Use and Validation of Administrative Data for Suicide Research

By

Jason R Randall

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirement of the degree of

Doctor of Philosophy

Department of Community Health Sciences

Rady Faculty of Health Sciences

University of Manitoba

Winnipeg

ABSTRACT

Health system administrative databases are valuable sources of data about health system use. Most of these databases now extend multiple decades and therefore can be used to perform birth cohort studies on children, youth and young adults. The goal of this thesis was to use the administrative data available in Manitoba, Canada to assess the incidence of mental health disorders, suicide attempts, and deaths, as well as to assess the validity of detecting suicide attempts with these data.

Chapter I examined the incidence of diagnosis with anxiety, mood and adjustment, personality, schizophrenia, and substance use disorders in a cohort of individuals born in Manitoba who were living in the city of Winnipeg on their tenth birthday. The estimates were higher than those provided by studies relying on recall of diagnosis. The results supported the idea that recall for diagnosis with mental disorders might be a poor method of determining lifetime history of these illnesses.

Chapter II examined the occurrence of suicide attempts/deaths in the cohort from Chapter I. This chapter showed that the suicide attempts were fairly common in the sample. It also noted that individuals with personality and schizophrenia disorders had the highest occurrence of attempts/deaths and that the occurrence of behaviours was particularly high after the first instance of diagnosis with one of the five disorders examined in the previous chapter.

Chapter III assessed the validity of using International Classification of Diseases (ICD) codes for detecting suicide attempts from hospital discharge abstracts. We found that these codes have good specificity and positive predictive validity, but miss most of the suicide attempts that are admitted.

Chapter IV used latent class analysis to examine whether the current data in the hospital discharge abstracts, medical claims and emergency department information system (EDIS) can accurately identify individuals presenting with self-harm to the emergency department. This study found that these data sources are currently insufficient at identifying these individuals.

Overall, this thesis used administrative data to perform an epidemiological cohort study on mental illness and suicide attempts, but also highlighted some of the limitations of this method of epidemiological study.

ACKNOWLEDGMENTS

I want to acknowledge the guidance and support of my PhD advisors; Dr. Bolton, Dr. Roos, and Dr. Martens. I would like to thank the other members of my thesis committee, Dr. Lix and Dr. Katz. I also want to thank the staff of the Manitoba Centre for Health Policy for their support during the process of receiving data access and their advice around the use of administrative data, including Carole Taylor, Dan Chateau, Charles Burchill, and Elaine Burland. I would also like to thank Aynslie, Shivoan, Taru, and Elizabeth for uncounted instances of advice and support during classes and my thesis.

During my study I received support from the University of Manitoba Graduate Extension of Tri-council funding (GETS) program. I would also like to acknowledgment the support of the Manitoba Centre for Health policy through the Evelyn Shapiro Award.

I would also like to acknowledge the Manitoba Centre for Health Policy (MCHP) for use of data contained in the Population Research Data Repository under project # 2015-017 (HIPC # 2014/2015 - 41). The results and conclusions are those of the authors and no official endorsement by MCHP, Manitoba Health, or other data providers is intended or should be inferred. Data used in this thesis were provided by Manitoba Health, the Winnipeg Regional Health Authority (WRHA), and Manitoba Vital Statistics Agency.

DEDICATION

I would like to dedicate this thesis to my parents, Brian and Sharon Randall, who supported me and my academic interests and ensured I was able to pursue them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBERVIATIONS

BIC:	Bayesian information criterion
C-CASA:	Columbia Classification Algorithm of Suicide Assessment
CI:	confidence interval
DF:	degrees of freedom
DSM:	Diagnostic and Statistical Manual of Mental Disorders
ED:	emergency department
EDIS:	emergency department information system
ICD-9-CM:	International Classification of Diseases, 9th revision – clinical modification
ICD-10-CA:	International Classification of Diseases, 10th revision, Canada
LCA:	latent class analysis
NPV:	negative predictive value
PPV:	positive predictive value
SA:	suicide attempt
SAFE:	suicide assessment form in the emergency department
SD:	standard deviation
SH:	self-harm

CHAPTER 1: INTRODUCTION

1.1 Background and Overview of the thesis

1.1.1 Background

Suicide and self-harm is a leading cause of mortality and morbidity for adolescents and young adults.[1,2] Determining risk factors for suicide and suicidal behaviours has been a major portion of research on suicide.[1–4] Mental illnesses have been found in over 90% of suicide deaths.[5,6] Mental illness often manifest in adolescence and the affected individuals often suffer for decades and experience high risk of mortality.[7–10] There has been some dispute over the use of retrospective methods for research on mental illness.[11] Undercounting is a potential issue with these methods.[11–13]

Administrative data derived from government provided services have been used to study mental illness and suicidal behaviours. In publically-funded systems, such as those in Manitoba, these data can provide the framework for comprehensive population-based analysis. However, there are limitations to using these data for research purposes. These data were not collected for research purposes and conditions and events of interests are not always explicitly coded in these data. Determining the occurrence of physical conditions using these data has been shown to have acceptable validity,[14–17] but there have been few projects examining the validity of detecting suicide attempts using these codes.[18] The studies that have been done have generally been small and have not provided exhaustive statistics for the validity of the codes.

This thesis will focus on two main projects, which are further subdivided into two sub-projects. These projects will focus on using administrative health records for mental health research. The first project will validate the coding of self-harm in administrative records and in a

clinically collected dataset used for research. The second project will estimate the cumulative incidence of treated mental illness and suicide attempts in a recent birth cohort.

1.1.2 Definitions and terminology

There has been some disagreement on terminology in the suicide research literature and various methods of classifying suicidal and self-harming behaviours have been suggested.[19–23] This thesis will use clear and specific definitions. Proper definitions in this area need to differentiate the *lethality* and *intent* of the behaviours. Behaviours that involved an *intent to die* will be discussed using the terms 'suicide,' 'suicide attempt,' 'death(s) from suicide,' and 'suicidal behaviour(s).' 'Suicide attempts' specify non-lethal actions and 'death(s) from suicide' specifies lethal actions, other terms will refer to both non-lethal and lethal behaviours, together.

The term 'self-harm' is often used to mean different outcomes by researchers. Since the term 'self-harm' is non-specific, it will be used to refer to *any* self-injury without respect to lethality or intention. 'Self-harm' will be used as the general catch-all phrase to denote the full spectrum of self-directed violence behaviours. For non-lethal behaviours without suicidal intent the term 'non-suicidal self-injury' (NSSI) will be used.

The term 'serious suicide attempt' is used to refer to attempts that result in death, or fail to result in death through chance (i.e. surviving a jump from a fatal height). 'Medically-treated' should be interpreted to mean that a particular incident of self-harm resulted in medical treatment, without clear indication of the severity of the injury/attempt. Due to limitations inherent to the administrative data available the severity of attempts is difficult to ascertain and will not be a focus of this thesis. However, the self-harm examined in this thesis will be entirely 'medically-treated' incidents of self-harm. A small part of this thesis will explore whether

current methods identify suicide attempts specifically and distinctly from other types of self-harm.

1.1.3 Overview of the thesis

This thesis will be composed of four papers subdivided into two sections. Each section will contain two related papers. The first section will contain papers detailing the epidemiology of several mental disorders and admission to inpatient units for suicide attempts or other forms of self-harm for a birth cohort between the ages of 10 and 25 years old. The second section will contain papers assessing the validity of identifying instances of suicide attempts that result in inpatient treatment using the diagnostic codes available in the hospital discharge records. The conclusion will summarize the results and their possible implication for mental health epidemiology, and discuss the potential strengths and limitations of using administrative data sources for long-term epidemiological research on suicide attempts and deaths.

1.1.4 Data sources

This thesis utilizes data from the Population Research Data Repository (the Repository) housed at the Manitoba Centre for Health Policy. The Repository contains the administrative data records for the publically administered health services in the province of Manitoba, Canada. Individuals within the data are identified and linkable using their Provincial Health Insurance Number. A de-identified version of this number is used to link individuals in the Repository to help protect the anonymity of individuals in the data. The specific datasets used included the medical claims records which contain the records of physician treatment in the fee-for-service single-payer health insurance program of the province of Manitoba. Also included were the hospital discharge abstracts dataset which contained data on inpatient treatment in the province. Vital statistics records were included to identify deaths. The Provincial Health Insurance

Registry data set, which includes coverage dates, birth dates, sex, and postal codes, was also used to identify individuals that matched the inclusion criteria for the papers in the epidemiology section of the thesis. The Emergency Department Information System (EDIS) contains records for individual presenting to emergency department and records triage information for these individuals. Lastly, the Suicide Assessment form in the Emergency Department (SAFE), a non-administrative clinical data set collected from routine psychiatric assessment in the two largest hospitals in Winnipeg, will be used in the validation section.

1.2 Epidemiology of mental health and suicidal behaviour

1.2.1 Emergence of psychiatric conditions in adolescence

Prior studies have shown that adolescence and early adulthood is a critical time of risk for new disorders.[24–27] These disorders not only emerge in youth, but frequently cause a significant burden of morbidity.[7–9] Studies on disability-adjusted life years have estimated that almost half of all disability in people aged 20-24 is due to mental and substance use disorders.[7] In the U.S., mental disorders are the third leading cause of hospital stays among youth.[28]

The National Comorbidity Survey Replication Adolescent Supplement estimated the 12-month prevalence of any DSM-IV disorder at 40.3% among 12-17 year olds and estimated the 30-day prevalence at 23.4%.[24] The mental illnesses of interest in this thesis are five groups of disorders; mood and adjustment disorders, anxiety, schizophrenia, personality disorders, substance use disorder. The 12-month prevalence of mood disorders in the National Comorbidity Survey was estimated at 10%.[24] This was considerably lower than the estimated prevalence for anxiety disorders during the same period (24.9%). Substance use disorder was almost as common as mood disorder with 8.3% experience a disorder within the last 12 months. Schizophrenia is widely considered to have a lifetime incidence of approximately 1%, but some

recent studies suggest the real incidence may be lower.[29,30] However, these recent studies may be affected by poor methodology.[29]

A prospective administrative data-based study in Denmark estimated the lifetime treatment incidence of disorders at different ages.[31] They estimated an incidence of 3-4% for mood disorders and 6-7% for anxiety disorders at age 25. Personality disorders and schizophrenias were estimated to have occurred in 2% and 1.6% percent of individuals by age 25. The pattern of new incidences in Pedersen et al.[31] showed an increase in incidence around age 14 which peaked around the age of 20. The results for the Pedersen et al. and the National Comorbidity Survey illustrate the sharp difference in the prevalence of treated disorders and the much higher rate from surveys.

Disorders have differential incidence and prevalence between the sexes.[24] Young females have higher prevalence of mood, and anxiety disorders.[24] Other disorders are more common among young males, including substance use disorders.[24] Schizophrenia is considered more commonly diagnosed among males than females, but a systematic review of prevalence studies did not support this.[30] Some specific personality disorders are more commonly diagnosed among women, whereas other personality disorders are more common among men.[32]

Some research has suggested that mental illnesses are becoming more common in this age group in recent years.[33–39] However, not all research has supported this claim – including longitudinal research using Canadian data.[40–44] More research is needed to determine whether there is a trend, whether the trend if not consistent across western countries, and what the underlying causes of the trend could be.

1.2.2 Emergence of self-harm in adolescence

Adolescence presents an opportunity to examine the earliest emergence of self-harm. Unfortunately, much of the research that has been done in this area has focused on western, developed countries and information is scarce or inaccurate for Africa, the Middle East, and India.[4] Until puberty these behaviours are uncommon.[45] Little research focuses on self-harm before age 14.[1,46] Previous research has detected a range of prevalence rates for adolescents. One-year prevalence rates for developed countries have been estimated to be in the range of 3% to 10%.[46] Lifetime prevalence rates for adolescents has been estimated to be between 1.9-19%.[1,46–49] This wide range of prevalence estimates suggests that larger and better quality studies are needed to estimate these rates.

There are many risk factors for suicide. Distal risk factors include genetics, personality, fetal and perinatal issues, early traumatic events, and neurobiological disturbances.[4] Proximal risk factors include psychiatric disorder, physical disorders (e.g., cancer), psychosocial crises, access to lethal means of suicide, and exposure to models of suicidal behaviour.[4,50] Much research on suicide and self-harm has been directed toward the causal effect of mental illness on these behaviours.[1,51,52] Approximately 90% of deaths from suicide had psychiatric conditions based on a systematic review of psychological autopsies.[53] This rate may be lower in among adolescent suicide deaths.[5,54–57]

Non-fatal self-harm has generally been found to be more common among females,[58] but males are more likely to die from suicide.[58] The male-to-female ratio is higher in western countries than Asian countries, and in China more women than men die from suicide.[4] In western countries, including Canada, approximately 80% of suicide deaths have been among males.[58–60]

It is possible that rates of self-harm have been increasing recently, including death from suicide, after several decades of apparent decline.[1,4,5,58,61,62] The decline in suicide death in Canada during the 1990's was driven specifically by a decrease in the rate of suicide among males.[5] The most recent year of data (2012) from Statistics Canada estimated the rate of death by suicide at 1.8 per 100,000 for those aged 10 to 14, 10.2 per 100,000 for those aged 15 to 19, and 12.1 per 100,000 for those aged 20 to 24.[60]

This thesis assesses the cumulative treatment incidence of suicidal behaviour for those with 5 groups of disorders; mood and adjustment disorders, anxiety, schizophrenia, personality disorder, and substance use disorder. There is a substantial body of research linking these groups to suicidal behaviours.[1,5,54,63–71] Therefore, knowing the cumulative incidence of these disorders can not only aid in the treatment of these disorders but also inform prevention and treatment of self-harm.

1.2.3 Potential issues with prior research

A paper by Scott Patten, and echoed by other researchers, illustrates the fundamental concern with lifetime incidence results derived from retrospective methods.[11,27] The specific concern discussed was that the lifetime incidence of mental health disorders derived from retrospective studies was similar across the ages. The expected result was that lifetime incidence would continue to increase throughout life. Two main causes were identified by Patten that could explain the flat trend in lifetime incidence; recall failure and excess mortality in those with the condition. Patten's paper used equations proposed in a paper by GB Hill for estimating incidence from prevalence in order to estimate the rate of recall failure per year that could explain a flat trend in lifetime incidence by age.[72] The mathematical modelling showed that the annual recall failure needed to explain the flat incidence was less than 1.5% per year, and this number

decreased with age. This number was also consistently lower for men, drastically so for the model with high relative mortality. This indicates that estimates for men may be more biased than women. There have been studies that illustrate that recall failure is common for conditions like depression.[73,74] Therefore the relatively low rate of recall failure required by Patten's estimation is highly plausible.

There have been several notable long-term birth cohorts that have examined incidence of mental disorders. However, there are some notable limitations to these disorders which may bias their results and interpretation. One issue is that they frequently experience significant loss to follow-up.[75] Since individuals with mental disorders or histories of self-harm are at increased risk of death, or otherwise being lost to follow-up, then this could cause a significant undercounting. Another issue is that the follow-up periods are often uneven, or span one or more years. These long periods between assessment risk increasing recall failure for disorders and self-harm. The previously discussed paper by Pedersen et al., while not strictly speaking a birth cohort, provides good prospective data on the population of Denmark.[31] The main drawback to this study is that it focuses on a strict definition of disorders and likely only captures disorders of higher severity. More research is needed that uses prospective data methods in order to estimate the true incidence rate. Administrative data are potentially a very strong source of data for this purpose since they are both prospective and do not suffer from the issue of loss to follow up as severely as traditional survey-based birth cohort studies. Using administrative data to derive an estimate of cumulative incidence that can be compared to self-reported incidence is a potential method for examining and addressing the issues just discussed.

1.2.4 Project papers

Paper I examines the cumulative incidence of several major mental illness groups in a birth cohort spanning the fiscal years of 1979 to 1992. These groups are; mood and adjustment disorders, anxiety, psychosis, substance use, and personality disorders. These groups are common and/or serious mental illnesses that are associated with a considerable portion of self-harm and death from suicide. Cumulative incidence is determined by using the cumulative incidence function, sometimes referred to as the cumulative incidence competing risk method.[76,77] This method adjusts for competing risks, in this case death, to provide a more accurate estimate of the cumulative incidence of these disorders. Incidences for the disorders at age 15, 20, and 25 will be produced. Further incidences will be produced based on sex and by grouping individuals based on whether they were born in the early or late half of the birth cohort. Paper II will examine the cumulative incidence of suicide attempts using the same cohort. Incidence will be shown both in the whole population, and by the mental disorders examined in Paper I. The analysis will also be performed grouped by sex and early/late birth periods.

1.3 Validation issues in administrative research

In order to utilize administrative data for research purposes; algorithms need to be derived to identify and classify patients. Administrative data do not clearly delineate whether a person has a specific condition. During treatment, physicians assign diagnostic codes to patients. These codes are recorded in the administrative data and used to identify people with the conditions of interest. Someone's condition may be improperly coded according to the ICD-9/10 framework. This thesis is interested in identifying self-harm presentations to the emergency department.

1.3.1 Signal detection theory and validation research

Signal detection theory is the foundation of validation research. Signal detection theory describes how to make decision during uncertainty.[78] Administrative data contain a degree of 'noise' that obscures the true occurrence of diseases and events. What researchers are interested in is the 'signal' within the data. Methods of interpreting the data are needed in order to separate the signal from the noise. This paper will refer to these methods as criterion or algorithms. Several methods have developed to assess how effectively algorithms and criterion separate the signal from the noise and detect a condition. The main statistics in use are sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).[79,80] Sensitivity is the percent of positive individuals that are detected by the screen. Specificity is the percent of negative individuals that are correctly identified as negatives by the screen. PPV is the percent with the condition among those screened as positive. NPV is the percent of individuals that are negative among those screened as negative. Another method is to use the Kappa statistic.[81] This method compares the observed accuracy with the expected accuracy that would occur through random chance. A kappa of 0.0 indicates random chance at detection the condition, whereas a kappa of 1.0 indicates perfect detection of cases and non-cases. Kappa values of less than 0 are also possible if disagreement occurs above chance levels.

There has been some research on validation of mental health coding in hospital discharge abstracts and emergency department data that used chart reviews to examine agreement between the original coders and expert reviewers.[82] They found good agreement for main diagnosis and major comorbidities, but suggested that there is potential for miscoding of comorbid conditions. The emergency department report found that although the main presentation issue was identified with high agreement, there was some difference in the specific codes used to describe the main

issue due to uncertainty over the proper coding of some issues.[83] Other research specifically on case identification for mental health disorders in administrative data has found the accuracy of these data sources to be acceptable, but have noted a trade-off between requiring multiple diagnosis codes to be assigned a case (which produces good specificity, but lower sensitivity) or requiring as few as one code (which produces good sensitivity, but lower specificity).[84–86] Research validating suicide attempt and self-harm harm related coding has been conducted that has shown that these codes under-identify true cases of suicide attempt. One specific issue is that apparent suicide attempts and self-harm are being categorized as injuries of undetermined intent. However, a major limitation of this prior research has been small sample sizes and a lack of a good gold standard to establish the validity of these codes.[18,87,88]

1.3.2 Validation research when there is no gold standard

There are many possible ways of deriving a criteria or algorithm.[79] The main issue with validating administrative data is that there is usually no gold standard available. Methods exist for validation when there is no gold standard. One method is to use Latent Class Analysis (LCA).[79] The main advantages of the LCA method is that it is a flexible, data driven method of validation. It also allows validation without having to perform a chart review with expert coders to re-abstract the admission records, which is a costly and time consuming procedure. Also, if poor documentation is a main source of poor accuracy then re-examining the charts may overestimate the validity of codes if the reviewers consistently make the same errors as the original medical coders.

LCA validation, in contrast, can be performed with data already present in the administrative records. LCA assumes that variables that are derived from data, in this case administrative records, are indicators of membership in a unmeasured class – a latent class.[89] For

example, if a researcher is interested in alcohol and drug use classes then they can perform a latent class analysis using indicators of frequency of use for drugs and alcohol. The LCA would then group people in classes based on the pattern of their drug and alcohol use, allowing the researcher to identify the major groups of substance users and obtain estimates for the proportion in each of the classes. For this study, potential indicators for the occurrence of a suicide attempt presenting to the emergency department can be entered into a LCA model to attempt to identify the members of one or more suicide attempter latent classes. The effectiveness of these indicators at identifying suicide attempt presentations can be assessed using this latent class. The proportion of class members that are identified by the indicator variables is the sensitivity of the specific indicator. The specificity, NPV and PPV of the indicators can also be assessed.

The latent class of suicide attempt presentations can also be used as a pseudo-gold standard.[79] This pseudo-gold standard can then be used to assess the validity of tests and variables at detecting cases. LCA methods are frequently used to validate methods of detecting infectious diseases.[79] In that scenario, the interest is in determining the effectiveness of specific tests in isolation. However, once the latent class of suicide attempters has been identified then methods such as regression can be used to derive mathematical models predicting membership in the suicide attempters' latent class. These models can be used to predict who is a case and then the validity statistics of the models can be assessed.

