslie and Leonard MacWilliam

How do Educational Outcomes Vary With Socioeconomic Status? Key Findings from the Manitoba Child Health Atlas 2004, by Marni Brownell, Noralou Roos, Randy Fransoo, Anne Guèvremont, Leonard MacWilliam, Shelley Derksen, Natalia Dik, Bogdan Bogdanovic, and Monica Sirski

Using Administrative Data to Develop Indicators of Quality in Family Practice, by Alan Katz, Carolyn De Coster, Bogdan Bogdanovic, Ruth-Ann Soodeen, and Dan Chateau

Patterns of Health Care Use and Cost at the End of Life, by Verena Menec, Lisa Lix, Carmen Steinbach, Okechukwu Ekuma, Monica Sirski, Matt Dahl, and Ruth-Ann Soodeen

2003

Pharmaceuticals: Therapeutic Interchange and Pricing, by Steve Morgan, Anita Kozyrskyj, Colleen Metge, Noralou Roos, and Matt Dahl

Pharmaceuticals: Focussing on Appropriate Utilization, by Colleen Metge, Anita Kozyrskyj, Matt Dahl, Marina Yogendran, and Noralou Roos

Supply, Availability and Use of Family Physicians in Winnipeg, by Diane Watson, Bogdan Bogdanovic, Petra Heppner, Alan Katz, Robert Reid, and Noralou Roos

Manitoba RHA Indicators Atlas: Population-Based Comparisons of Health and Health Care Use, by Patricia J Martens, Randy Fransoo, *The Need to Know Team*, Elaine Burland, Laurel Jebamani, Charles Burchill, and others.

Why is the Health Status of Some Manitobans Not Improving? The Widening Gap in the Health Status of Manitobans, by Marni Brownell, Lisa Lix, Okechukwu Ekuma, Shelley Derksen, Suzanne De Haney, and others.

Discharge Outcomes for Long-Stay Patients in Winnipeg Acute Care Hospitals, by Anita Kozyrskyj, Charlyn Black, Elaine Dunn, Carmen Steinbach, and Dan Chateau

Key Events and Dates in the Manitoba Health Care System, 1990 to 2003, compiled by Fred Toll

2002

Improving Children's Health: How Population-Based Research Can Inform Policy - The Manitoba Experience, Canadian Journal of Public Health, Volume 93, Supplement 2, November/December 2002

Monitoring the Acute Care Sector: Key Measures and Trends, Healthcare Management Forum Supplement, Winter 2002

Estimating Personal Care Home Bed Requirements, by Norman Frohlich, Carolyn De Coster, and Natalia Dik

The Health and Health Care Use of Manitoba's Seniors: Have They Changed Over Time? by Verena Menec, Leonard MacWilliam, Ruth-Ann Soodeen, and Lori Mitchell

Profile of Medical Patients Who Were Assessed as Requiring Observation-Level Services at Winnipeg Acute Care Hospitals in 1998/99, by Sharon Bruce, Charlyn Black, and Charles Burchill

Projecting Hospital Bed Needs for 2020, by David Stewart, and Robert Tate, Greg Finlayson, Leonard McWilliam, and Noralou Roos

Copies of MCHP publications are available for download free of charge at **http://www.umanitoba.ca/centres/mchp/reports.htm**
Hard copies of our reports are available, free of charge, by contacting us at:

Manitoba Centre for Health Policy
University of Manitoba
4th Floor, Room 408
727 McDermot Avenue
Winnipeg, Manitoba, Canada R3E 3P5
Email: reports@cpe.umanitoba.ca
Phone: 204-789-3819 **Fax**: 204-789-3910