1.3.3 Project papers

Paper I examines the validity of administrative data at detecting suicide attempts using variations of current coding methods for suicide attempts in hospital data.[18] It produced sensitivity, specificity, PPV, NPV, and Kappa statistics for ICD-10-CA codes to assess the validity of identifying patients admitted with self-harm or a suicide attempt. This paper adds to previous

results by using a good gold standard (clinician assessment of suicidality during presentation) combined with a sample large enough to make accurate estimates of the validation statistics.

Paper II aims to determine the comprehensiveness of the available administrative datasets assessment of suicide attempters. Specifically, it uses LCA to determine the latent class of suicide attempters using the C-CASA classification from a linked clinical assessment form, as well as other injury and mental health indicators derived from administrative data. This will allow an estimation of the number of suicide attempters and self-harmers that are not identified in the SAFE data. It will also provide the framework, based on the use of the suicide attempter latent class as a pseudo-gold standard, to derive better methods of detecting suicide attempts and self-harm presentation to the emergency department. It will also provide estimates for the proportion of self-harmers with each of the indicator variables. No prior research using LCA to validate administrative data sources was found prior to this study. Therefore this study is a novel attempt to examine how to identify cases in imperfect administrative data using LCA.

1.4 References

1 Hawton K, Saunders KE a, O'Connor RC. Self-harm and suicide in adolescents. *Lancet* 2012; **379**: 2373–82.

2 Johnson GR, Krug EG, Potter LB. Suicide among adolescents and young adults: a cross-national comparison of 34 countries. *Suicide Life Threat Behav* 2000; **30**: 74–82.

3 Johnson JG, Cohen P, Gould MS, Kasen S, Brown J, Brook JS. Childhood adversities, interpersonal difficulties, and risk for suicide attempts during late adolescence and early adulthood. *Arch Gen Psychiatry* 2002; **59**: 741–9.

4 Hawton K, van Heeringen K. Suicide. *Lancet* 2009; **373**: 1372–81.

5 Steele MM, Doey T. Suicidal behaviour in children and adolescents. part 1: etiology and risk factors. *Can J Psychiatry* 2007; **52**: 21S–33S.

6 American Academy of Child and Adolescent Psychiatry. Practice parameter for the assessment and treatment of children and adolescents with suicidal behavior. American Academy of Child and Adolescent Psychiatry. *J Am Acad Child Adolesc Psychiatry* 2001; **40**: 24S–51S.

7 Whiteford H, Ferrari A, Degenhardt L. Global Burden Of Disease Studies: Implications For Mental And Substance Use Disorders. *Health Aff* 2016; **35**: 1114–20.

8 Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, *et al.* Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2197–223.

9 Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: An analysis from the global burden of disease study 2010. *PLoS One* 2015; **10**: 1–14.

10 Eaton WW, Roth KB, Bruce M, Cottler L, Wu L, Nestadt G, *et al.* The relationship of mental and behavioral disorders to all-cause mortality in a 27-year follow-up of 4 epidemiologic catchment area samples. *Am J Epidemiol* 2013; **178**: 1366–77.

11 Patten SB. Recall bias and major depression lifetime prevalence. *Soc Psychiatry Psychiatr Epidemiol* 2003; **38**: 290–6.

12 Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G, *et al.* How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychol Med* 2010; **40**: 899–909.

13 Hamdi NR, Iacono WG. Lifetime prevalence and co-morbidity of externalizing disorders

and depression in prospective assessment. *Psychol Med* 2014; **44**: 315–24.

14 Chan KS, Fowles JB, Weiner JP. Review: electronic health records and the reliability and validity of quality measures: a review of the literature. *Med Care Res Rev* 2010; **67**: 503–27.

15 Baker DW, Persell SD, Thompson JA, Soman NS, Burgner KM, Liss D, *et al.* Automated review of electronic health records to assess quality of care for outpatients with heart failure. *Ann Intern Med* 2007; **146**: 270–7.

16 Tang PC, Ralston M, Arrigotti MF, Qureshi L, Graham J. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *J Am Med Inform Assoc* 2007; **14**: 10–5.

17 Faulconer ER, de Lusignan S. An eight-step method for assessing diagnostic data quality in practice: chronic obstructive pulmonary disease as an exemplar. *Inform Prim Care* 2004; **12**: 243–54.

18 Walkup JT, Townsend L, Crystal S, Olfson M. A systematic review of validated methods for identifying suicide or suicidal ideation using administrative or claims data. *Pharmacoepidemiol Drug Saf* 2012; **21 Suppl 1**: 174–82.

19 Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. *Am J Psychiatry* 2007; **164**: 1035–43.

20 Kapur N, Cooper J, O'Connor RC, Hawton K. Non-suicidal self-injury v. attempted suicide: new diagnosis or false dichotomy? *Br J Psychiatry* 2013; **202**: 326–8.

21 Skegg K. Self-harm. *Lancet* 2005; **366**: 1471–83.

22 Platt S, Bille-Brahe U, Kerkhof A, Schmidtke A, Bjerke T, Crepet P, *et al.* Parasuicide in Europe: the WHO/EURO multicentre study on parasuicide. I. Introduction and preliminary analysis for 1989. *Acta Psychiatr Scand* 1992; **85**: 97–104.

23 Posner K, Brown GK, Stanley B, Brent DA, Yershova K V., Oquendo MA, *et al.* The Columbia–Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings From Three Multisite Studies With Adolescents and Adults. *Am J Psychiatry* 2011; **168**: 1266–77.

24 Kessler RC, Avenevoli S, Costello EJ, Georgiades K, Green JG, Gruber MJ, *et al.* Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Arch Gen Psychiatry* 2012; **69**: 372–80.

25 Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. Prior juvenile

diagnoses in adults with mental disorder. *Arch Gen Psychiatry* 2003; **60**: 709–17.

26 Costello EJ, Egger H, Angold A. 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *J Am Acad Child Adolesc Psychiatry* 2005; **44**: 972–86.

27 Costello EJ, Foley DL, Angold A. 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: II. Developmental epidemiology. *J Am Acad Child Adolesc Psychiatry* 2006; **45**: 8–25.

28 Mulye TP, Park MJ, Nelson CD, Adams SH, Irwin CE, Brindis CD. Trends in adolescent and young adult health in the United States. *J Adolesc Health* 2009; **45**: 8–24.

29 Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, *et al.* Lifetime Prevalence of Psychotic and Bipolar I Disorders in a General Population. *Arch Gen Psychiatry* 2007; **64**: 19.

30 Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med* 2005; **2**: e141.

31 Pedersen CB, Mors O, Bertelsen A, Waltoft BL, Agerbo E, McGrath JJ, *et al.* A Comprehensive Nationwide Study of the Incidence Rate and Lifetime Risk for Treated Mental Disorders. *JAMA Psychiatry* 2014; **71**: 573.

32 Grant BF, Hasin DS, Stinson FS, Dawson D a, Chou SP, Ruan WJ, *et al.* Prevalence, correlates, and disability of personality disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. *J Clin Psychiatry* 2004; **65**: 948–58.

33 Compton WM, Conway KP, Stinson FS, Grant BF. Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991-1992 and 2001-2002. *Am J Psychiatry* 2006; **163**: 2141–7.

34 Goldney RD, Eckert K a, Hawthorne G, Taylor AW. Changes in the prevalence of major depression in an Australian community sample between 1998 and 2008. *Aust N Z J Psychiatry* 2010; **44**: 901–10.

35 Achenbach TM, Dumenci L, Rescorla LA. Are American children's problems still getting worse? A 23-year comparison. *J Abnorm Child Psychol* 2003; **31**: 1–11.

36 Collishaw S, Maughan B, Goodman R, Pickles A. Time trends in adolescent mental health. *J Child Psychol Psychiatry* 2004; **45**: 1350–62.

37 Lewinsohn PM, Rohde P, Seeley JR, Fischer SA. Age-cohort changes in the lifetime occurrence of depression and other mental disorders. *J Abnorm Psychol* 1993; **102**: 110–20.

38 Zwaanswijk M, van Dijk CE, Verheij RA. Child and adolescent mental health care in

Dutch general practice: Time trend analyses. *BMC Fam Pract* 2011; **12**: 133.

39 Kelleher KJ, McInerny TK, Gardner WP, Childs GE, Wasserman RC. Increasing identification of psychosocial problems: 1979-1996. *Pediatrics* 2000; **105**: 1313–21.

40 McMartin SE, Kingsbury M, Dykxhoorn J, Colman I. Time trends in symptoms of mental illness in children and adolescents in Canada. *Can Med Assoc J* 2014; **186**: E672–8.

41 Mattisson C, Bogren M, Nettelbladt P, Munk-Jorgensen P, Bhugra D. First incidence depression in the Lundby Study: A comparison of the two time periods 1947-1972 and 1972-1997. *J Affect Disord* 2005; **87**: 151–60.

42 Walters K, Rait G, Griffin M, Buszewicz M, Nazareth I. Recent trends in the incidence of anxiety diagnoses and symptoms in primary care. *PLoS One* 2012; **7**. doi:10.1371/journal.pone.0041670.

43 Simpson KRS, Meadows GN, Frances AJ, Patten SB. Is mental health in the Canadian population changing over time? *Can J Psychiatry* 2012; **57**: 324–31.

44 Murphy JM, Horton NJ, Laird NM, Monson RR, Sobol AM, Leighton AH. Anxiety and depression: A 40-year perspective on relationships regarding prevalence, distribution, and comorbidity. Acta Psychiatr. Scand. 2004; **109**: 355–75.

45 Gould MS, Greenberg T, Velting DM, Shaffer D. Youth suicide risk and preventive interventions: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 2003; **42**: 386–405.

46 Stallard P, Spears M, Montgomery A a, Phillips R, Sayal K. Self-harm in young adolescents (12-16 years): onset and short-term continuation in a community sample. *BMC Psychiatry* 2013; **13**: 328.

47 O'Connor RC, Rasmussen S, Hawton K. Adolescent self-harm: a school-based study in Northern Ireland. *J Affect Disord* 2014; **159**: 46–52.

48 Kidger J, Heron J, Lewis G, Evans J, Gunnell D. Adolescent self-harm and suicidal thoughts in the ALSPAC cohort: a self-report survey in England. *BMC Psychiatry* 2012; **12**: 69.

49 Nock MK, Borges G, Bromet EJ, Cha CB, Kessler RC, Lee S. Suicide and suicidal behavior. *Epidemiol Rev* 2008; **30**: 133–54.

50 Turecki G, Brent DA. Suicide and suicidal behaviour. *Lancet (London, England)* 2016; **387**: 1227–39.

51 Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. *Br J Psychiatry* 1997; **170**: 205–28.

52 Fliege H, Lee J-R, Grimm A, Klapp BF. Risk factors and correlates of deliberate self-

harm behavior: a systematic review. *J Psychosom Res* 2009; **66**: 477–93.

53 Cavanagh JTO, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. *Psychol Med* 2003; **33**: 395–405.

54 Shaffer D, Gould MS, Fisher P, Trautman P, Moreau D, Kleinman M, *et al.* Psychiatric diagnosis in child and adolescent suicide. *Arch Gen Psychiatry* 1996; **53**: 339–48.

55 Beautrais AL. Child and young adolescent suicide in New Zealand. *Aust N Z J Psychiatry* 2001; **35**: 647–53.

56 Brent DA, Baugher M, Bridge J, Chen T, Chiappetta L. Age- and sex-related risk factors for adolescent suicide. *J Am Acad Child Adolesc Psychiatry* 1999; **38**: 1497–505.

57 Grøholt B, Ekeberg O, Wichstrøm L, Haldorsen T. Suicide among children and younger and older adolescents in Norway: a comparative study. *J Am Acad Child Adolesc Psychiatry* 1998; **37**: 473–81.

58 Dilillo D, Mauri S, Mantegazza C, Fabiano V, Mameli C, Zuccotti GV. Suicide in pediatrics: epidemiology, risk factors, warning signs and the role of the pediatrician in detecting them. *Ital J Pediatr* 2015; **41**: 49.

59 Bridge J a., Goldstein TR, Brent D a. Adolescent suicide and suicidal behavior. *J Child Psychol Psychiatry Allied Discip* 2006; **47**: 372–94.

60 Statistics Canada. Suicides and suicide rate, by sex and by age group. 2015. (http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/hlth66d-eng.htm).

61 Värnik A, Kõlves K, Allik J, Arensman E, Aromaa E, van Audenhove C, *et al.* Gender issues in suicide rates, trends and methods among youths aged 15-24 in 15 European countries. *J Affect Disord* 2009; **113**: 216–26.

62 Curtin SC, Warner M, Hedegaard H. Increase in suicide in the United States, 1999-2014. NCHS data brief, no 241. , 2016 (http://www.cdc.gov/nchs/products/databriefs/db241.htm).

63 Marttunen MJ, Aro HM, Henriksson MM, Lönnqvist JK. Mental disorders in adolescent suicide. DSM-III-R axes I and II diagnoses in suicides among 13- to 19-year-olds in Finland. *Arch Gen Psychiatry* 1991; **48**: 834–9.

64 Spirito A, Esposito-Smythers C. Attempted and completed suicide in adolescence. *Annu Rev Clin Psychol* 2006; **2**: 237–66.

65 Randall JR, Walld R, Finlayson G, Sareen J, Martens PJ, Bolton JM. Acute risk of suicide and suicide attempts associated with recent diagnosis of mental disorders: a population-based, propensity score-matched analysis. *Can J Psychiatry* 2014; **59**: 531–8.

66 Poorolajal J, Haghtalab T, Farhadi M, Darvishi N. Substance use disorder and risk of

suicidal ideation, suicide attempt and suicide death: a meta-analysis. *J Public Health (Oxf)* 2016; **38**: e282–91.

67 Oldham JM. Borderline Personality Disorder and Suicidality. *Am J Psychiatry* 2006; **163**: 20–6.

68 American Psychiatric Association. Practice guideline for the assessment and treatment of patients with suicidal behaviors. *Am J Psychiatry* 2003; **160**: 1–60.

69 Kessler RC, Berglund P, Borges G, Nock M, Wang PS. Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990-1992 to 2001-2003. *JAMA* 2005; **293**: 2487–95.

70 Carlborg A, Winnerbäck K, Jönsson EG, Jokinen J, Nordström P. Suicide in schizophrenia. *Expert Rev Neurother* 2010; **10**: 1153–64.

71 Palmer BA, Pankratz VS, Bostwick JM. The lifetime risk of suicide in schizophrenia: a reexamination. *Arch Gen Psychiatry* 2005; **62**: 247–53.

72 Hill GB, Forbes WF, Kozak J. A simple method for estimating incidence from prevalence. *Chronic Dis Can* 1999; **20**: 151–3.

73 Newman SC, Bland RC. Incidence of mental disorders in Edmonton: Estimates of rates and methodological issues. *J Psychiatr Res* 1998; **32**: 273–82.

74 Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G. Recall of depressive episode 25 years previously. *Psychol Med* 1999; **29**: 787–91.

75 Power C, Kuh D, Morton S. From Developmental Origins of Adult Disease to Life Course Research on Adult Disease and Aging: Insights from Birth Cohort Studies. *Annu Rev Public Health* 2013; **34**: 7–28.

76 Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant* 2013; **28**: 2670–7.

77 Lin G, So Y, Johnston G, Nc C. Analyzing Survival Data with Competing Risks Using SAS ® Software. *SAS Glob Forum* 2012; **344**: 1–8.

78 Heeger D. Signal Detection Theory. New York University. New York, New York. 1997: 1–10.

79 Rutjes a WS, Reitsma JB, Coomarasamy a, Khan KS, Bossuyt PMM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007; **11**: iii, ix-51.

80 Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to x-ray techniques. *Public Health Rep* 1947; **62**: 1432–49.

81 Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; **37**: 360–3.

82 Juurlink D, Preyra C, Croxford R, Chong A, Austin P, Tu J, *et al.* Canadian Institute for Health Information Discharge Abstract Database: A Validation Study ICES Investigative Report. *Toronto Inst Clin Eval Sci* 2006; : 69.

83 Canadian Institute for Health Information. CIHI Data Quality Study of Ontario Emergency Department Visits for Fiscal Year 2004-2005: Volume II of IV - Main Study Findings. , 2007.

84 Steele LS, Glazier RH, Lin E, Evans M. Using administrative data to measure ambulatory mental health service provision in primary care. *Med Care* 2004; **42**: 960–5.

85 Kurdyak P, Lin E, Green D, Vigod S. Validation of a Population-Based Algorithm to Detect Chronic Psychotic Illness. *Can J Psychiatry* 2015; **60**: 362–8.

86 Martens PJ, Burchill C, Fransoo R, De Coster C, McKeen N, Ekuma O, *et al.* Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study. , 2004.

87 Bethell J, Rhodes AE. Identifying deliberate self-harm in emergency department data. *Heal reports* 2009; **20**: 35–42.

88 Bethell J, Bondy SJ, Lou WYW, Guttmann A, Rhodes AE. Emergency department presentations for self-harm among Ontario youth. *Can J Public Health* 2013; **104**: e124-30.

89 Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS Procedure for Latent Class Analysis. *Struct Equ Model A Multidiscip J* 2007; **14**: 671–94.

CHAPTER 2: INCIDENCE OF DIAGNOSED MENTAL DISORDERS DURING ADOLESCENCE AND EARLY ADULTHOOD IN A POPULATION-BASED BIRTH COHORT

2.1 Chapter overview

This manuscript describes a birth cohort of individuals born in Manitoba between the years of 1979 and 1993 who lived within the city of Winnipeg on their tenth birthday. Administrative data on diagnoses from medical claims and hospital discharge abstracts were used to estimate cumulative incidence of diagnoses between the age of 10 and 25 years old for five mental disorder clusters (anxiety, mood and adjustment, personality, substance use, and schizophrenia). Previous research on the incidence of these disorders has relied heavily on self-reported lifetime incidence which has been criticized as potentially undercounting true incidence. Using diagnoses recorded in administrative data sets may be a method of estimating these incidences with less bias.

2.2 Abstract

Objective: Mental disorders often emerge during adolescence and early adulthood. Previous research has relied on self-reported diagnosis to determine lifetime incidence, but this significantly underestimates true incidence. The objective of this study was to use administrative data to estimate the cumulative incidence of 5 groups of mental disorders and therefore provide estimates that do not rely on individual self-reported diagnosis.

Methods: This study used administrative data, including physician billing claims and hospital admission records, from Winnipeg, Canada to estimate incidence of diagnosed mental disorders in the general population. Individuals born in Manitoba between the fiscal years of 1979/1980 and 1992/1993 were included in the sample cohort if they were living in the city of Winnipeg on their tenth birthday. Cumulative incidence rates and their 95% confidence intervals (CIs) were estimated at 15, 20 and 25 years of age for: mood and adjustment, anxiety, personality, schizophrenia, and substance use disorders.

Results: Cumulative incidence rates were highest for diagnosed mood disorders at 445 (95%CI: 432 to 458) per 10,000 at age 15; rates increased to 2440 (95%CI: 2410 to 2470) per 10,000 at age 25. Diagnosed anxiety disorders incidence rates were similar, with 351 (95%CI: 339 to 363) at 15, rising to 1960 (95%CI: 1930 to 1980) at age 25. Diagnosed substance use disorder rates were estimated at 84 (95%CI: 78 to 90) per 10,000 at age 15 and 896 (95%CI: 876 to 916) at age 25. Personality disorders and schizophrenia were the least common diagnoses.

Conclusion: Mental disorders affect many adolescents and young adults. Some estimates of lifetime incidence relying on self-report may be underestimating diagnosis of these disorders and birth cohorts research using administrative data is a potentially useful method of obtaining epidemiological data.

2.3 Background

Mental disorders are associated with a significant burden of disease.[1] Unlike many other illnesses associated with a high level of morbidity, these disorders frequently emerge in adolescence and cause a substantial burden among the young.[1–3] Many individuals who have these disorders experience their effects for decades, while others will die prematurely due to the excess risk of mortality from suicide and other causes associated with these conditions.[2–4] Previous studies have reported estimates of lifetime prevalence of mental disorders between 26% and 50% in adults. [5–9] These estimates have proven to be contentious. Some researchers have claimed that the estimates are inflated by the inclusion of non-clinical personalities or the inclusion of mild disorders.[10] Others have suggested that the incidence of disorders is often underestimated due to methodological flaws.[11–13]

Previous epidemiological research has frequently relied on self-reported lifetime incidence of mental disorders.[11,12,14] This method of estimating lifetime incidence has been criticized as being vulnerable to significant bias with even minor rates of recall failure, or other causes of underreporting of mental disorders such as a lack of self-identification as having a disorder even after a physician diagnosis.[11,15] People admitted to hospital with depressive disorders often fail to recall them years later, even when they are serious enough to warrant admission.[15] One study estimated that retrospective studies provide estimates that are half the size of those found using prospectively collected data.[12]This bias will cause a significant underestimation of the occurrence of disorders. Such bias has resulted in a jarring disconnect between fairly high estimates of short-term incidence and prevalence with unexpectedly low rates of lifetime prevalence, particularly among middle-age and older individuals.[11] This effect is

compounded by the high mortality rates of individuals with a mental disorder. People who die from suicide at age 17 cannot complete a survey questionnaire about their lifetime history of mental disorder when they are 25. Therefore long-term cohort studies that involve repeated or continuous measurement of outcomes and focus on cumulative incidence are required to derive estimates for the true incidence of mental disorders. The incidence of these disorders during adolescence and young adulthood are particularly important since this is when most mental disorders begin to manifest. Treatment in this period of life has the potential to significantly affect the life trajectory of individuals with mental disorders, allowing them to establish functional adult lives. Many studies of this group have estimated prevalence at specific times,[16–21] but comprehensive information on cumulative incidence from childhood into adulthood is lacking. Information on incidence and prevalence of these disorders is useful for decisions about resourcing mental health care.[5]

Administrative health data provides a source of data on the occurrence of physician diagnosed disorders. These data are also suitable at determining cumulative incidences for the first occurrence of a mental disorder diagnosis. However, there has been little validation research of ICD-based identification of mental disorder from administration data, and there is no accepted standard for classifying individuals from administrative records using ICD codes.[22–24] One study assessed the accuracy of similar codes for depression, anxiety, and psychosis using American Veterans Health Administration data on diabetic patients.[22] A noticeable discrepancy between administrative data and self-report was found with particular error in false positives for psychosis and anxiety disorders. However the use of a reference standard that likely undercounts cases (self-report) and a population with significant morbidity may be affecting that validation and makes conclusion from that study difficult. A systematic review on depression coding by

Townsend et al. suggested that detection of depression is poor due to a substantial number of untreated individuals not being detected, but agreement is better when compared against medical charts.[24] This suggests that ICD coding may be a reasonable method of detecting diagnosed cases of depression. However a key difference between estimates using administrative data and survey-derived estimates is that administrative data estimates are based on treated and diagnosed conditions. Survey-based research relies either on self-reported and self-identified occurrence of mental illness or an approximation of a clinical assessment through a structured assessment or interview. Since the two methods do not measure exactly the same underlying construct this may result in additional differences in estimates.

This paper aims to assess the lifetime incidence of diagnosed mental disorder among a youth birth cohort in Winnipeg, Canada by using a comprehensive population-based repository of administrative health records, with physician diagnosis of mental illness. The study aims to estimate the cumulative incidence of diagnosed mental disorders in this cohort and test for differences by sex, and birth year.

2.4 Methods

2.4.1 Data sources

Data were acquired from the Population Research Data Repository at the Manitoba Centre for Health Policy. The Data Repository contains linked data for multiple government agencies and services located in the province of Manitoba, with a focus on health service data. Included are records for births that occur within the province. The medical claims database contains records for all service billed by physicians within the province. These records include dates, tariff codes for treatment, and diagnostic codes. The diagnosis is coded according to the International

Classification of Diseases version 9 clinical modification (ICD-9-CM; National Center for Health Statistics 2005; Canadian Institute for Health Information n.d.). The hospital discharge abstract database contains information on inpatient treatment in the province. This includes 25 diagnostic fields coded using ICD-10-CA since April 1, 2004. Prior to April 1, 2004 these records were coded using ICD-9-CM. Data up to and including fiscal year 2013 were used. The Provincial Health Insurance Registry includes information on all individuals registered in the provincial health insurance plan, incorporating data on when individuals begin and end their healthcare coverage.

2.4.2 Cohort creation

Individuals with birth dates spanning the fiscals years of 1979 to 1992 were identified from the registry. These records identify approximately 98% of births occurring within the province[27] with the missing births occurring among those born to Canadian Military and Royal Canadian Mounted Police. These birth years were chosen due to previous research with this cohort, the ages between 10 and 25 years old being of particular interest, and due to data being only available until March 31st, 2014.This study focuses on the population living within the Winnipeg Metropolitan area on their 10th birthday. This focus was chosen due to the unique environment and health care issues that affect the northern and remote regions of the province, whereas the Winnipeg metropolitan area should be more generalizable to other Western nation's urban centres. This region comprises Winnipeg and other communities in its metropolitan area with a population of more than 700,000 people according to the 2011 Statistics Canada Census.[28] Individuals within the cohort were followed until one of the following events occurred: death, out of province migration, or their 25th birthday.

2.4.3 Variable coding

Age was calculated using birth date provided in the registry. Visit dates from hospital discharge abstracts and physician claims and birth dates from the registry were used to derive the dates of diagnosis required to estimate cumulative incidence. Sex was also recorded in the registry file. Physician diagnosed mood and adjustment disorders, anxiety disorders, psychotic disorders, personality disorders, and substance use disorders were derived based on ICD codes recorded in physician billing claims and hospital discharge abstracts. Only diagnosis codes occurring after the age of 10 were used. Individuals could be included in multiple disorders groups if they meet the criteria for more than one disorder. The codes used to determine mental disorders, based on previous work at the Manitoba Centre for Health Policy,[29,30] are reported in **Appendix 2-A**.

2.4.4 Statistical analysis

Frequency of individuals in the Winnipeg cohort, and diagnosis of the disorders was calculated. Disorder-specific follow-up percentages, medians and ranges were determined. Since data ended before the 25th birthday of some individuals, the percent with follow-up until the end of available data was also calculated. In order to examine the lifetime incidence of these disorders between ages 10 and 25 a method of calculating incidences across time was needed. Kaplan-Meier estimation of survival functions can estimate the cumulative incidence of an outcome over time. However, the Kaplan-Meier method cannot account for competing risks, which are events that prevent the outcome of interests from occurring for an individual.[31] Due to the length of the follow-up period, death is a competing risk that may slightly bias estimates upward if those deaths are treated as censoring events.[32] A nonparametric extension of the Kaplan-Meier survival

analysis known as the cumulative incidence function calculates the probability of a specific outcome prior to each time point while adjusting for the occurrence of competing risks.[31]

Cumulative incidence rates were calculated. Graphs of cumulative incidence were derived for the study cohort. Cumulative incidence was estimated per 10,000 at ages 15, 20, and 25 for the overall sample, by sex, and by year of birth. Birth year effect was assessed by splitting the sample into early births (pre-1987) and late birth (1987 and later), with roughly equal numbers in each group. The counting process method of determining variance for cumulative incidences was used to calculate 95% CIs.[31,33] Tests of differences between sexes and birth year groups was conducted using Gray's method.[31,34]

The Epanechnikov Kernel-smoothed Hazard functions for the mental disorders were plotted by age.[35] All analyses were conducted with SAS software version 9.4.[36]

2.4.5 Ethical approval

Ethical approval was obtained from the University of Manitoba's Health Research Ethics Board (HREB #: H2015:009) and approval for data access was obtained from the Manitoba Health Information Privacy Commission (HIPC #- 2014/2015 – 41).

2.5 Results

2.5.1 Study Cohort

The cohort contained 92,890 individuals; frequency of the disorders in the sample and the outcome specific follow-up rates are reported in **Table 2-1**. Approximately half (51.1%) of the cohort members were male. The 13 fiscal years in the data included between 5748 (1980) and

7240 (1990) individuals. Median follow-up time was 15 years (range: 1 week to 15 years) for each disorder-specific analysis.

2.5.2 Mental disorder diagnosis incidence

A total of 30,661 (33%) cohort members had a diagnosis for at least one of the five examined disorders (**Table 2-1**). Anxiety and mood disorders were more common than the other disorders with 16176 (17.4%) and 19736 (21.4%) individuals respectively with an incident diagnosis. Schizophrenia was uncommon with only 795 (0.9%) diagnosed incident cases. Substance use disorders and personality disorders had 6981 (7.5%) and 2014 (2.2%) incident cases, respectively.

Figure 2-1 shows the smoothed hazard function for a new diagnosis of the 5 disorders while cumulative incidence by age is presented in **Table 2-2**. Anxiety and mood disorders were diagnosed at an earlier age than other disorders with many new diagnoses occurring shortly after age 10. Anxiety and mood disorder diagnoses had cumulative incidence rates of 351 (95%CI: 339 to 363) and 445 (95%CI: 432 to 458) per 10,000 by age 15, respectively. The rate of new diagnoses increased significantly for anxiety, mood disorders, and substance use around 12 years of age and peaked near age 18. At this point mood disorders and substance use had a steady rate of new diagnoses until age 25, but the rate for anxiety dropped below the rate for substance use disorder shortly before the end of the follow-up time. The cumulative incidences of these disorders at age 25 were 24.4% for mood disorders and 19.6% for anxiety. Personality disorders and schizophrenia do not change dramatically in hazard rate by age, but instead show a slow increase until around age 15 (personality disorders) and age 18 (schizophrenia) and a stable rate afterwards.

Sex stratified analyses are shown in **Table 2-3**. Substantial differences in incidence between the sexes across these disorders were observed. Anxiety disorders had a cumulative incidence of 2530 diagnoses per 10,000 (95%CI: 2490 to 2570) among females, but only 1370 diagnoses (95%CI: 1340 to 1400) among males at age 25. The estimated incidence of diagnosed mood disorders was 3050 (95%CI: 3000 to 3090) and 1800 (95%CI: 1760 to 1840), respectively. Personality disorders were also more common among females, 281 (95%CI: 265 to 298) versus 220 (95%CI: 207 to 235) per 10,000 individuals. Schizophrenia and substance use disorders were more commonly diagnosed among males. Schizophrenia was diagnosed in 143 per 10,000 males (95%CI: 132 to 155) and 53 per 10,000 females (95%CI: 46 to 61) at age 25. Substance use was diagnosed in 918 males per 10,000 (95%CI: 890 to 946) and 845 per 10,000 females (95%CI: 818 to 873).

Comparing those born before 1987 and those in 1987 onward produced some differences in cumulative incidence for these disorders (depicted in **Table 2-4**). The later cohort had higher diagnosed incidence of anxiety (375 per 10,000; 95%CI: 357 to 392 versus 330; 95%CI: 314 to 346) and mood disorders (475 per 10,000; 95%CI: 453 to 492 versus 420; 95%CI: 402 to 438) at age 15. Substance use disorder diagnosis was also higher at age 15 in the later cohort (96 per 10,000; 95%CI: 88 to 106 versus 72; 95%CI: 65 to 80). At age 20, anxiety disorders were no longer more commonly diagnosed in the later cohort, but mood (1480 per 10,000; 95%CI: 1450 to 1510 versus 1430; 95%CI: 1400 to 1460) and substance use (470 per 10,000; 95%CI: 450 to 490 versus 440; 95%CI: 420 to 460) retained slightly higher incidences. The incidence of diagnosed personality disorders was similar at age 15, but lower in the later cohort (127 per 10,000; 95%CI 117 to 138) than the early cohort (150 per 10,000; 95%CI: 140 to 160) at age 20. Schizophrenia diagnosis had similar incidences in the two cohort strata.

2.6 Discussion

Our estimates for diagnosed mental disorder incidence ranged from 24.4% and 19.6% for diagnosed mood and anxiety disorders down to 2.5% and 1.0% for diagnosed personality disorder and schizophrenia at age 25. These estimates are in one sense higher than expected based on some previous estimates of lifetime risk.[5–9] Earlier research on the lifetime incidence of mental disorders has relied heavily upon self-reported lifetime incidence, which may be biased due to poor recall.[11,15] The diagnosed incidence presented here represents a likely more accurate assessment of diagnosed mental disorders than self-reported lifetime incidence. However, interpreting these results requires considering undiagnosed disorders that are missed, these could consist of a large number of individuals.[37] This means it is possible that the combined incidence of diagnosed and undiagnosed disorders is higher. However, the issue of undercounting due to individuals that do not seek treatment can be contrasted with individuals that are misclassified as having disorders they do not. This is potentially an issue with physicians giving patients provisional diagnoses during the initial stages of a disorder before correctly diagnosing the true condition later. These data do not provide any information that could be used to identify these provisional diagnoses or to rule out previously incorrect diagnostic coding. These results also provide incidence estimates at several age points. The incidence of diagnosed anxiety and mood and adjustment disorders affects a substantial percentage of the population by age 25, suggesting that previous estimates of these disorders may be undercounting the true lifetime incidence.[8] However, comparisons between prevalence estimates based on survey data (survey prevalence) and estimates based on treatment prevalence (such as this study using administrative data) may be different because they are measuring different constructs. Survey estimates may reflect self-perception of mental illness or the presence of symptom clusters, while administrative data

studies only detect individuals that seek treatment and assign diagnosis based on clinical judgment.

This study indicates that a substantial portion of the population received a physician diagnosis for a mental disorder at a young age. Incidence rates of anxiety and mood disorders were particularly high and affected approximately one fifth and one quarter of the cohort, respectively, by age 25. Approximately 9% of individuals were also diagnosed with a substance use disorder. The difference in incidence of new cases of depressive disorders and anxiety disorders is particularly interesting. While mood disorders have a fairly consistent occurrence of new cases by age, new cases of anxiety disorders increase rapidly until age 18 before declining. The estimated incidence of a new diagnosis of anxiety at age 25 is much closer to the hazards for schizophrenia or personality disorder than for substance use and mood disorders. This peak in anxiety incidences might be driven by the stresses of transitioning into adulthood – including employment-related issues, college pressures, moving away from home and changes in social circles after the end of high school. These ages in particular may be worth more thorough examination with respect to the epidemiology and treatment of anxiety disorders.

An American study estimated that 20.4% of 14 to 18 year olds had experienced at least one major depressive disorder episode.[38] Another American survey found adolescent estimates in line with ours by surveying both the adolescents and their parents, though this research had an extremely high prevalence of anxiety disorders due to the inclusion of untreated phobias.[21] Many of the identified individuals in that study might not have sought treatment; therefore the similarity to our results for diagnosed cumulative incidence might still indicate undercounting of the total lifetime incidence. The estimates for disorders here are lower than those found by a prospective American study.[39] The use of repeated mental assessment may have identified more

cases of untreated mental disorder. This study also started at age 9 and provided an estimated incidence at age 16 – a longer time span than the 10 – 15 age period being compared to in this research.

A noticeably higher incidence of mood disorders is reported here than has been previously noted for depression and related disorders in Canada. Survey-based Canadian studies estimated the lifetime incidence of a major depressive episode to be 11.3% and 12.2% and a major depressive disorder to be 9.9% and 10.8%.[14,40] The 1-year prevalence of depression was estimated at 14% in other work using Canadian electronic medical records.[41] Our study found that by age 25, 24.4% of those in the cohort had received at least one diagnosis of a mood or adjustment disorder. Explanations for the higher rate in our research include the wider definition of mood disorders, accounting for individuals with a diagnosis who did not live to age 25, the high risk of recall bias in the previous studies, and the differences in self-reported mental disorders from survey studies and treated mental disorders.

Administrative data have been used to examine the diagnosis incidence of several major mental disorder clusters.[5] This study is the most likely to agree with our study due to its measurement of treatment incidence rather than survey incidence/prevalence. Our study found lifetime cumulative incidences higher than some of the incidences found at 25 years in Pedersen et al;[5] they estimated a cumulative incidence of ~3-4% for mood disorders, and anxiety disorders with ~6-7%. Their estimates of personality disorders at ~2% and schizophrenia and related disorders at ~1.6% at age 25, findings that are very close to the 2.5% and 1.0% found in this research. Pedersen et al's more restrictive method of identifying individuals diagnosed with mental disorder likely undercounts individuals with disorders, since many never require inpatient treatment. Similar patterns of new cases were obtained in both studies. A large increase in new

cases starting around age 14/15 which peaks at around 20 years of age was found for these disorders.[5] Our work noted a very pronounced peak around age 18 for anxiety disorders with a rapid decline afterwards. The decline in new cases after the peak was much slower among the other disorders in our study compared to Pedersen et al. We also showed an early separation between males and females with respect to schizophrenia incidence, while Pedersen et al. found cumulative incidence to begin to diverge only during the early twenties. Previous research in the province of Manitoba (Martens et al. 2004) had estimated a 5-year cumulative prevalence of 24% for these disorders in those aged 10 and older (compared to the 24.4% with mood disorders alone in this study, and 33% being diagnosed with any disorder without adjustment for censoring). Since mental disorders are often episodic many individuals with a lifetime history would not be identified within the 5-year window.

The major strength of this study is that a population-based birth cohort was used to derive incidence estimates. The use of comprehensive, population-based administrative data ensures the capture of all diagnoses made by healthcare providers within the province. Competing risks and censoring events were taken into account during the estimation of the cumulative incidence. Previous work using adult self-report to determine lifetime incidence are potentially biased due to the increased mortality associated with these disorders. Long term birth cohorts have also suffered substantially from loss to follow-up and uneven occurrence of follow-up assessments.[42] The population of Winnipeg is relatively stable with less people lost to follow-up compared to prospectively collected birth cohorts, resulting in over 86% of people being followed to the end of the observation period. The use of administrative data means that the exact dates of diagnosis are known; this avoids the issues of uneven follow-up and missed events due to recall error.

One study limitation is the coding of mental disorders in administrative data is subject to misclassification bias. Some conditions might be misdiagnosed or some people diagnosed with a mental disorder when they do not fulfil the accepted DSM criteria for that illness. These data do not allow the coding of provisional diagnoses which might result in additional misclassification that could be inflating the estimates. The exact date of onset is also uncertain due to potential delays in seeking treatment. It is unclear whether misdiagnosed disorders or missed disorders causes a larger bias, therefore the net bias of overall incidence of disorders is uncertain. Males seek treatment less for these disorders and therefore may be undercounted. Decreased stigma over time may increase the diagnosis and treatment of these disorders without representing an actual increase in incidence. Information about severity of diagnosed conditions is not available in administrative data. Diagnosis of personality disorders in adolescents is contentious, with some clinicians and researchers believing that only adults should be diagnosed with these disorders.[43] It is possible that this may lead to adolescents not being diagnosed with a personality disorder even though they meet the criteria. Physician claim codes are not sufficiently specific to isolate bipolar disorders from unipolar depressive disorders. Therefore these disorders need to be combined when using physician claims data to determine incidence. This study uses one physician visits to diagnosis most conditions whereas coding diagnosis using multiple visits may reduce the occurrence of misclassification. The reduction in prevalence estimates between using a single-encounter coding method (as is being done for most conditions in this study) and more restrictive methods is around 25.7% to 34.9%.[23] Another limitation is that knowing these cumulative incidence rates may not have direct clinical implications because these conditions are frequently episodic and often do not require lifetime treatment. Another issue is that ICD-10-CA coding is used in hospital discharge abstracts starting in 2004 in Manitoba instead of ICD-9-CM.

Coding of mental disorders is not identical between the two systems and this may affect the estimates of mental disorders for individuals born later in the cohort. Improvement in the coding schema should result in more accurate diagnosis using ICD-10-CA, however agreement between ICD-10 and ICD-9 was high for mental disorders in validation assessment by the WHO.[44] Another issue is the diagnoses occurring before the age of ten were not included. The relationships between paediatric, adolescent, and adult disorders is still not fully understood. Including diagnoses occurring before age 10 would increase the incidence estimates for the disorders, but these paediatric diagnoses may not be sufficiently similar to later onset diagnoses. Only diagnosis after age 10 was used in order to minimize the blurring of paediatric and adolescent/adult disorders and to be consistent with prior research using these data.[29-30] This age should help to maintain good specificity at detecting non-paediatric conditions without causing a marked decrease in incidence estimates.

2.7 Conclusions

These data provide a better understanding of the cumulative incidences of mental disorders in youth and underscore the importance of policy initiatives to identify, prevent, and treat mental disorder early. A large portion of individuals receive one or more psychiatric diagnoses at a young age. These conditions emerge early in life; without successful treatment, many adolescents and young adults are at risk of engaging in serious suicidal behaviour and dying at a young age.[45] The acuity of the risk for serious suicidal behaviour after initial diagnosis underscores the clinical importance of early interventions that target this population.[46]

2.8 References

1 Whiteford H, Ferrari A, Degenhardt L. Global Burden Of Disease Studies: Implications For Mental And Substance Use Disorders. *Health Aff* 2016; **35**: 1114–20.

2 Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, *et al.* Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2197–223.

3 Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: An analysis from the global burden of disease study 2010. *PLoS One* 2015; **10**: 1–14.

4 Eaton WW, Roth KB, Bruce M, Cottler L, Wu L, Nestadt G, *et al.* The relationship of mental and behavioral disorders to all-cause mortality in a 27-year follow-up of 4 epidemiologic catchment area samples. *Am J Epidemiol* 2013; **178**: 1366–77.

5 Pedersen CB, Mors O, Bertelsen A, Waltoft BL, Agerbo E, McGrath JJ, *et al.* A Comprehensive Nationwide Study of the Incidence Rate and Lifetime Risk for Treated Mental Disorders. *JAMA Psychiatry* 2014; **71**: 573.

6 Robins L, Regier D. *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study*. Free Press, 1991.

7 Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, *et al.* Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994; **51**: 8–19.

8 Kessler RC, Angermeyer M, Anthony JC, DE Graaf R, Demyttenaere K, Gasquet I, *et al.* Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry* 2007; **6**: 168–76.

9 Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, *et al.* Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004; **420** : 21–7.

10 Kessler RC, Merikangas KR, Berglund P, Eaton WW, Koretz DS, Walters EE. Mild Disorders Should Not Be Eliminated From the DSM-V. *Arch Gen Psychiatry* 2003; **60**: 1117.

11 Patten SB. Recall bias and major depression lifetime prevalence. *Soc Psychiatry Psychiatr Epidemiol* 2003; **38**: 290–6.

12 Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G, *et al.* How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychol Med* 2010; **40**: 899–909.

13 Hamdi NR, Iacono WG. Lifetime prevalence and co-morbidity of externalizing disorders and depression in prospective assessment. *Psychol Med* 2014; **44**: 315–24.

14 Patten SB, Williams JVA, Lavorato DH, Wang JL, McDonald K, Bulloch AGM. Descriptive epidemiology of major depressive disorder in Canada in 2012. *Can J Psychiatry* 2015; **60**: 23–30.

15 Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G. Recall of depressive episode 25 years previously. *Psychol Med* 1999; **29**: 787–91.

16 Reinherz HZ, Giaconia RM, Lefkowitz ES, Pakiz B, Frost AK. Prevalence of psychiatric disorders in a community population of older adolescents. *J Am Acad Child Adolesc Psychiatry* 1993; **32**: 369–77.

17 Costello EJ, Angold A, Burns BJ, Stangl DK, Tweed DL, Erkanli A, *et al.* The Great Smoky Mountains Study of Youth. Goals, design, methods, and the prevalence of DSM-III-R disorders. *Arch Gen Psychiatry* 1996; **53**: 1129–36.

18 Cohen P, Cohen J, Kasen S, Velez CN, Hartmark C, Johnson J, *et al.* An epidemiological study of disorders in late childhood and adolescence--I. Age- and gender-specific prevalence. *J Child Psychol Psychiatry* 1993; **34**: 851–67.

19 Maalouf FT, Ghandour LA, Halabi F, Zeinoun P, Shehab AAS, Tavitian L. Psychiatric disorders among adolescents from Lebanon: prevalence, correlates, and treatment gap. *Soc Psychiatry Psychiatr Epidemiol* 2016. doi:10.1007/s00127-016-1241-4.

20 Kessler RC, Avenevoli S, Costello EJ, Georgiades K, Green JG, Gruber MJ, *et al.* Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Arch Gen Psychiatry* 2012; **69**: 372–80.

21 Merikangas K, Jian-ping H, Burstein M, Swanson S, Avenevoli S, Lihong C, *et al.* Lifetime Prevalence of Mental Disorders in US Adolescents: Results from the National Comorbidity Study-Adolescent Supplement. *J Am Acad Child Adolesc Psychiatry* 2011; **49**: 980–9.

22 Frayne SM, Miller DR, Sharkansky EJ, Jackson VW, Wang F, Halanych JH, *et al.* Using administrative data to identify mental illness: what approach is best? *Am J Med Qual* 2010; **25**: 42–50.

23 Bauer MS, Lee A, Miller CJ, Bajor L, Li M, Penfold RB. Effects of diagnostic inclusion criteria on prevalence and population characteristics in database research. *Psychiatr Serv* 2015; **66**: 141–8.

24 Townsend L, Walkup JT, Crystal S, Olfson M. A systematic review of validated methods for identifying depression using administrative data. *Pharmacoepidemiol Drug Saf* 2012; **21**: 163–73.

25 National Center for Health Statistics. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). , 2005.

26 Canadian Institute for Health Information. ICD-9/CCP and ICD-9-CM. (https://www.cihi.ca/en/data-and-standards/standards/classification-and-coding/icd-9ccp-and-icd-9-cm).

27 Wall-Wieler E, Roos LL, Chateau DG, Rosella LC. What predictors matter: Risk factors for late adolescent outcomes. *Can J Public Heal* 2016; **107**: e16–22.

28 Statistics Canada. Focus On Geography Series, 2011 Census. , 2012.

29 Fransoo R, Martens P, Burland E, The Need to Know Team, Prior H, Burchill C. Manitoba RHA Indicators Atlas 2009. , 2009 (http://mchp-appserv.cpe.umanitoba.ca/reference/RHA_Atlas_Report.pdf).

30 Martens PJ, Burchill C, Fransoo R, De Coster C, McKeen N, Ekuma O, *et al.* Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study. , 2004.

31 Lin G, So Y, Johnston G, Nc C. Analyzing Survival Data with Competing Risks Using SAS ® Software. *SAS Glob Forum* 2012; **344**: 1–8.

32 Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant* 2013; **28**: 2670–7.

33 Andersen P, Borgan O, Gill R, Keiding N. *Statistical Models Based on Counting Processes*. Springer, 1995.

34 Gray R. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat* 1988; **16**: 1141–54.

35 SAS. The LIFETEST Procedure. 2016. (http://support.sas.com/documentation/cdl/en/statug/68162/HTML/default/viewer.htm#statug_lifetest_details14.htm).

36 SAS Institute Inc. SAS software. 2012.

37 Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, *et al.* Worldwide Use of Mental Health Services for Anxiety, Mood, and Substance Disorders: Results from 17 Countries in the WHO World Mental Health Surveys. *Lancet* 2007; **370**: 841–50.

38 Lewinsohn PM, Rohde P, Seely JR. Major depressive disorder in older adolescents: Prevalence, risk factors, and clinical implications. *Clin Psychol Rev* 1998; **18**: 765–94.

39 Costello EJ. Prevalence and Development of Psychiatric Disorders in Childhood and Adolescence. *Arch Gen Psychiatry* 2003; **60**: 837.

40 Patten SB, Jian LW, Williams JVA, Currie S, Beck CA, Maxwell CJ, *et al.* Descriptive epidemiology of major depression in Canada. *Can J Psychiatry* 2006; **51**: 84–90.

41 Wong ST, Manca D, Barber D, Morkem R, Khan S, Kotecha J, *et al.* The diagnosis of

depression and its treatment in Canadian primary care practices: an epidemiological study. *C open* 2014; **2**: E337-42.

42 Power C, Kuh D, Morton S. From Developmental Origins of Adult Disease to Life Course Research on Adult Disease and Aging: Insights from Birth Cohort Studies. *Annu Rev Public Health* 2013; **34**: 7–28.

43 Gask L, Evans M, Kessler D. Personality disorder. *BMJ* 2013; **347**: f5276–f5276.

44 Maier W, Philipp M, Zaudig M. Comparison of the ICD-10-classification with the ICD-9- and the DSM-III-classification of mental disorders. *Pharmacopsychiatry* 1990; **23 Suppl 4**: 183–7.

45 Hawton K, Saunders KE a, O'Connor RC. Self-harm and suicide in adolescents. *Lancet* 2012; **379**: 2373–82.

46 Randall JR, Walld R, Finlayson G, Sareen J, Martens PJ, Bolton JM. Acute risk of suicide and suicide attempts associated with recent diagnosis of mental disorders: a population-based, propensity score-matched analysis. *Can J Psychiatry* 2014; **59**: 531–8.

Table 2-1: ICD codes and number of visits used to determine diagnoses of mental disorder

Anxiety	Hospital records (1+ visits): ICD–9–CM: 300.0, 300.2, 300.3; ICD-10–CA: F40, F41.0, F41.1,F41.3, F41.8, F41.9, F42 Physician claims (3+ visits): ICD–9–CM: 300
Mood and adjustment disorders	Hospital records(1+ visits): ICD–9–CM: 296.2–296.8, 300.4, 309, 311; ICD–10–CA: F31, F32, F33, F34.1, F38.0, F38.1, F41.2, F43.1, F43.2, F43.8, F53.0, F93.0 Physician claims (1+ visits): ICD–9–CM: 296, 309, 311
Personality disorders	Hospital records or physician claims (1+ visits): ICD-9-CM: 301; ICD-10-CA: F34.0, F60-F62, F68.1,F68.8, F69
Substance use	Hospital records or physician claims (1+ visits): ICD–9–CM: 291, 292, 303, 304, 305; ICD–10–CA: F10–F19, F55
Schizophrenia	Hospital records or physician claims (1+ visits): ICD–9–CM: 295; ICD–10–CA: F20, F21, F23.2, F25

Number of visits indicates how many physician claims, or inpatient records with the diagnosis are required before classifying an individual as a case

Table 2-2: Study cohort characteristics and follow-up

	N	%	Outcome-specific follow up (%)	
			Until age 25	To end of study
Study cohort (age 10)	92890	100		
Mental disorder diagnosis*				
Anxiety	16176	17.4	65.1	87.7
Mood and adjustment	19736	21.2	66.7	88.4
Personality	2014	2.2	60.2	86.7
Schizophrenia	795	0.9	59.8	86.6
Substance use	6981	7.5	62.0	87.3

Table 2-3: Cumulative Incidence per 10,000 of selected diagnosed mental disorders at three ages

Mental disorder diagnosis	Age 15	95% CI	Age 20	95% CI	Age 25	95% CI
Anxiety	351	339 to 363	1270	1250 to 1290	1960	1930 to 1980
Mood and adjustment	445	432 to 458	1450	1430 to 1480	2440	2410 to 2470
Personality	41	37 to 45	139	132 to 147	254	244 to 265
Schizophrenia	5.5	4.1 to 7.1	51	46 to 55	101	94 to 108
Substance use	84	78 to 90	452	438 to 465	896	876 to 916

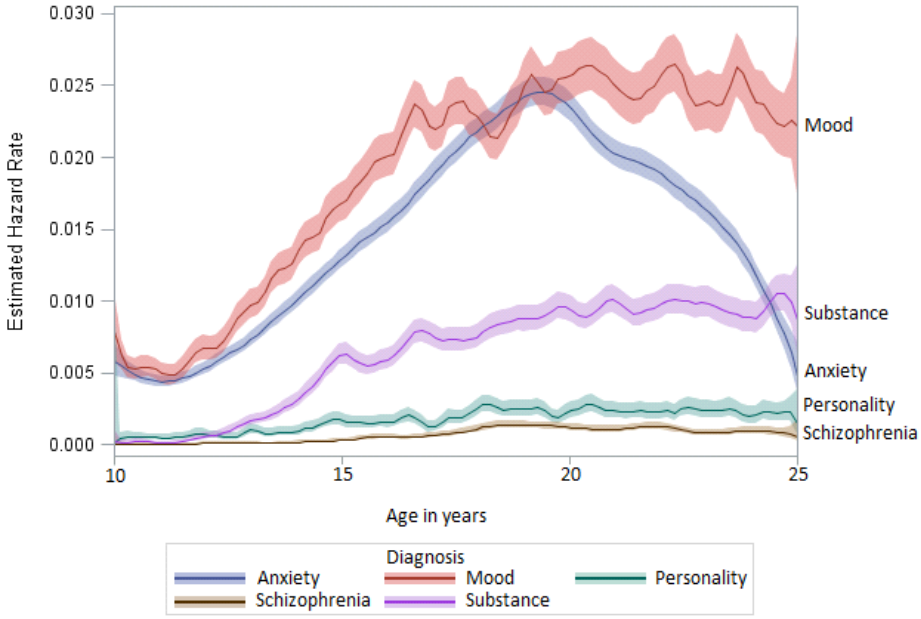
Table 2-4: Cumulative incidence per 10,000 of selected mental disorders at three ages by sex

Mental disorder diagnosis	Males					
	Age 15	95% CI	Age 20	95% CI	Age 25	95% CI
Anxiety	240	226 to 254	813	788 to 848	1370	1340 to 1400
Mood/adjust.	366	350 to 383	1060	1030 to 1080	1800	1760 to 1840
Personality	41	36 to 47	121	111 to 131	220	207 to 235
Schizophrenia	7.4	5.3 to 10.3	70	63 to 78	143	132 to 155
Substance use	67	60 to 74	439	421 to 458	918	890 to 946
Mental disorder diagnosis	Females					
	Age 15	95% CI	Age 20	95% CI	Age 25	95% CI
Anxiety	461	442 to 480	1720	168 to 175	2530	2490 to 2570
Mood/adjust.	517	497 to 538	1840	181 to 188	3050	3000 to 3090
Personality	39	34 to 46	154	143 to 165	281	265 to 298
Schizophrenia	3.5	2.1 to 5.7	29	24 to 34	53	46 to 61
Substance use	100	91 to 109	453	434 to 473	845	818 to 873

Table 2-5: Cumulative incidence per 10,000 of disorders by birth cohort

	1979-1986			
Mental disorder diagnosis	Age 15	95% CI	Age 20	95% CI
Anxiety	330	314 to 346	1260	1230 to 1290
Mood/adjust	420	402 to 438	1430	1400 to 1460
Personality	41	36 to 47	150	140 to 160
Schizophrenia	6.3	4.4 to 8.9	49	43 to 56
Substance use	72	65 to 80	439	421 to 458
	1987-1993			
	Age 15	95% CI	Age 20	95% CI
Anxiety	375	357 to 392	1280	1250 to 1310
Mood/adjust	475	453 to 492	1480	1450 to 1510
Personality	40	35 to 46	127	117 to 138
Schizophrenia	4.5	2.9 to 6.9	52	45 to 59
Substance use	96	88 to 106	470	450 to 490

Figure 2-1: Epanechnikov kernel-smoother hazard functions for diagnosis with a mental disorder with 95% confidence limits



CHAPTER 3: INCIDENCE OF SUICIDE AND SUICIDE ATTEMPTS IN A POPULATION-BASED ADOLESCENT BIRTH COHORT

3.1 Chapter overview

Suicidal behaviour is a significant concern among adolescents and young adults. These behaviours emerge in young adulthood and are responsible for a disproportionate amount of mortality and morbidity. Part of the reason for the emergence of these behaviours at this age is because it is also a period for diagnosis of mental disorders. The previous paper showed that the cumulative incidence of anxiety, mood and adjustment disorders, personality disorders, schizophrenia, and substance use disorders increases around 15 years of age and remains high thereafter (with the sole exception of anxiety disorders). Given that previous research shows risk of suicide and suicide attempt is particularly high in the first year after diagnosis there should be a concurrent rise in suicide attempts along with the diagnosis of new mental disorders.

The next paper seeks to examine the incidence of hospitalized attempts and deaths from suicide. This paper will determine at what age the hospitalized attempts and deaths begin to occur and whether the incidences differ by year of birth, and sex. It will also examine the incidence of suicidal behaviours after the first recorded treatment of the disorders examined in the previous paper. The results of this paper will illustrate how the occurrence of suicidal behaviours and show how they are associated with the emergence of mental disorders during this developmental period.

3.2 Abstract

Objective: The emergence of adult mental disorders in adolescence coincides with an increase in suicide attempts and deaths. Suicide attempts and deaths are common among youths, but challenging to accurately estimate. Previous incidence estimates using self-report methods may be underestimating the true incidence. This study aimed to estimate the incidence of suicide deaths and hospitalized suicide attempts among youth.

Method: A birth cohort of individuals born between the fiscal years of 1979/1980 and 1992/1993 and living in Winnipeg, Manitoba on their 10th birthday was derived from administrative data. This cohort allowed the first occurrence of mental disorder diagnosis and hospitalization due to a suicide attempt after age 10 to be identified for members of the cohort. Administrative data from vital statistics, medical claims and hospital discharge abstracts was used to identify deaths, hospitalized attempts and mental health diagnosis. The incidence of hospitalized suicide attempts and suicide death between ages 10 and 25 years was estimated. Incidence within 10-years of diagnosis for anxiety, mood and adjustment disorders, schizophrenia, personality and substance disorders were also estimated.

Results: Approximately one-percent (1.03%; 95% confidence interval: 0.97% to 1.11%) of individuals had a hospital admission for a suicide attempt by age 25. Incidence was higher among the mentally ill – particularly within 1 year of diagnosis. 129 individuals (95% confidence interval: 106 to 156) per 100,000 died from suicide. Suicide death occurred particularly more often among those with personality, schizophrenia, and substance use (1-2% after 10 years).

Conclusions: Suicidal behaviours are common, particularly within one year of a new mental illness diagnosis.

3.3 Background

Suicide and suicide attempts are a major cause of mortality and morbidity among youth.[1,2] Youth suicide rates increased between 1960 and 1990 but have seen a decline since 1990.[1,3] However, a CDC report suggests that this trend has reversed in recent years in the United States.[4] The rate of adolescent suicide in Canada, specifically, decreased between 1991 and 1997, but this was driven by a decrease in the rate among males.[5] Mental illnesses are often considered a major, if not the primary, underlying force that leads to suicidal behaviour. Mental illnesses have been found to occur in over 90% of those who have died from suicide.[5,6] Mental health research frequently relies on self-report for estimating the incidence and prevalence of suicidal behaviour.[7,8] This method is prone to undercounting due to recall bias.[7,9] As an example, people admitted with depressive disorders serious enough to warrant inpatient admission often fail to recall being diagnosed with depression.[9] In addition, the high mortality rates of individuals with histories of suicidal behaviour can exacerbate this bias since they may not survive long enough to be surveyed. Therefore, estimates of lifetime incidence of suicidal behaviour, and specifically for those with mental illnesses, should be obtained using methods that do not rely on patient recall.

This study examines a birth cohort in Manitoba, Canada containing births in the fiscal years of 1979/80 to 1992/93. The study aims to examine the cumulative incidences for suicide and hospital-treated suicide attempts and how they vary based on birth year, sex, and mental illness diagnosis within this cohort between the ages of 10 and 25 years old.

3.4 Methods

3.4.1 Data sources

Data were acquired from the Population Research Data Repository (The Repository) at the Manitoba Centre for Health Policy. The Repository holds linkable administrative data from government agencies and services in the province of Manitoba, Canada, including the province's publically-funded health system. The specific data sources used were the hospital discharge abstract, medical claims, and Provincial Health Insurance Registry databases. Births in the province are recorded in the Provincial Health Insurance Registry database. The hospital discharge abstract database contains information on inpatient treatment in the province. These data capture up to 25 diagnosis fields coded using the Canadian version of the 10th edition of the International Classification of Diseases (ICD-10-CA) since April 1, 2004. Prior to April 1, 2004 these abstracts were coded using the Clinical modification of the 9th edition of the International Classification of Diseases (ICD-9-CM). Data up to and including fiscal year 2013 were used. The medical claims database contains records for all services billed by physicians within the province. These records include treatment dates and diagnostic codes. Diagnosis is coded using (ICD-9-CM).[10] The registry database has information on all individuals registered in the provincial health insurance plan. This includes data on place of residence and when individuals begin and end their coverage. The Repository contains records for 98% percent of births that occur within the province, individuals born to parents working for the Canadian Military, and the Royal Canadian Mounted Police may be missed due to health care for these individuals being provided via federal funding.

3.4.2 Cohort creation

Individuals born between April 1st 1979 and March 31st 1993 were included in the cohort if they resided in the Winnipeg Regional Health Authority region, the largest health region in the province, on their 10th birthday. Winnipeg and its metropolitan area have a population of approximately 793,000 people.[11] Individuals were followed until one of the following occurred: death, moving out of the province, or their 25th birthday.

3.4.3 Variable coding

Age was taken from the birth date in the registry. Physician-diagnosed mood and adjustment disorders, anxiety, psychotic disorder, personality disorder, and substance use disorder were derived based on ICD coding in medical claims and hospital discharge abstracts. Only diagnostic codes which occurred after the age of 10 years were used. The codes used to determine mental illness, suicide attempt hospitalization and suicide deaths, based on previous work at the Manitoba Centre for Health Policy,[12,13] are located in **Table 3-1**. Individuals could be classified as having multiple disorders. Dates from hospital discharge abstracts and medical claims, and birth dates from the registry were used to derive the dates required to calculate the cumulative incidences. Sex was recorded in the registry file.

3.4.4 Statistical analysis

Frequencies were estimated for births in the province, Winnipeg residents at age 10, attempt hospitalizations, deaths from suicide, and diagnosis of the mental disorders. Information on follow-up were examined for the sample by outcome. The percent of individuals followed until age 25 or until an outcome occurred was estimated. Since the data did not incorporate all individuals until their 25th birthday, the percent of individuals followed until the end of available

data was also determined. Follow-up was also examined within the mental disorder groups using suicide death as the outcome.

A graph of the cumulative incidence was derived for both outcomes using the cumulative incidence function method of estimating cumulative incidence.[14] This method is also known as the cumulative incidence competing risk method and is a non-parametric method of estimated incidence in the presence of competing risks.[14,15] This method adjusts the cumulative incidence estimates to reflect that those with competing events, in this instance death, are no longer at risk for the event of interest, rather than censoring them.[15] Additional graphs were produce by sex, birth year (1979 to 1986, 1987 to 1993), and mental health diagnosis. For the mental health diagnoses, start time for the survival analysis was the date of first recorded diagnosis after the age of 10. For the other graphs the starting time was age 10. Cumulative incidence was extracted from the graphs at ages 15, 20, and 25 for the overall sample and the demographic groups. Cumulative incidence functions were estimated by mental health group at 1-year, 5-year, and 10-year time points after first recorded diagnosis. Individuals with multiple diagnoses were included in estimation for all their diagnosed disorders. Additional cumulative incidences for suicide deaths and suicide attempts were derived starting from the date of first attempt among the suicide attempters. These incidences show the cumulative incidence of second and third admitted attempts. These incidences were also calculated using 1-year, 5-year, and 10-year time points after the first detected admission. A graph of the Epanechnikov Kernal-Smoothed Hazard functions by sex was obtained for the suicide attempt outcome.[16] This graph shows the rate of new incidences of suicide attempt admissions at each age point. Statistical significance of demographic differences was derived using the Fine and Gray test for Equality of Incidence Functions.[17] Analyses were performed using SAS software version 9.4.[18]

3.4.5 Ethical approval

Approval was obtained from the University of Manitoba's Health Research Ethics Board (HREB #: H2015:009) and approval for data access was provided by the Government of Manitoba's Health Information Privacy Commission (HIPC #- 2014/2015 – 41).

3.5 Results

3.5.1 Sample

The birth cohort contained 92,890 individuals; descriptive statistics and outcome specific follow-up rates are located in **Table 3-2**. A total of 47,428 (51.1%) cohort members were male. The 13 full fiscal years in the data included between 5748 (1980) and 7240 (1990) individuals. Due to the large sample and high follow-up rate the median follow-up time was 15 years with a range of 0 to 15 years.

3.5.2 Mental health diagnosis incidence

One or more of these five disorders were found among 30,661 (33%) individuals (**Table 3-2**). Anxiety and mood and adjustment disorders were common, with 17.4% and 21.4% individuals having a recorded diagnosis. Substance use disorders and personality disorders had 7.5% and 2.2% cases. Schizophrenia was the least common diagnosis with only 0.9% incidences.

3.5.3 Suicidal behaviour by birth cohort and sex

There were 105 suicide deaths identified for the cohort. **Table 3-3** contains the estimated cumulative incidence for suicide deaths. There were 8.9 (95%CI= 4.3 to 17.0) suicide deaths per 100,000 at age 15; this number increased to 60 (95%CI= 46 to 78) and 129 (95%CI= 106 to 156) per 100,000 at ages 20 and 25. There was no significant difference in suicide incidence based on

year of birth (p=0.77). The incidences of suicide at age 15 were similar for males and females, but by age 25 there was a significantly higher occurrence of suicide death among males (p=0.0378).

Among those in the cohort 875 hospitalized attempts were recorded. **Table 3-4** contains the estimates for suicide attempt admissions. The incidence per 10,000 was 20 at age 15, increasing to 71 and 103 at ages 20 and 25. **Figure 3-1** shows the Epanechnikov Kernal-Smoothed hazard rate by sex. This figure reveals a rapid increase in the occurrence of first admissions for suicide attempts starting around age 11 for females and growing rapidly till age 15. The rate then decreases with similar rapidity until it approximates the hazard among males around age 18. Admissions begin to increase for males around age 13, reach a peak around age 17, and begin a slow decline. The incidence of suicide attempt admissions for those born in 1986 or later (92 per 10,000) decreased significantly compared to the earlier half of the cohort (114 per 10,000; p=0.0011). The difference is largest at age 15 (15 per 10,000 versus 24 per 10,000) and shrinks over time.

3.5.4 Mental disorder and suicidal behaviour

The incidence of both suicide attempt admission and suicide deaths were elevated among those diagnosed with any of the 5 investigated mental disorders. The 1-year cumulative incidence of suicide deaths for the personality disorders and substance use groups both exceeded the estimated incidence at age 25 for the population as a whole (190 and 150 versus 129 per 100,000). Attempt admissions were high for those with personality disorders or schizophrenia during the first year (271 per 10,000 for both). Anxiety disorders were associated with the lowest occurrence of both suicide deaths and attempts at all three time points. Schizophrenia and

personality disorders showed the highest incidence of attempts. These disorders and substance use disorder had the highest incidence of suicide death within 10-years of diagnosis. Schizophrenia had the highest incidence of deaths (2103 per 100,000) but also a large confidence interval (552 to 5678) due to the small number of people diagnosed with the condition and having a sufficient follow-up period.

3.5.5 Incidence of suicide admissions and deaths after first suicide attempt

The incidence of repeat attempt admission and of suicide death was calculated for the 875 individuals with attempt admissions. The cumulative incidence for death was 4.6 (95%CI: 1.6 to 11.3) deaths per 1,000 individuals within one year of admission, increasing to 13.2 (95%CI: 6.7 to 23.5) and 28.6 (95%CI: 15.4 to 48.3) suicide deaths by 5 and 10 years. Males with an attempt admission were more likely to die from suicide compared to females (p=0.0346). Incidence for males was 9.6 (95%CI: 2.7 to 26.1) cases per 1,000 after 1-year compared to 1.8 (95%CI: 0.2 to 9.7) cases for females. These incidences increased to 27.1 (95%CI: 11.7 to 53.4) and 34.1 (95%CI: 15.3 to 65.1) for males at 5 and 10 years from first attempt. Females had incidences of 6.1 (95%CI: 1.7 to 16.8) and 22.6 (95%CI: 9.4 to 46.1) at 5 and 10 years. The incidence of deaths between those born early or late in the cohort (p=0.1429) showed no significant differences.

The incidence of a second attempt admission was 9.1 (95%CI: 7.3 to 11.1) cases per 100 during the first year after the first admission. This increased to 13.6 (95%CI: 11.3 to 16.1) and 17.2 (95%CI: 14.3 to 20.4) after 5 and 10 years. The incidence of a third admission was 2 (95%CI: 1.2 to 3.1) cases per 100 after 1 year, and 4.7 (95%CI: 3.3 to 6.4) and 6.8 (95%CI: 5.0 to 9.0) cases after 5 and 10 years. The sexes did not differ significantly in the incidence of a

second attempt admission (p=0.2), but females showed a significantly higher incidence of a third admission (p=0.0349). There were 0.6 (95%CI: 0.1 to 2.2) cases with three admissions per 100 for males compared to 2.7 (95%CI: 1.6 to 4.4) cases per 100 for females. These increased to 2.7 (95%CI: 1.1 to 5.5) and 5.2 (95%CI: 2.0 to 10.6) for males, and 5.8 (95%CI: 4.0 to 8.1) and 7.9 (95%CI: 5.6 to 10.7) for females at 5 and 10 years. There was a significantly higher incidence of two admissions for those born early in the cohort (p=0.0257), but no difference for the incidence of three admissions. Those born before 1987 had second attempt incidence of 9.7 (95%CI: 7.3 to 12.4) and 15.8 (95%CI: 12.7 to 19.2) at 1 and 5 years from the first admission. Second attempt incidence was 8.3 (95%CI: 5.7 to 11.5) and 10.4 (95%CI: 7.4 to 13.9) for those born in 1987 or later.

3.6 Discussion

This study provides estimates for the cumulative suicide attempt admission incidence using administrative data not prone to recall bias. Serious suicidal behaviour was common. Previous research on the lifetime incidence of suicidal behaviour has relied heavily upon self-reported lifetime incidence, which is heavily biased due to poor recall.[7,9] Suicide was responsible for approximately 1 death per 1,000 individuals by age 25, and 1 out of every 100 people were hospitalized due to a suicide attempt. Death by suicide is particularly high among those with personality disorders and schizophrenia, with 1 and 2 out of every 100 individuals with a diagnosis for these disorders dying from suicide within 10-years of diagnosis. These results illustrate the importance of treatment and prevention efforts for suicidal behaviour in this group.

The estimates for cumulative incidence of suicide agree with the rates of suicide found across Canada in Statistics Canada data.[21] The incidence of suicides and attempts by sex agrees with previous research showing more suicides deaths among males, but more suicide attempts

among females.[1,2] In our study males had a significantly higher incidence of deaths, whereas females had a higher occurrence of attempts. Rates of suicide among 15 to 24 year old males was estimated at 20.2 per 100,000, compared to 5.5 for females in Canada for the year 2000,[2] which is greater than the 156 to 101 disparity incidence in our study. The rates across Canada were more similar among 5-14 to year olds, 1.4 and 0.9 respectively, which agrees with our findings in this cohort. Recent American data show that youth suicides for males decreased by approximately 1.5 percent per year between 1994 and 2012.[22] Females had a non-significant increase of 0.7 percent per year. The incidence of suicide deaths grew among the later cohort, though this increase is not significant. The population of Winnipeg may also be experiencing an increase in suicide deaths recently among females, but this trend is not detectable due to insufficient power. Previous research has suggested that males comprise around 80% of deaths from suicide,[2,23] but our study found males to comprise around 60% of deaths by age 25.

The pattern of serious suicidal behaviour following diagnosis varies by diagnosed mental disorder. Personality disorders and schizophrenia are marked by particularly acute short-term risk. Mood disorders closed some of the gap with both personality disorders and schizophrenia around 10 years post-diagnosis. Previous research has also suggested a period of acute risk within the first year after a diagnosis is made.[24] These results suggest that 20% of the suicide mortality for personality disorders over the first 10 years occurs within the first year. Around 15% of suicide deaths among those with substance use disorders occur within the first year, compared to 10% for anxiety and mood disorders. The finding of a U-shaped mortality pattern for mood disorders, and to a lesser extent anxiety, is interesting. Suicide risk may increase over time for those with poorly controlled anxiety or mood disorders. This may partially explain the occurrence of a peak in suicide rate in the 20-24 age group in Canada followed by a decline in

the 25-29 age group and increasing rates from 30-34 onward.[21] These results should be interpreted with the understanding that individuals can be included in multiple diagnosis groups. This could mean that a significant portion of the incidence of attempts and deaths after diagnosis could be driven by comorbid individuals who are being counted in multiple diagnosis categories.

Previous work has shown that the risk of repeated suicide attempt or death from suicide is high among those with previous attempts. This study found 9.1% of individuals admitted with suicide attempts were re-admitted due to a second attempt within 1 year. After the first detected attempt admission 0.46% of attempters died due to suicide within 1 year and 1.3% within 5 years. These incidence rates are high, but lower than incidence found in another study using register data.[25] Christiansen and Jensen[25] found repeat attempts in 31.3% of their sample and deaths in 2.33% despite mean follow up periods of only 2.88 and 3.88 years. That study used all presenting attempts between 1995 and 2001 for that sample. Therefore age and cohort effects may explain the large difference in risk between the two samples. Another study found that 25.1% of their admitted attempters had a repeat admission.[26] Research in Taiwan found rates that were close to this study.[27] A systematic review determined that the median repeat rate within 1 year was 16% (IQR= 12 to 25%).[29] Median rate of death from suicide within 1 year was 1.8% (IQR= 0.8 to 2.6%). This study's use of first lifetime occurrence of suicide attempt to form the sample may explain the lower rates. Other work that relies on hospital-based sampling, or the inclusion of all attempts within a small cross-section of years, will likely oversample individuals with chronic suicidal behaviours.

A major study strength is the population-based birth cohort. The occurrence of suicidal behaviours treated in inpatient wards can be counted for cohort members throughout their life. Therefore the measured incidence should be a less biased estimate of the true occurrence of

treated disorders, treated suicide attempts, and deaths. The use of administrative data for outcome measures ensures the capturing of all hospitalizations within the province. Prior research using self-report of adults to estimate cumulative incidence may not fully account for the occurrence of mortality and how excess mortality will affect suicidal individuals more than people with no history of suicidal behaviour. This excess mortality will cause people with history of suicidal behaviour to be less likely to be sampled in retrospective studies. Competing risks and censoring events were taken into account during the estimation of the cumulative incidence. Capturing events in the data as they occurred represents an improvement over using birth cohorts with uneven follow-ups.[30] This study also does not experience the significant attrition found in many previous birth cohort studies,[30] with 86.5% or better follow-up until the end of available data.

One limitation of this study is in not including suicide attempts that present to the emergency department but are not admitted or are not treated medically. This group of patients is also of interest, since the emergency department is a key location for interventions. This group was omitted due to the lack of injury intent codes from emergency department treatment in the province, and because no validated method of detecting attempts in the emergency department is available that does not use ICD injury intent codes. Moreover, although all the hospitalizations in the province will be captured in the data, ICD codes do not detect all occurrences of suicide attempts.[31] Of particular concern is that the codes used will not detect some cases, and therefore will underestimate the actual cumulative incidence.[31] Another limitation is that the codes used to detect an attempt do not attempt to separate self-injury based on the intent of the injury and therefore may contain various types of self-harm as well as suicide attempts.

The coding of mental illnesses may be subject to misclassification error in hospital and physician records[32] and this may bias the analysis. This misclassification is compounded by the potential for disorders to be initially misdiagnosed by the physician. There is no indicator in the records to indicate a provisional diagnosis or a prior misdiagnosis. Therefore the diagnostic categories likely contain some individuals that do not have the disorder of interest, but rather a disorder with similar symptoms. The diagnosis of personality disorders during adolescence is controversial; this may affect the diagnosis of these disorders in the administrative data.[33] The physician claims in the Repository cannot separate bipolar disorder from other mood disorders. The incidences associated with mental illnesses only show the occurrence of suicidal behaviours after a diagnosis. However, the incidence after diagnosis is clinically useful information. The actual onset of these disorders may have occurred well before the first diagnosis date. Moreover, individuals with these disorders who do not receive a diagnosis will be missed.

3.7 Conclusions

The incidence of admission for a suicide attempt after a first diagnosis with a mental disorder was between 1% (anxiety disorder) and 5.7% (personality disorder and schizophrenia) during the first 5 years. These conditions emerge early in life and, without successful treatment, many adolescents and young adults are at risk of engaging in serious suicidal behaviour with potential long-term consequences – including a high risk of death at a young age. The acuity of the risk for serious suicidal behaviour after initial diagnosis underscores the clinical importance of early interventions that target this population. Our study results underscore the need for policy initiatives to address risk factors for suicidal behaviour.

3.8 References

1 Hawton K, Saunders KE a, O'Connor RC. Self-harm and suicide in adolescents. *Lancet* 2012; **379**: 2373–82.

2 Bridge JA, Goldstein TR, Brent D a. Adolescent suicide and suicidal behavior. *J Child Psychol Psychiatry Allied Discip* 2006; **47**: 372–94.

3 Värnik A, Kõlves K, Allik J, Arensman E, Aromaa E, van Audenhove C, *et al.* Gender issues in suicide rates, trends and methods among youths aged 15-24 in 15 European countries. *J Affect Disord* 2009; **113**: 216–26.

4 Curtin SC, Warner M, Hedegaard H. Increase in suicide in the United States, 1999-2014. NCHS data brief, no 241. , 2016 (http://www.cdc.gov/nchs/products/databriefs/db241.htm).

5 Steele MM, Doey T. Suicidal behaviour in children and adolescents. part 1: etiology and risk factors. *Can J Psychiatry* 2007; **52**: 21S–33S.

6 American Academy of Child and Adolescent Psychiatry. Practice parameter for the assessment and treatment of children and adolescents with suicidal behavior. American Academy of Child and Adolescent Psychiatry. *J Am Acad Child Adolesc Psychiatry* 2001; **40**: 24S–51S.

7 Patten SB. Recall bias and major depression lifetime prevalence. *Soc Psychiatry Psychiatr Epidemiol* 2003; **38**: 290–6.

8 Patten SB, Williams JVA, Lavorato DH, Wang JL, McDonald K, Bulloch AGM. Descriptive epidemiology of major depressive disorder in Canada in 2012. *Can J Psychiatry* 2015; **60**: 23–30.

9 Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G. Recall of depressive episode 25 years previously. *Psychol Med* 1999; **29**: 787–91.

10 Canadian Institute for Health Information. ICD-9/CCP and ICD-9-CM. (https://www.cihi.ca/en/data-and-standards/standards/classification-and-coding/icd-9ccp-and-icd-9-cm).

11 Statistics Canada. Population of census metropolitan areas. 2016. (http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo05a-eng.htm).

12 Fransoo R, Martens P, Burland E, The Need to Know Team, Prior H, Burchill C. Manitoba RHA Indicators Atlas 2009. , 2009 (http://mchp-appserv.cpe.umanitoba.ca/reference/RHA_Atlas_Report.pdf).

13 Martens PJ, Burchill C, Fransoo R, De Coster C, McKeen N, Ekuma O, *et al.* Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-

Based Study. , 2004.

14 Lin G, So Y, Johnston G, Nc C. Analyzing Survival Data with Competing Risks Using SAS ® Software. *SAS Glob Forum* 2012; **344**: 1–8.

15 Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant* 2013; **28**: 2670–7.

16 SAS. The LIFETEST Procedure. 2016. (http://support.sas.com/documentation/cdl/en/statug/68162/HTML/default/viewer.htm#statug_lifetest_details14.htm).

17 Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc* 1999; **94**: 496–509.

18 SAS Institute Inc. SAS/ACCESS® 9.4. 2012. Cary, NC: SAS Insititute Inc.

19 Kann, L., Kinchen, S., Shanklin, S.L., Flint, K.H., Hawkins, J., Harris, W.A., Lowry, R., Olsen, E.O., McManus, T., Chyen, D., Whittle, L., Taylor, E. Demissie, Z., Brener, N., Thornton, J., Moore, J., Zaza S. Youth Risk Behavior Surveillance — United States, 2013. *Surveill Summ* 2014; **63**: 1–168.

20 Fergusson DM, Horwood LJ, Ridder EM, Beautrais AL. Suicidal behaviour in adolescence and subsequent mental health outcomes in young adulthood. *Psychol Med* 2005; **35**: 983–93.

21 Statistics Canada. Suicide and suicide rate, by sex and by age group (Both sexes rate). 2015. (http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/hlth66d-eng.htm).

22 Sullivan EM, Annest JL, Simon TR, Luo F, Dahlberg LL, Centers for Disease Control and Prevention (CDC). Suicide trends among persons aged 10-24 years--United States, 1994-2012. *MMWR Morb Mortal Wkly Rep* 2015; **64**: 201–5.

23 Dilillo D, Mauri S, Mantegazza C, Fabiano V, Mameli C, Zuccotti GV. Suicide in pediatrics: epidemiology, risk factors, warning signs and the role of the pediatrician in detecting them. *Ital J Pediatr* 2015; **41**: 49.

24 Randall JR, Walld R, Finlayson G, Sareen J, Martens PJ, Bolton JM. Acute risk of suicide and suicide attempts associated with recent diagnosis of mental disorders: a population-based, propensity score-matched analysis. *Can J Psychiatry* 2014; **59**: 531–8.

25 Christiansen E, Jensen BF. Risk of repetition of suicide attempt, suicide or all deaths after an episode of attempted suicide: a register-based survival analysis. *Aust N Z J Psychiatry* 2007; **41**: 257–65.

26 O'Connor RC, Smyth R, Williams JMG. Intrapersonal positive future thinking predicts repeat suicide attempts in hospital-treated suicide attempters. *J Consult Clin Psychol* 2015; **83**: 169–76.

27 Chen C-Y, Yeh H-H, Huang N, Lin Y-C. Socioeconomic and clinical characteristics associated with repeat suicide attempts among young people. *J Adolesc Health* 2014; **54**: 550–7.

28 Owens D, Horrocks J, House A. Fatal and non-fatal repetition of self-harm. Systematic review. *Br J Psychiatry* 2002; **181**: 193–9.

29 Owens D, Wood C, Greenwood DC, Hughes T, Dennis M. Mortality and suicide after non-fatal self-poisoning: 16-year outcome study. *Br J Psychiatry* 2005; **187**: 470–5.

30 Power C, Kuh D, Morton S. From Developmental Origins of Adult Disease to Life Course Research on Adult Disease and Aging: Insights from Birth Cohort Studies. *Annu Rev Public Health* 2013; **34**: 7–28.

31 Randall JR, Roos LL, Lix LM, Katz LY, Bolton JM. Emergency department and inpatient coding for self-harm and suicide attempts: valdiation using clinician assessment data. *Int J Methods Psychiatr Res* 2017. doi:10.1002/mpr.1559.

32 O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. *Health Serv Res* 2005; **40**: 1620–39.

33 Gask L, Evans M, Kessler D. Personality disorder. *BMJ* 2013; **347**: f5276–f5276.

Table 3-1: ICD codes and number of visits used to determine diagnoses of mental disorder

Anxiety	Hospital records (1+ visits): ICD–9–CM: 300.0, 300.2, 300.3; ICD-10–CA: F40, F41.0, F41.1,F41.3, F41.8, F41.9, F42 Physician claims (3+ visits): ICD–9–CM: 300
Mood and adjustment disorders	Hospital records(1+ visits): ICD–9–CM: 296.2–296.8, 300.4, 309, 311; ICD–10–CA: F31, F32, F33, F34.1, F38.0, F38.1, F41.2, F43.1, F43.2, F43.8, F53.0, F93.0 Physician claims (1+ visits): ICD–9–CM: 296, 309, 311
Personality disorders	Hospital records or physician claims (1+ visits): ICD-9-CM: 301; ICD-10-CA: F34.0, F60-F62, F68.1,F68.8, F69
Substance use	Hospital records or physician claims (1+ visits): ICD–9–CM: 291, 292, 303, 304, 305; ICD–10–CA: F10–F19, F55
Schizophrenia	Hospital records or physician claims (1+ visits): ICD–9–CM: 295; ICD–10–CA: F20, F21, F23.2, F25
Suicide death/attempt	Hospital records or vital statistics records(1+ visits): ICD-9-CM: E950-E959, E980-E989; ICD-10-CA: X60-X84, Y10-Y34

Number of visits indicates how many physician claims, or inpatient records with the diagnosis are required before classifying an individual as a case

Table 3-2: Birth cohort characteristics and follow-up

	N	(%)*	Outcome-specific Follow up (%)	
			Until Age 25	Until end of study
Total births	236809	(--)		
Study cohort (age 10)	92890	(100)		
Suicide attempt	875	(0.9)	59.8	86.6
Suicide death	105	(0.1)	59.5	86.5
Mental disorder diagnosis**				
Anxiety	16176	(17.4)	68.1	92.8
Mood and adjustment	19736	(21.2)	66.2	90.9
Personality	2014	(2.2)	69.4	90.1
Schizophrenia	795	(0.9)	68.3	91.7
Substance use	6981	(7.5)	66.8	89.5

* % of study cohort

** physician diagnosis by age 25; follow up calculated using suicide death outcomes by disorder
Note: some members of the cohort were younger than 25 at the end of data, therefore attrition rates are shown for age 25 and for attrition before the end of available data.

Table 3-3: Cumulative incidence of suicide deaths per 100,000 population

		Age 15	95%CI	Age 20	95%CI	Age 25	95%CI	P-Value
Full study cohort		8.9	4.3 to 17	60	46 to 78	129	106 to 156	N/A
Birth cohort	1979-1986	8.5	3.0 to 22	50	33 to 74	127	97 to 164	0.77
	1987-1993	9.3	3.3 to 24	71	49 to 100	*	*	
Sex	Male	8.7	3.1 to 22	71	50 to 102	156	121 to 199	0.0378
	Female	9.1	3.2 to 23	48	31 to 734	101	74 to 137	
		1-year**	95%CI	5-year**	95%CI	10-year**	95%CI	
Mental disorder diagnosis								
	Anxiety	30	12 to 70	132	81 to 223	318	201 to 488	
	Mood and adjustment	50	26 to 90	177	122 to 252	491	354 to 668	
	Personality	190	60 to 480	519	256 to 967	1005	492 to 1870	
	Schizophrenia	*	*	575	158 to 610	2103	552 to 5678	
	Substance use	150	80 to 270	500	333 to 728	967	668 to 1361	

* estimates unreliable due to small number of events

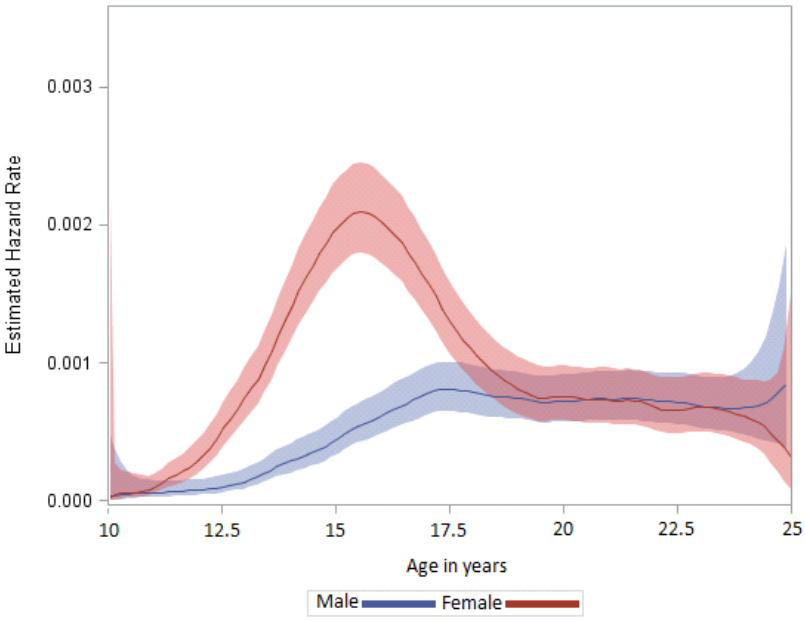
** years from initial diagnosis

Table 3-4: Cumulative incidence of suicide attempt per 10,000 population

		Age 15	95%CI	Age 20	95%CI	Age 25	95%CI	P-Value
Full study cohort		20	17 to 23	71	65 to 76	103	97 to 111	N/A
Birth cohort	1979-1986	24	20 to 29	81	73 to 89	114	104 to 124	0.0011
	1987-1993	15	12 to 19	59	52 to 67	92	83 to 103	
Sex	Male	7.2	5 to 10	42	36 to 48	76	68 to 84	<0.0001
	Female	33	28 to 38	101	92 to 110	133	121 to 144	
		1-year*	95%CI	5-year*	95%CI	10-year*	95%CI	
Mental disorder diagnosis								
	Anxiety	29	22 to 38	98	83 to 116	203	172 to 238	
	Mood and adjustment	82	70 to 95	213	192 to 236	389	350 to 432	
	Personality	271	207 to 348	573	470 to 688	838	684 to 1010	
	Schizophrenia	271	175 to 400	568	406 to 767	961	575 to 1470	
	Substance use	104	82 to 130	302	259 to 349	643	527 to 773	

*years from initial diagnosis

Figure 3-1: Epanechnikov kernel-smoother hazard functions for first suicide attempt admission with 95% confidence limits



CHAPTER 4: EMERGENCY DEPARTMENT AND INPATIENT CODING FOR SELF-HARM AND SUICIDE ATTEMPTS: VALIDATION USING CLINICIAN ASSESSMENT DATA

4.1 Chapter overview

The previous two chapters examined the epidemiology of diagnosed mental illness and hospitalized suicide attempts. Administrative data were the source of information on diagnosis and suicide attempt-related hospitalization. Some health researchers have questioned how well these data can identify those with mental illnesses and suicide attempts. Although some validation work has been done in this area more research is needed to guide use of data, particularly for suicide attempts where validation research is sparse and heavily focused in the United States.

This chapter assesses the agreement between hospital coded cause of injury codes and psychiatrist determined suicide attempt and self-harm presentation status from a clinical dataset. In addition to assessing agreement between these two methods of determining the suicidality of presentation, this chapter will use the psychiatrist determined presentation status as a gold standard to determine the sensitivity, specificity, and negative/positive predictive values for hospital coding of self-harm and suicide attempts.

Publication details:

Randall JR, Roo LL, Lix LM, Katz LY, Bolton JM (2017). Emergency department and inpatient coding for self-harm and suicide attempts: validation using clinician assessment data. *International Journal of Methods in Psychiatric Research.*

4.2 Abstract

Objective: Administrative data have been used to determine the occurrence of suicide attempts and deliberate self-harm, but research about the accuracy of these sources is limited. This study aimed to assess the accuracy of inpatient coding for individuals that were assessed by psychiatrist for suicidality during an emergency department visit that resulted in hospitalization.

Methods: This study used a clinical sample (n=5719) containing psychiatry consultations from the emergency departments and inpatient units of the two major tertiary hospitals in Winnipeg, Canada to validate the accuracy of inpatient hospital diagnosis codes at identifying presentations for self-harm and suicide attempts. The Columbia Classification Algorithm of Suicide Assessment (C-CASA) was used as the gold standard. Administrative data from hospital discharge abstracts provided International Classification of Diseases version 10-CA codes for intentional self-harm, undetermined intent injury, and accidental poisoning were assessed. Measures of validity included Kappa (*K*), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Results: Sensitivity of hospitalized attempts was low using intentional intent codes (36.9%, 95% confidence interval [CI]: 32.4% to 41.4%) but improved using unknown intent and accidental poisoning codes (44.8%, 95%CI: 40.2% to 49.4%). Agreement for suicide attempts did not increase with the addition of unknown intent and accidental poisoning codes (*K* = 0.465 to 0.481), but were better for any self-harm (*K* = 0.395 to 0.478).

Conclusion: Hospital diagnosis codes undercount attempts and self-harm admissions. Including more data sources might improve the detection of events.

4.3 Background

Suicidal behaviour has been the subject of a substantial amount of research. Many studies of suicidal behaviour have focused on assessing potential risk factors.[1–5] Understanding the factors leading to suicidal behaviour is expected to improve treatment for patients and to prevent the occurrence of the significant morbidity and mortality associated with suicidal behaviour.[6] In the last few decades research has increasingly used administrative databases and registries to provide information on the occurrence of risk factors and suicide attempts and deaths.[7–15] These registries contain records of the use of health services by individuals and allow population-based work with significant power to be conducted. However, administrative data are not primarily intended to be used for research and require validation of their contents and appropriateness for use as research variables.[16,17] Prior research has examined the accuracy of the coding in Canadian hospitals for mental illnesses and found that accuracy was acceptable.[18–21]

The validity of the occurrence of suicide attempts and deaths derived from these data is important to the interpretation of work relying on this information. Research assessing methods of identifying suicide attempts in administrative data is insufficient and a recent systematic review determined that more research is needed.[22] This review identified only six small sample studies and none of which used the newest version of the International Classification of Disease coding system, version number 10 (ICD-10). These studies also chiefly relied on chart reviews to establish a gold standard; such work relies on proper documentation in the charts. Larger samples with better gold standards are required to understand the validity of diagnostic coding schemes for these outcomes.

Some prior research has suggested that many suicide attempts are not being detected in administrative data.[22] Attempts that are never treated medically will not be detected using these

data. However, even when treated in the emergency department or an inpatient unit, a substantial portion of suicide attempts might be missed. Analysis of Canadian emergency department data found that the number of individuals recorded as having an injury of undetermined intent is a noticeable portion of presentations and almost as many presentations as were classified as intentional self-injury.[23] This means that relying solely on individuals with intentional self-injury codes and ignoring those coded with injuries of undetermined intent might lead to an undercounting of self-harm in administrative data.[23,24]

This can be due to several factors including physicians listing underlying mental illnesses as the diagnosis and not properly charting the occurrence of suicidal behaviour, and gaps in administrative data (such as limited information on ED triage presenting complaint). Using different combinations of ICD codes can improve the sensitivity of detecting suicide attempts but may cause a high rate of false positives. Properly understanding the accuracy of identifying suicide attempts with these codes is crucial to future work. Currently there is no evidence-based consensus on the best way to code for suicide attempts and deliberate self-harm using ICD-10 diagnostic codes. Research on the relationship among ICD-10 codes for external cause of injury, suicide attempts and deliberate self-harm is needed to guide researchers and understand the strengths and limitations of studies using these outcomes.

This paper examined the validity of using ICD-10 diagnostic codes in hospital discharge records to identify suicide attempts and deliberate self-harm presentations admitted to inpatient units. It used clinician assessment of the occurrence of self-injury as a gold standard to determine whether an admission was a suicide attempt and/or deliberate self-harm (including both suicidal and non-suicidal self-injury). Diagnostic codes of interest included those for intentional self-harm, self-harm of undetermined intent, and accidental poisonings derived from hospital

discharge data. These analyses were focused on admitted patients due to ICD-10 codes not being recorded for patients discharged from the ED. This study also examined the triage complaints from the emergency department records and their occurrence among those presenting with suicide attempts, self-harm, or with no self-harm. We hypothesized that the intentional self-harm ICD codes would have high specificity but modest sensitivity, and that including ICD codes for self-harm of undetermined intent and accidental poisoning codes in the definition would improve the sensitivity slightly while reducing the specificity. We also hypothesized that triage complaints would not be able to accurately identify those with suicide attempts and self-harm.

4.4 Methods

4.4.1 Sample

The study sample was obtained from emergency departments of the two tertiary care hospitals in the city of Winnipeg, Manitoba, Canada. Due to the single-payer health care system in the province, these hospitals are accessible to all residents of the province and are not limited to individuals based on insurance coverage or other factors. Those hospitals have fulltime psychiatrists and psychiatry residents. As part of their assessment, the psychiatric staff and residents servicing these emergency departments are required to fill out the Suicide Assessment Form in Emergency (SAFE). Data from all psychiatric consultations were collected between January 1, 2009 and December 31, 2012. These consultations consisted of all consecutive adult patients referred for psychiatry consultation after presentation to these emergency departments. The first visit for each individual in the specified time period was selected to form the analysis sample.

4.4.2 Clinical assessment for suicide attempt and self-harm

SAFE contains several risk factor assessments as well as information on date of presentation and the Columbia Classification Algorithm of Suicide Assessment (C-CASA). The C-CASA is a standardized tool that classifies suicidal and self-harm behaviours, ideation and planning, sorting individuals into mutually exclusive categories.[25] The C-CASA classification served as the gold standard in this study. The following C-CASA categories were used: suicide attempts, self-injurious behaviour (no suicidal intent), self-injurious behaviour (intent unknown), suicidal ideation, and preparatory acts toward imminent suicidal behaviour. Self-harm presentations included any presentation assigned to the following C-CASA categories: suicide attempts, self-injurious behaviour (no suicidal intent), self-injurious behaviour (intent unknown). Classification is based on the current presentation, and past self-harming thoughts and behaviours are not included in this assessment.

4.4.3 Administrative data linkage

Records from the SAFE data set were linked to administrative records housed at the Manitoba Centre for Health Policy (the Repository) using patients' Provincial Health Insurance Numbers collected during the assessment. The Repository contains information on medical services provided by the province of Manitoba's single-payer health care system. The Repository contains individual-level data on over 99% of the province's population. Two databases from the Repository were used: the Emergency Department Information System (EDIS), and the Hospital Discharge Abstract Database. EDIS is used in both of the hospitals to track and record visits by individuals to the emergency department. EDIS captures the triage presenting complaint; relevant categories of complaints include 'Mental health', 'Substance use', and 'Trauma'. More specific categorizations within these categories include 'Depression, suicidal, self-harm',

'Overdose', and 'Laceration'. Moreover, supplementary information can be used to identify the following: 'Attempted suicide, clear plan', 'Active thoughts', 'High risk/unknown substance, and 'Altered Level of Consciousness'. Only the chief complaint is recorded in EDIS, therefore patients presenting with issues other than suicide attempts might have those issues listed as the chief complaint.

The Hospital Discharge Abstract Database identified individuals admitted to inpatient units following emergency department presentation and was the source of diagnosis codes for validation. This database contains records for all admissions in the province and captures up to 25 diagnosis codes for each inpatient admission. The diagnoses are coded using International Statistical Classification of Diseases and Related Health Problems version 10 Canadian Enhancement (ICD-10-CA; CIHI 2001). Coding was performed by trained medical coders, coding of disorders is thorough and comprehensive based on the documentation of clinicians. The following codes were used to determine suicide attempts: Intentional self-harm (ICD-10-CA codes X60-X84), injury of undetermined intent (Y10-Y34), and accidental poisoning (ICD-10-CA codes X40-49). These codes were selected because of their use in previous validation and epidemiology research.[22,27] All 25 diagnoses were scanned to detect the occurrence of these codes. These codes are not available for patients discharged from the ED.

No variable recorded in the SAFE dataset can link individual SAFE records to specific emergency department and inpatient visit records held in the Repository; only date and personal health insurance numbers are available. Therefore, the administrative records were grouped together based on dates. Emergency department visits were identified and records on the same date or subsequent dates were combined into one record, or event. A space of two days without any administrative record was used as an indication that an event had ended. Hospital records

provide both admission and discharge dates for this purpose. All analysis is based on these grouped records, which will be referred to as 'events' to differentiate them from 'emergency department presentations'. Since multiple emergency department visits are possible within the same event, a specific event may have multiple triage codes. Between-facility transfers may lead to one event being linked to multiple hospital records. All hospital discharge records were scanned for the diagnostic codes of interest.

4.4.4 Statistical Analysis

Frequencies were calculated for individuals based on their EDIS complaint categorization and whether they were coded as a suicide attempt or undetermined intent self-harm in hospital records. The distribution of these classifications was assessed for those with a suicide attempt and self-harm. The agreement of the ICD-10-CA codes from administrative records and suicide attempts/self-harm determined by C-CASA was assessed using the kappa (K) statistic.[28] Kappa was interpreted using the following groups; K <0 – less than chance, 0.01 -0.20 – slight agreement, 0.21-0.4 – fair agreement, 0.41-0.60 – moderate agreement, 0.61-0.80 – Substantial agreement, >0.81 – Almost perfect agreement.[28,29] These heuristics should be used with caution since one number statistics provide a limited assessment of validity and usefulness and the kappa statistic weighs sensitivity and specificity equally. Three definitions of ICD codes were used – X60-X84; X60-X84 and Y10-Y34; and X60-X84, Y10-Y34 and X40-49. Frequencies comparing C-CASA classification with these codes were derived twice – once for all of the emergency department events and again with the sample restricted to events with inpatient admissions. Further validity statistics are determined for admitted patients since ICD-10 codes are only available for admitted patients. Predictive statistics and their 95% confidence intervals (CIs) were derived, including sensitivity, specificity, positive predictive value (PPV), and

negative predictive value (NPV). Frequencies and standardized differences for the five most relevant EDIS complaints among admitted patients were compared between the C-CASA determined self-harm/attempt patients and those detected by the ICD coding algorithms to assess for potential bias from using ICD codes.

3.4.5 Ethical approval

Ethics approval was obtained from the University of Manitoba's Health Research Ethics Board (HREB #: H2015:009) and data access was granted by the Government of Manitoba's Health Information Privacy Commission (HIPC #- 2014/2015 – 41).

4.5 Results

The sample consisted of 10150 assessment records in the SAFE dataset, of which 9319 were linkable to the Repository. A total of 6025 index visits were identified, 5719 of these had C-CASA assessment information in the SAFE data and comprised the study sample (**Table 4-1**). Among the linked presentations, 3013 were male (52.7%) with a mean age of 41.4 years (median = 39; SD = 17.4). 780 presentations were classified as suicide attempts according to the C-CASA classification. 1147 presentations were due to self-harm (20.1%; including those previously assessed as attempts). Visits categorized as having imminent plans consisted of an additional 190 (3.3%), while suicidal ideation was the classification for 1495 (26.1%).

Table 4-2 contains the frequencies of individuals classified into the four C-CASA categories according to the inpatient diagnosis records (X60-X84). The estimate of kappa for suicide attempt categorization for all emergency department visits (both admitted and non-admitted visits) was K = 0.295 (95%CI= 0.252 to 0.339), which is considered fair agreement.[28] Agreement between self-harm and the hospital records was K = 0.229 (95%CI= 0.187 to 0.270, fair agreement). Restricting the sample to the admitted visits improved K to 0.465 (95%CI=

0.414 to 0.516) and 0.393 (95%CI= 0.347 to 0.443) respectively; these figures translate into moderate and fair agreement. Using the broader suicide attempt definition including unknown intent codes led to better agreement in all cases, but particularly improved the accuracy of identifying self-harm. The value of *K* after including unknown intent codes for predicting self-harm admissions was 0.434 (95%CI= 0.388 to 0.479). Suicide attempt admissions had *K* = 0.482 (95%CI= 0.433 to 0.531). Further adding the Accidental Poisoning codes had little impact on the kappa for suicide attempts (0.481, 95%CI= 0.434 to 0.528) but improved *K* to 0.478 (95%CI= 0.435 to 0.521) for self-harm. Frequencies for the two expanded definitions are displayed in **Table 4-3**.

Using the C-CASA assessment as the gold standard, the predictive ability of the administrative data coding for the admitted patients is shown in **Table 4-4**. The sensitivity of the diagnostic codes was poor. The X60-X84 codes identified only 36.9% (95%CI= 32.4% to 41.4%) of those admitted with an attempt, that is 63.1% of admitted suicide attempts were not detected using the intentional self-harm codes. Adding the Y10-Y34 codes marginally improved the sensitivity to 41.0% (95%CI= 36.4% to 45.6%). The sensitivity for capturing self-harm was lower at 29.7% (95%CI= 26.2% to 33.3%), with sensitivity increasing to 34.2% (95%CI= 30.5% to 38.0%) with the Y10-Y34 codes. The PPV and specificity were fairly good, however. These statistics were slightly higher for the more general self-harm group than for the suicide attempt group. Adding accidental poisoning increased the sensitivity for self-harm admissions to 40.0% (95%CI= 36.2% to 43.9%). The poisoning codes only slightly increased the sensitivity for suicide attempt admissions (43.6%, 95%CI= 40.2% to 49.4%) while decreasing the PPV to 66.8% (95%CI= 61.4% to 72.1%). In all instances the majority of people that were determined to have made a suicide attempt or engaged in self-harm by the clinician were not detected using all

of these codes. Since non-admitted patients do not receive ICD-10-CA codes they will be missed entirely. Sensitivity for all events was 21% (95%CI= 18.2% to 23.9%) for the X60-X84 codes, increasing to 23.3% and 25.5% when the Y10-Y34 codes and X40-X49 were added. Without further data on emergency department treatment, a solid majority of emergency department events were not detected.

Table 4-5 presents the occurrence of the various EDIS triage complaint categorization by suicide attempt and self-harm classification. Mental health was the most common primary classification with 48.7% of the self-harm presentations in this group compared to 73.7% of the people with no self-harm. Of these presentations, 83.5% were classified under 'Depression, suicidal, self-harm'. Approximately a third of the self-harmers were categorized under the 'substance use' and 'overdose' presentation categories. There were 195 (25.0%) suicides attempts and 233 (20.3%) self-harmers with the 'Attempted suicide, clear plan' classification, but 162 (3.5%) non-self-harmers were also given this classification. **Table 4-6** shows the occurrence of the 5 most relevant EDIS complaints among admitted patients and how common they were between individuals classified as self-harm or attempt patients by C-CASA, and the three ICD coding methods. A downward bias was apparent for three of the complaints; 'Mental Health', 'Depression/suicidal/self-harm', and 'Attempt suicide, clear plan'. 'Mental health' and 'Depression/suicidal/self-harm' was approximately half as common among the group detected through ICD codes than expected based on the frequency in the C-CASA attempt group (ratios between 0.6 and 0.49 relative to C-CASA attempt group frequency). These are equal to standardized differences of between 0.37 and 0.49. The bias for these three complaints was least when using only X60-X84 codes. Overdose triage complaints were overrepresented in those detected by ICD codes by 1.09 to 1.23 times the expected frequency compared to the CASA

attempt group. In this instance the 'X60-X84, Y10-Y34' coding was the least biased. The standardized differences for overdose are between 0.06 and 0.15. The 'Attempted suicide, clear plan' was 0.89 to 0.7 times less common among those identified with the ICD codes compared to the C-CASA attempt group. The standardized differences between these groups were between 0.06 and 0.18.

4.6 Discussion

Our findings suggest that the use of ICD-10 codes in administrative data for detecting hospital-treated suicide attempts and self-harm is an imperfect method of detecting these outcomes. The main issue is the poor sensitivity of the codes. Even among those admitted with a suicide attempt, less than half of the sample was identified from the intentional self-harm codes. This finding has substantial impact for population-based administrative data research relying on suicide attempts as an outcome measure. The PPV of the codes is better, but still not perfect. Adding the undetermined intent codes improved the sensitivity, although at the expense of a reduced PPV. When the codes are used to detect suicide attempts, approximately half of the false positives are due to other instances of self-harm. This study suggests that at least 75% of suicide attempts presenting to the emergency department will be missed if ICD-10 codes from hospital discharge records are used. Even the majority of individuals admitted to the hospital are missed if ICD-10 codes are the only data used to detect attempts.

A recent systematic review of validation studies on using ICD codes to detect suicide attempts treated at hospitals determined that research in this area is lacking.[22] No studies examining ICD-10 coding were located, as all previous validation studies were based on ICD-9. A current literature search did not locate any additional relevant studies. Using presenting complaints in systems like EDIS may prove helpful in improving the accuracy of detecting self-

harm using ICD codes. However, the EDIS third level category for 'attempt suicide, clear plan' appears to capture very few of those with self-harm and a large number of non-self-harming individuals. This field is also optional and is often left blank. Conversely, two of the secondary categories (Overdose and Depression/suicidal/self-harm) both managed to identify a large number of those with self-harm. The overdose category also had a good PPV in the sample, though this is likely to decrease in a sample containing all emergency department visits. Using overdose presentation in conjunction with mental health ICD codes or with psychiatric treatment/assessment tariffs may prove effective in detecting a sizeable portion of the self-harm presentations. A limitation of triage systems similar to EDIS is the reliance of a single presenting complaint. This means that other comorbid conditions may be triaged as the main complaint instead of self-harm, and this will occur more often among those with less medically serious self-harm. This may explain why some individuals do not appear to have a relevant triage complaint.

The PPV estimates found in this study are similar to the previous results for intentional self-harm derived from ICD-9 codes E950-E959.[30,31] The PPV of the codes were acceptable, especially for self-harm without specifying intent. About half of the false positives that occur when deriving suicide attempt outcomes appear to be due to other forms of self-harm. This will hinder research aiming to use suicide attempts, specifically, as the outcome. However, this should not be a serious impediment, since individuals that attempt suicide and individuals that self-injure share similarities across many risk factors and non-suicidal self-injurers are the minority of those detected by the codes.[32] These data also suggest that these codes are specific enough to determine that people missed will be largely true negatives. It is likely that the NPV from a sample of all hospital admissions would be higher than the estimates here.

Few studies have examined the sensitivity of ICD codes. The work validating suicide attempt coding in the systematic review did not determine the sensitivity of the codes at detecting hospital-treated suicide attempts, only the sensitivity of the codes at detecting suicide deaths.[22] Depending on the importance of PPV or sensitivity in detection of suicide attempts, researchers can opt to restrict to X60-X84 codes or use the expanded definition including Y10-Y34 and X40-X49. For deriving self-harm, the use of the expanded definition appears to be superior.

The ability to separate non-suicidal self-harm from suicidal self-harm is a potential challenge when using administrative data. Some researchers have opted to label events derived from these data as being self-harm events whereas others have persisted in describing these events as suicidal. This study suggests that using expanded criteria (including Y10-Y34) in order to detect self-harm outcomes is likely the most accurate method of coding. The accuracy of detecting suicide attempts specifically is similar when using this expanded coding as well. Therefore using these codes to detect suicide attempts is feasible. However, restricting the definition to intentional codes (X60-X84) would reduce the occurrence of false positives and is advisable when PPV is more valued than sensitivity. This is potentially the case when conducting risk factor or causal studies and this is supported by the smaller bias in EDIS triage complaints found when using only the X60-X84 codes. Including the unknown intent codes would be superior when sensitivity is the primary concern. Specifically, this would be advisable when estimating incidences of suicide attempts. Expanding the codes used in detection would produce numbers more closely approximating the actual incidence of self-harm and suicide attempts. Including accidental poisoning improved the sensitivity and kappa agreement, but in practice the opposite might occur as non-mental health related accidental poisonings will be

identified. Adding an additional requirement of at least one mental health diagnostic code or treatment code before including accidental poisonings would potentially address this issue.

These results have implications for the use of administrative data to track the prevalence of suicide attempts and as the data source for suicide outcomes for causal studies. The low sensitivity indicates that there is likely a significant undercounting of the number of admitted patients estimated using these codes. The implication is less clear for causal studies. Missing cases are likely to bias estimates towards the null if the missing cases are similar to the identified individuals. However, if the identified individuals are not similar to the missed cases then other biases could also occur. It is possible that more severe cases would be detected disproportionately and this could cause risk factors to be differentially biased away from or towards the null, depending on whether they are associated more often with serious or less serious attempts. Overdoses presentations appear to be overrepresented according to the bias assessment potentially due to these reasons.

These results should be generalizable to other similar acute care settings. Institutions with comprehensive record keeping and trained medical coders deriving ICD codes should produce similar results. Areas with standardized psycho-social assessment of suicidality included in the chart should have increased validity using these codes. Hospitals without psychiatric assessment available 24 hours a day (e.g., rural hospitals) might perform poorer due to less thorough psycho-social assessment and documentation. Areas without comprehensive coding of charts will likely also perform poorly and poor assessment and coding will likely result in more false negatives particularly.

The main reason for errors in coding is probably poor documentation in the charts. Individuals miscoded with Y10-Y34 codes are a good example as these are likely individuals

with injuries and overdoses listed in the chart, but no clear documentation of intention. Medically minor self-harm might be treated in the emergency department, or not require any treatment at all, while the inpatient records might fail to mention self-harm due to a focus on what the treating physician determined to be the main issue (e.g., depression). The coding of charts is performed by trained medical coders who are taught to code diagnoses comprehensively; they should correctly code self-harm and attempts that are clearly documented in the charts.

This study used consecutive sampling of patients receiving psychiatric consultations in the emergency department. The C-CASA assessment of treating psychiatrists/psychiatric residents was used to determine whether specific visits were attempts, any form of self-harm, or not. Previous validation efforts have normally used chart reviews to determine whether visits were suicidal in nature. However, such reviews allow for a considerable degree of interpretation and assume that charting contains all of the relevant information. Classification of the presentations using the C-CASA completed by the clinician who performed the psychiatric assessment is likely to be more accurate than chart reviews. This work also used a sample considerably larger than previous studies.

A study limitation was the inability to assess the accuracy of determining suicide attempts using emergency department ICD codes; external cause of injury codes specifying intent are not recorded for visits to emergency departments in Manitoba. Another limitation is that approximately 13% of the sample either was not linkable to specific emergency department visits in the administrative records or was missing the C-CASA assessment. This was mostly due to missing dates or missing health insurance numbers in the SAFE data. While this information appears to be missing randomly, the possibility of bias cannot be entirely dismissed. However, due to the relatively small percent missing, the bias is unlikely to be large enough to substantially

affect the statistics obtained. Another limitation is this study's restriction to individuals assessed by the psychiatry consult service. Without assessment by psychiatry there is no C-CASA assessment to use as a gold standard. Individuals not consulted to psychiatry may have engaged in less serious self-harming behaviour or to have less severe comorbid conditions at presentation. The psychiatry consult service also only assesses adult patients. Therefore this study was limited to assessing the accuracy of these codes for those 18 years of age and older. Another limitation is that these results reflect the accuracy of coding done an inpatient setting and may not be generalizable to other psychiatric or outpatient settings. Other settings are likely to have less thorough coding and documentation and the performance of these codes could potentially be worse. This is likely to be reflected with even lower sensitivity for suicide attempts or self-harm, and the possibility of false positives exists but is likely less common. Generalization of these results outside of inpatient settings is not recommended. Another limitation is the use of linkage by date to match the SAFE data set and the administrative data records. This may cause some misclassification if records are linked in that are not related to the SAFE assessment presentation. This method does combine a small number of individuals with multiple ED presentations (about 1% of presentations). However this misclassification is not likely to affect the statistics since it occurs infrequently.

4.7 Conclusion

This study assessed the accuracy of ICD-10 diagnosis codes at detecting patients admitted due to a suicide attempt or self-harm from among all patients consulted to psychiatry from the emergency department and non-psychiatric units. Relevant individuals were likely to be accurately identified by the codes, but the sensitivity of the codes is low. These findings suggest that research using these codes to identify suicidal behaviour outcomes is likely missing in the

vicinity of one half to two thirds of their outcomes. This is likely to cause significant undercounting of prevalence and incidence and may bias the results of causal studies. Future research should continue in this area. Examining the potential usefulness of non-ICD-based data sources, such as EDIS, to derive detection algorithms is a potential area that should be examined.

4.8 References

1 Johnson GR, Krug EG, Potter LB. Suicide among adolescents and young adults: a cross-national comparison of 34 countries. *Suicide Life Threat Behav* 2000; **30**: 74–82.

2 Johnson JG, Cohen P, Gould MS, Kasen S, Brown J, Brook JS. Childhood adversities, interpersonal difficulties, and risk for suicide attempts during late adolescence and early adulthood. *Arch Gen Psychiatry* 2002; **59**: 741–9.

3 Hawton K, van Heeringen K. Suicide. *Lancet* 2009; **373**: 1372–81.

4 Hawton K, Saunders KE a, O'Connor RC. Self-harm and suicide in adolescents. *Lancet* 2012; **379**: 2373–82.

5 Turecki G, Brent DA. Suicide and suicidal behaviour. *Lancet (London, England)* 2016; **387**: 1227–39.

6 Christiansen E, Larsen KJ, Agerbo E, Bilenberg N, Stenager E. Risk factors and study designs used in research of youths' suicide behaviour—An epidemiological discussion with focus on level of evidence. *Nord J Psychiatry* 2014; **68**: 513–23.

7 Randall JR, Walld R, Finlayson G, Sareen J, Martens PJ, Bolton JM. Acute risk of suicide and suicide attempts associated with recent diagnosis of mental disorders: a population-based, propensity score-matched analysis. *Can J Psychiatry* 2014; **59**: 531–8.

8 Bolton JM, Walld R, Chateau D, Finlayson G, Sareen J. Risk of suicide and suicide attempts associated with physical disorders: a population-based, balancing score-matched analysis. *Psychol Med* 2015; **45**: 495–504.

9 Morrison KB, Laing L. Adults' use of health services before death by suicide in Alberta. , 2011 (http://www.statcan.gc.ca/pub/82-003-x/2011003/article/11516-eng.pdf).

10 Tran T, Luo W, Phung D, Harvey R, Berk M, Kennedy RL, *et al.* Risk stratification using data from electronic medical records better predicts suicide risks than clinician assessments. *BMC Psychiatry* 2014; **14**: 1–9.

11 Spirito A, Esposito-Smythers C. Attempted and completed suicide in adolescence. *Annu Rev Clin Psychol* 2006; **2**: 237–66.

12 Christiansen E, Stenager E. Risk for attempted suicide in children and youths after contact with somatic hospitals: a Danish register based nested case-control study. *J Epidemiol Community Health* 2012; **66**: 247–53.

13 Spittal MJ, Pirkis J, Miller M, Carter G, Studdert DM. The Repeated Episodes of Self-Harm (RESH) score: A tool for predicting risk of future episodes of self-harm by hospital patients. *J Affect Disord* 2014; **161**: 36–42.

14 Nordentoft M, Mortensen PB, Pedersen CB. Absolute risk of suicide after first hospital

contact in mental disorder. *Arch Gen Psychiatry* 2011; **68**: 1058–64.

15 Katz LY, Au W, Singal D, Brownell M, Roos N, Martens PJ, *et al.* Suicide and suicide attempts in children and adolescents in the child welfare system. *CMAJ* 2011; **183**: 1977–81.

16 Roos LL, Gupta S, Soodeen R-A, Jebamani L. Data quality in an information-rich environment: Canada as an example. *Can J Aging* 2005; **24 Suppl 1**: 153–70.

17 Lix LM, Smith M, Levy A, Dai S, Sanmartin C, Quan H. *Quality of Administrative Data in Canada: A Discussion Paper*. Population Health Data Laboratory, 2012.

18 Juurlink D, Preyra C, Croxford R, Chong A, Austin P, Tu J, *et al.* Canadian Institute for Health Information Discharge Abstract Database: A Validation Study ICES Investigative Report. *Toronto Inst Clin Eval Sci* 2006; : 69.

19 Canadian Institute for Health Information. CIHI Data Quality Study of Ontario Emergency Department Visits for Fiscal Year 2004-2005: Volume II of IV - Main Study Findings. , 2007.

20 Steele LS, Glazier RH, Lin E, Evans M. Using administrative data to measure ambulatory mental health service provision in primary care. *Med Care* 2004; **42**: 960–5.

21 Kurdyak P, Lin E, Green D, Vigod S. Validation of a Population-Based Algorithm to Detect Chronic Psychotic Illness. *Can J Psychiatry* 2015; **60**: 362–8.

22 Walkup JT, Townsend L, Crystal S, Olfson M. A systematic review of validated methods for identifying suicide or suicidal ideation using administrative or claims data. *Pharmacoepidemiol Drug Saf* 2012; **21 Suppl 1**: 174–82.

23 Bethell J, Rhodes AE. Identifying deliberate self-harm in emergency department data. *Heal reports* 2009; **20**: 35–42.

24 Bethell J, Bondy SJ, Lou WYW, Guttmann A, Rhodes AE. Emergency department presentations for self-harm among Ontario youth. *Can J Public Health* 2013; **104**: e124-30.

25 Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. *Am J Psychiatry* 2007; **164**: 1035–43.

26 CIHI. The Canadian Enhancement of ICD-10. , 2001.

27 Fransoo R, Martens P, Burland E, The Need to Know Team, Prior H, Burchill C. Manitoba RHA Indicators Atlas 2009. , 2009 (http://mchp-appserv.cpe.umanitoba.ca/reference/RHA_Atlas_Report.pdf).

28 Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; **37**: 360–3.

29 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74.

30 Iribarren C, Sidney S, Jacobs DR, Weisner C. Hospitalization for suicide attempt and completed suicide: epidemiological features in a managed care population. *Soc Psychiatry Psychiatr Epidemiol* 2000; **35**: 288–96.

31 Simon GE, Savarino J. Suicide attempts among patients starting depression treatment with medications or psychotherapy. *Am J Psychiatry* 2007; **164**: 1029–34.

32 Nock MK, Joiner TE, Gordon KH, Lloyd-Richardson E, Prinstein MJ. Non-suicidal self-injury among adolescents: diagnostic correlates and relation to suicide attempts. *Psychiatry Res* 2006; **144**: 65–72.

Table 4-1: Descriptive statistics of index emergency department events captured in the SAFE datasets

	N	%
Sample	5719	100
Sex(male)	3013	52.7
Admitted to hospital	3328	58.2
C-CASA classification		
Suicide attempt	780	13.6
Self-harm*	1147	20.1
Imminent plans	190	3.3
Suicidal ideation	1495	26.1
	Mean	SD
Age	41.4	17.4

*includes suicide attempts

Table 4-2: Administrative data diagnostic codes compared to C-CASA classification codes

		ICD codes for Intentional self-harm (X60-X84)			
		All emergency visits		Admitted to hospital	
C-CASA classification		Yes	No	Yes	No
Suicide attempt	Yes	164	616	164	280
	No	36	4900	36	2848
Self-harm	Yes	185	962	185	437
	No	15	4554	15	2691

Table 4-3: Administrative data diagnostic codes with expanded definitions compared with C-CASA classification

		X60-X84, Y10-Y34			
		Emergency visits		Admitted to hospital	
C-CASA classification		Yes	No	Yes	No
Suicide attempt	Yes	182	598	182	262
	No	60	4876	60	2824
Self-harm	Yes	213	934	213	409
	No	29	4540	29	2677
		X60-X84, Y10-Y34, X40-X49			
Suicide attempt	Yes	199	581	199	245
	No	99	4837	99	2785
Self-harm	Yes	249	898	249	373
	No	49	4520	49	2657

X60-X84 are the ICD codes for intentional self-harm
Y10-Y34 are the ICD codes for undetermined intent
X40-X49 are the ICD-10 codes for accidental poisoning

Table 4-4: Validation statistics for admitted patients using different coding methods

	X60-X84	X60-X84, Y10-Y34	X60-X84, Y10-Y34, X40-X49
		Suicide Attempt	
Statistic	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
Sensitivity	36.9% (32.4% to 41.4%)	41.0% (36.4% to 45.6%)	44.8% (40.2% to 49.4%)
Specificity	98.8% (98.3% to 99.2%)	97.9% (97.4% to 98.4%)	96.6% (95.9% to 97.2%)
PPV	82.0% (76.7% to 87.3%)	75.2% (69.8% to 80.6%)	66.8% (61.4% to 72.1%)
NPV	91.0% (90.0% to 92.0%)	91.5% (90.5% to 92.5%)	91.9% (90.9% to 92.9%)
		Self-harm	
Statistic	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
Sensitivity	29.7% (26.2% to 33.3%)	34.2% (30.5% to 38.0%)	40.0% (36.2% to 43.9%)
Specificity	99.4% (99.2% to 99.7%)	98.9% (98.5% to 99.3%)	98.2% (97.7% to 98.7%)
PPV	92.5% (88.8% to 96.2%)	88.0% (83.9% to 92.1%)	83.6% (79.3% to 87.8%)
NPV	86.0% (84.8% to 87.2%)	86.7% (85.6% to 87.9%)	87.7% (86.5% to 88.9%)

X60-X84 are the ICD codes for intentional self-harm
Y10-Y34 are the ICD codes for undetermined intent
X40-X49 are the ICD-10 codes for accidental poisoning

Table 4-5: Emergency Department Information System (EDIS) triage category by C-CASA classification

	Suicide Attempt N = 780		Self-harm N = 1147		No Self-harm N = 4569	
Primary triage category	N	column %	N	column %	N	column %
Mental Health	344	44.1	559	48.7	3368	73.7
Substance misuse	276	35.4	348	30.3	104	2.3
Trauma	44	5.6	61	5.3	68	1.5
Multiple*	13	1.7	15	1.3	19	0.4
Secondary triage category	N	column %	N	column %	N	column %
Depression/suicidal/self-harm	326	41.8	467	40.7	1596	34.9
Overdose	269	34.5	335	29.2	46	1.0
Laceration	15	1.9	32	2.8	9	0.2
Multiple*	9	1.2	10	0.9	8	0.2
Tertiary triage category	N	column %	N	column %	N	column %
Attempted suicide, clear plan	195	25.0	233	20.3	162	3.5
Active thoughts	56	7.2	79	6.9	271	5.9
High risk/unknown substance	78	10.0	104	9.1	17	0.4
Altered Level-of-consciousness	29	3.7	42	3.7	62	1.4
Multiple*	c	c	c	c	c	c

*visit contains more than one of the above codes

**self-harm includes all C-CASA classifications involving self-injury

c= censored due to small number of events

Table 4-6: Assessment of bias in triage chief complaint among admitted patients using ICD-derived attempt and self-harm groups

	Suicide Attempt* N = 444		Self-harm* N = 622		X60-X84 N = 200		X60-X84, Y10-Y34 N = 242		X60-X84, Y10-Y34, X40-X49 N = 298	
Primary triage category	N	column %	N	column %	N	column %	N	column %	N	column %
Mental Health	205	46.2	311	50.0	55	27.5	63	26.0	71	23.8
Secondary triage category	N	column %	N	column %	N	column %	N	column %	N	column %
Depression/suicidal/self-harm	193	43.5	248	39.9	52	26.0	58	24.0	63	21.1
Overdose	141	31.8	159	25.6	76	38.0	84	34.7	116	38.9
Tertiary triage category	N	column %	N	column %	N	column %	N	column %	N	column %
Attempted suicide, clear plan	105	23.6	123	19.8	42	21.0	45	18.6	49	16.4

*C-CASA classification of suicide attempt or any self-harm
X60-X84 are the ICD codes for intentional self-harm
Y10-Y34 are the ICD codes for undetermined intent
X40-X49 are the ICD-10 codes for accidental poisoning

CHAPTER 5: USING LATENT CLASS ANALYSIS TO DETECT SELF-HARM EVENTS IN THE EMERGENCY DEPARTMENT

5.1 Chapter overview

The next paper examines the use of latent class analysis to determine the occurrence of self-harm presentations to the emergency department. Earlier work examined the ability of ICD-10-CA codes from hospitalization records to identify self-harm and suicide attempt presentations for individuals contained within the SAFE database of emergency department (ED) psychiatry consultations. **Chapter 4** showed that the codes had good positive predictive validity but only detected slightly less than half of the hospitalized suicide attempts and self-harm identified in the SAFE data. There is clearly room for improved methods of identifying suicide attempts using administrative data variables.

This chapter aimed to identify the individuals not linked in the previous study due to missing provincial health insurance numbers and assessment dates in the SAFE data. It also tried to identify self-harm presentations not consulted to psychiatry. **Chapter 5** will use latent class analysis to determine and describe the self-harm presentations using administrative and SAFE data. Latent class analysis is used to identify unmeasured groups (i.e., latent classes) from measured categorical variables (i.e., indicator variables) that are associated with those unmeasured groups. Available measures that may be effective at identifying the latent class of self-harm presentations include mental health diagnosis codes, injury codes, triage chief complaint and psychiatry assessment/treatment. The paper will discuss whether the measures used are capable of identifying presentations due to self-harm.

5.2 Abstract

Objective: To identify instances of self-harm and describe the variables associated with them.

Methods: Linked health administrative databases from Manitoba, Canada were used to identify emergency department events during routine psychiatric assessment data collection from 2009 to 2012. Injury, mental health, and psychiatric treatment information was extracted from administrative data. Triage chief complaints from the Emergency Department Information System (EDIS) including 'mental health', 'depression/suicidal/self-harm', 'overdose ingestion', and 'suicide attempt, clear plan' were used to construct the study measures. Psychiatrist-determined self-harm presentations were identified from patients consulted to psychiatry. The analysis sample included events with at least one mental health measure or triage complaint. One randomly selected presentation for each individual was included in the analysis sample. Latent class analysis was applied to the study measures using PROC LCA with SAS software. Bayesian information criterion was used to determine the number of classes. The identified latent classes were described and classes with potential self-harm presentations were identified.

Results: A latent class model with 7 classes was identified as optimal, including 1 clear self-harm class. The self-harm class contained 6.26% of the sample. The majority in this class had psychiatrist determined self-harm (73.7%) and a mood or adjustment disorder diagnosis (65.9%). Substance use disorders (33.7 %) and personality disorders (19.9%) were also common. 'Overdose ingestion'' (31.4%) and 'depression/suicide/self-harm' (40.5%) accounted for most of the triage complaints in the self-harm group. Another latent class had triage complaints similar to the self-harm class but without psychiatric assessment-related measures. This class could contain self-harm events not assessed by psychiatry.

Conclusion: The latent class model identified a clear self-harm class but also identified a second class of potentially under-documented self-harm, overdose, and ideation events. These results suggest that administrative data collection should be improved in order to provide a better single indicator of self-harm.

5.3 Background

The occurrence of suicide attempts presenting to the emergency department (ED) is used both an outcome for risk factor and treatment research and used in epidemiological studies of the occurrence of suicidal behaviour.[1–4] Estimates for the proportion of ED visits involving suicide attempts or self-harm are in the area of 0.4% to 0.5% of the total visits in the US.[5–8] However, using International Classification of Diseases (ICD) codes to identify ED events and inpatient admission that involved suicide attempts results in imperfect detection.[9,10] In particular, ICD codes are likely to undercount the true incidence of self-harm events. Better methods of identifying self-harm in administrative data are needed to improve research using these data.

Administrative data are regularly collected under the single-payer health care systems found among Canadian provinces. These data include ICD injury intent codes for admitted patients previously used in research involving suicide attempts.[1–4,11] However, determining the number of ED events in Manitoba is hindered by ICD injury intent codes not being recorded in administrative data for EDs– such codes are only available for those also admitted to inpatient units. Assessment of suicidality using the Columbia Classification Algorithm of Suicide Assessment (C-CASA)[12] is done as part of psychiatric consultation in the two main tertiary hospitals in Winnipeg, Manitoba. These data provide a potentially invaluable source of information to detect treated attempts in the ED and inpatient settings and provide better estimates of the incidence of self-harm events than provided using ICD injury codes alone.

However, there are limitations to the data available. First of all, C-CASA information is only recorded for individuals receiving a psychiatric assessment during their treatment. Although psychiatric referral for suicide attempts is expected to be high at these locations, the actual rate of consultation for these patients is unknown. An unknown number of missed patients could bias any new detection methods. Secondly, the C-CASA records are missing key data elements needed to link them with administrative records for approximately 15% of the consultations.

This article aims to use latent class analysis (LCA) and measures of self-harm from administrative data and the C-CASA assessment to identify self-harm events in the emergency department. This paper aimed to identify the broader group of all self-harm events (as opposed to specifically suicide attempt presentations) because the previous chapter indicated that separating these events based on intent was difficult with these data. LCA uses manifest variables to identify unmeasured groups. These unmeasured groups are referred to as latent classes. LCA is helpful in determining disease status when there is no gold standard as true disease status is a latent class.[13,14] Previous research has used LCA to assess the validity of diagnoses (including mental health related diagnoses) when there is no accurate confirmation of true disease status.[14] This method can identify the latent class of self-harm events from diagnostic codes, triage complaints, and clinical assessment available in administrative data. This will not only provide estimates for the occurrence of self-harm events and how many self-harm events are not being consulted to psychiatry as part of their assessment, but also provide information for improved validation of better methods to detect self-harm events using these measures. This method also has the added benefit of not requiring any chart review or other similar labour-intensive methods of determining a gold standard to compare against.

5.4 Methods

5.4.1 Data sources

This study utilizes databases from the Population Research Data Repository at the Manitoba Centre for Health Policy. This repository includes data on over 99% of individuals residing in Manitoba.[15] This study uses the hospital discharge abstract data, the Emergency Department Information System (EDIS) database, and the medical claims database. The hospital discharge abstract data contain information on patients admitted to the inpatient units of hospitals within the province, including diagnoses coded using the International Classification of Diseases, 10th revision, Canadian version (i.e., ICD-10-CA).[16] EDIS contains information on all individuals presenting to the emergency departments of the two main tertiary hospitals in Winnipeg. These data include several fields that contain individuals' presenting complaint. The medical claims database contains records of all fee-for-service payments made in the province using the 9th revision (Clinical Modification) of the ICD classification system (ICD-9-CM).[17] These records include one field for diagnosis, an indicator of physician specialty, referral information, and treatment information. However, ED physicians in the study locations are no longer required to submit medical claims for their services.

In addition to the databases from the Repository, this study also utilizes the Suicide Assessment Form in Emergency psychiatry (SAFE) data, which were obtained from the emergency departments of Winnipeg's two main tertiary hospitals. Those hospitals have fulltime psychiatric staff and psychiatry residents. Data on psychiatric consults were collected between January 1, 2009 and December 31, 2012. These data were linked to the Repository through Provincial Health Insurance Numbers and date of assessment. This form contains the C-CASA assessment of suicidality where individuals are sorted into mutually exclusive categories –

suicide attempt, self-injurious behaviour: no suicidal intent, self-injurious behaviour: intent unknown, suicidal ideation, preparatory acts towards imminent suicidal behaviour, unknown and no suicidality.[12]

5.4.2 Sample

EDIS was used to identify all presentations to the emergency departments of the two hospitals during the collection of SAFE data between January 1, 2009 and December 31, 2012. Date of treatment was used to link each SAFE assessment to an emergency department presentation. Events, which were the unit of analysis, were created by identifying all presentations in EDIS data and assessment records on the same date or subsequent dates in the SAFE dataset, along with relevant hospital discharge abstracts and medical claims data. Two days without any records in the data determined the end of an event.

5.4.3 Study measures

A series of study measures (i.e., present/absent or yes/no) were constructed. Injury intent ICD codes in hospital abstract data and physician claims with injury codes were identified and combined to form a 5-level categorical variable (intentional self-harm admission, injury of unknown intent admission, accidental poisoning admission, other injury physician claim, and no injury codes). The ICD codes for this indicator are found in **Table 5-1**. Diagnosis codes from hospital discharge abstracts and physician claims were used to ascertain five measures of mental illness disorders: mood/adjustment, anxiety, personality, schizophrenia, and substance use. The ICD codes for these indicator variables are shown in **Table 5-1**. Mood/adjustment, substance use, and personality disorders were combined into an 8 category joint-item. Physician claims records were also used to develop an indicator for psychiatry assessment/treatment.

This study used the triage presenting complaint fields in EDIS to extract additional measures. There are three hierarchical fields; the first two are mandatory and the third one is optional. The first field classified events under broad categories such as ‘Mental health’, while a second field allows for more specific categorization. Complaints of interest in the second field were ‘Depression, suicide, self-harm’, and ‘Overdose ingestion’. The last complaint field was scanned for individuals listed as ‘Attempted suicide, clear plan’. Binary measures for the ‘Mental health’, ‘Depression, suicide, self-harm’, ‘Overdose ingestion’, and ‘Attempted suicide, clear plan’ were combined into a single joint-item.

The C-CASA assessment was used to specify a series of indicator variables. The self-harm indicator combined the following C-CASA categories into a single binary indicator of any deliberate self-harm; suicide attempt, self-injurious behaviour: no suicidal intent, and self-injurious behaviour: intent unknown. A suicidal ideation indicator variable was created by combining the following C-CASA categories; suicidal ideation, and preparatory acts towards imminent suicidal behaviour. A C-CASA assessment of no suicidality was used to create a ‘No suicidal ideation/Self-harm’ indicator. Sex and age were combined into a 6-category indicator based on the following age groups; <25, 25 to 49, and 50+.

5.4.4 Statistical Analysis

Latent class analysis was used to determine self-harm and suicidal ideation classes. The outputs provided by latent class analysis include the proportion of the sample that is within each class, and the proportion of individuals within each class having each of the measures. The proportion of individuals having each of the measures within each class provides a measure of the association between those measures and membership in that class. A high proportion of

individuals with the C-CASA self-harm measure will indicate the self-harm latent class. Conditional dependence was a concern with some of the binary measures. Conditional dependence occurs when measures within a latent class are correlated. Conditional dependence can cause biased estimates for class membership and occurrence of measures within different classes. This dependence can be accounted for with several methods. Additional latent classes can be added to accommodate conditional independence but in this sample that would require adding a substantial number of non-informative classes. Instead of adding more classes, related binary measures were combined using the joint-item approach.[18,19] This entails combining several correlated indicator variables into one variable showing all possible combination of the component measures.

The latent class model includes the following measures; injury diagnosis, physician diagnosis (mood/adjustment disorders, substance use and personality disorders), two triage complaint measures (joint item combination of 'Depression, suicide, self-harm' triage, 'Overdose ingestion' triage, and 'Mental Health' triage complaints; combination of 'Attempt Suicide, clear plan' and 'Suicidal Ideation' triage complaints) and the C-CASA binary indicator variables (i.e., 'Self-Harm', 'Suicidal Ideation', 'No Self-harm/Ideation'). To simplify the presentation of results the joint-item combinations were presented as binary or categorical since the joint-item combinations produce many categories and are not straightforward to interpret.

Study measures that were potentially informative but not likely to be coded as a result of self-harm specifically (i.e., psychosis and anxiety disorders, and treatment by psychiatrists) were not included as measures in the model since they would cause several irrelevant subgroups of non-suicidal mental illness classes to be identified. These measures were included as covariates using

the covariates function of PROC LCA. This allows information from these classes to be used in classifying individuals into classes while reducing the number of non-substantial classes to be identified.

The sample was split into two equal halves. The first half was used to develop the latent class model and the second half was used to provide the estimates for the classes. The estimated percent of events classified into the self-harm group and rho parameters for suicide-related classes were compared between these two halves and a lack of agreement between the two could be a sign of poor internal validity.

LCA uses maximum likelihood estimation to determine the parameter coefficients, but this method is sensitive to the starting coefficients of the model and may identify a local maximum instead of the global maximum.[19,20] Therefore multiple models need to be run with different starting parameters (i.e., different seeds) in order to ensure that the correct model is arrived upon. The models were run using 100 seeds for different numbers of latent classes to determine the best fitting number of classes. These model runs started with 4 latent classes and continued until the number of classes with the smallest Bayesian Information Criteria (BIC) was identified. The model started with 4 classes since that the minimum number of distinct classes expected a priori. The BIC determined the number of classes in the model since prior research supports this criteria as the best at identifying the correct number of classes.[21] The entropy and the percent agreement among seeds were recorded. The entropy statistic assesses how confidently the model can assign individuals into latent classes. The entropy statistic ranges from 0 to 1 with 0.8 or higher considered an indication of good separation of classes.[22] The percent agreement among seeds shows how much of the seeds agree with the selected model. If seed agreement is low then the model might have poor identification or have identified a local minimum.[20] The G^2 statistic and

its degrees of freedom, which indicate whether the model had good fit, were also recorded. A poor measure of fit indicated by the G^2 statistic is a good indication that the conditional independence assumption is being violated.[18,20] The analysis was performed using SAS software version 9.4 (SAS Institute; Cary, NC).[23]

5.4.5 Ethics and Data Access Approvals

Ethical approval was obtained from the University of Manitoba's Health Research Ethics Board (HREB #: H2015:009) and data access was granted by the Government of Manitoba's Health Information Privacy Commission (HIPC #- 2014/2015 – 41).

5.5 Results

5.5.1 Sample descriptive statistics

A total of 313,057 adult ED events were identified with 29951 having at least one mental health diagnosis or triage complaint present. The analysis sample contained 19,363 events after reducing sample to one event per person. Descriptive statistics for the analysis sample, including study measures, are shown in **Table 5-2**. The analysis sample had a median age of 44 and 9386 (48.5%) were male. There were linkable SAFE assessments included for 4537 (23.4%) individuals with 883 (4.6%) and 1279 (6.6%) individuals classified as involving self-harm or suicidal ideation. The other study measures had the following occurrences; 'Overdose ingestion'(7.9%), 'Depression/suicide/self-harm' (17.3%), 'Other mental health' (19.8%), 'Attempted suicide, clear plan' (3.2%), 'Suicidal ideation' (3.0%), 'Intentional self-harm' (1.6%), 'Unknown intent' injury (0.7%), 'Accidental poisoning' (1.7%), other injuries (9.8%), mood/adjustment disorder (28.2%), substance use disorder (14.9%), personality disorder (4.0%),

anxiety disorder (18.4%), schizophrenia disorder (3.1%), psychiatry assessment/treatment (33.3%).

5.5.2 Latent Class Analysis Modelling

Information on model fit statistics used to determine the optimal number of classes is located in **Table 5-3**. The BIC statistic identified 7 classes as the best number of classes with a value of 4985. The G^2 statistic was 3086 (df= 22865) which indicates good model fit. Entropy was 0.85 and seed agreement was 41%. All models with 5 or more classes had good model fit based on the G^2 statistic.

5.5.3 Latent Class Analysis estimates

Table 5-4 presents estimates from the latent class analysis**.** Items combined for joint-item modelling are shown as proportions for the individual items to aid interpretation. Class 1 contained 6.26% of the events and these events likely involved self-harm as 73.7% of the class was screened as being self-harm events in the SAFE data. The 'Overdose ingestion' triage complaint occurred for 31.4% of the class, 40.5% had the depression/suicide/self-harm triage complaint and 19.7% had the 'attempted suicide, clear plan' triage indicator. The intentional self-harm admission codes occurred in 19.5% of the class with 2.7% and 3.7% having unknown intent and accidental poisoning codes. The self-harm class had the highest occurrence of both substance use (33.7%) and personality disorder (19.9%) codes. Mood and adjustment disorder codes were also elevated compared to most classes (65.9%), while anxiety and psychosis codes were less common at 17.7% and 6.5%. Members of this class were slightly more often female (53.1%) and slightly younger. Almost all of the members of this class were seen by a psychiatrist (98.3%).

Class 2 contains a high portion of people with suicidal ideation on the C-CASA (96.4%) and is the suicide ideation class. A total of 6.78% of the events belonged to this class. The Depression/Suicide/Self-harm triage complaint was found in a solid majority of this group (68.0%), and the 'Attempted Suicide, clear plan' complaint also occurred (9.1%). The 'Suicidal Ideation' complaint was highest in this class (15.4%). Injury codes and overdose triage complaint were very infrequent. Mental disorder codes were found at approximately the same frequency as in the self-harm group. Individuals in this class were also almost always seen by a psychiatrist (99.2%).

Individuals in class 3 had a triage complaint of overdose or depression/suicide/self-harm, but almost no other measures. In particular, this group had no recorded claims by a psychiatrist. The age profile of this class is also young like the self-harm class. Classes 4 through 7 all contain mixtures of mental disorder codes and/or mental health triage codes. Class 5 contains mostly mental health triage complaints, no psychiatry assessment/treatment, and no mental health ICD diagnosis which indicates that they were likely non-suicidal patients discharged from the ED. Class 6 was treated by psychiatry and received mental health ICD codes. Slightly more than half (68.2%) had a mental health triage complaint. This class is likely a combination of psychiatry assessed mental health ED patients and inpatient consultations based on a third not having a mental health triage complaint and the rarity of young patients. Class 7 all had anxiety codes and almost no other measures and are potentially anxious patients with primarily medical conditions.

5.5.4 Split-half agreement on self-harm proportion

The two halves of the sample produced similar estimates for the proportion of events that involved self-harm. The first half estimated that 6.4% (95% confidence interval: 5.8% to 6.9%)

of the events involved self-harm versus 6.3% (95% confidence interval: 5.7 to 6.8%) in the second half.

5.6 Discussion

Although the LCA model was capable of identifying a distinct self-harm class, some self-harm cases might be occurring within Class 3. The model itself was likely also only effective due to the use of C-CASA assessments which are not widely available in administrative data. Better measures or increased recording of C-CASA or similar assessments in the ED are needed to provide accurate detection.

The LCA models developed in this study estimated that 6.3% of the events identified in the analysis sample involved self-harm treated in the ED. An additional 6.8% of the events were identified as involving suicidal ideation, and a third class involving 10.3% of the events could contain events involving ideation or self-harm that lack sufficient measures to reliably classify them into either the self-harm or ideation group. This class had triage complaints that could indicate a self-harm event but were not assessed by psychiatry and not diagnosed with a mental health code. However, it is also possible that the group contained a combination of depression-related and non-mental health related overdoses, with few or none of the events involving self-harm. Currently available data are unable to determine whether these events involve self-harm or not. Therefore, to accurately detect suicide attempts there must either be improvements in the triaging of mental health patients recording through EDIS, or the SAFE assessment form needs to be expanded to all mental health patients.

A study using the American National Hospital Ambulatory Medical Care Survey (NHAMCS) provided estimates for mental health diagnosis codes among suicide attempt

events.[5] They found that 54% percent of those identified as suicide attempt events had at least one mental health code (ICD-9 290-319). The estimates here are higher than those from the NHAMCS survey, with mood and adjustment disorders alone being diagnosed in 65.9% compared to 34% found in that paper. They also found 12% had alcohol abuse (no other substance use statistics were reported). Another American study that relied on ICD codes to identify suicide attempts and to examine the mental disorders coded among them also reported considerably lower occurrence of these disorders than our results found [6]. Their sample reported 42.1% with mood disorders (another 2.5% reported adjustment disorders which were combined with mood disorders in our study), 12.1% had substance use disorders and 8.9% had alcohol use disorders recorded, 6.4% had anxiety disorders, and 3.6% had schizophrenia. Only 0.5% of their sample had a recorded diagnosis of a personality disorder. A systematic review from Japan also provided some pooled prevalence estimates for mental disorders [24]. Their estimates based on ICD codes were lower than those found in this study with the exception of their codes for schizophrenia, schizotypal and delusional disorders (estimated at 13% versus the 9.5% estimate for schizophrenia here). Personality disorders were found to be lower using ICD codes (13%), they were drastically higher in studies using questionnaire based classification (41%, 95% CI= 24% to 60%). Non-Japanese studies also had lower estimates for mental disorders, with the exception of a Korean study (finding that 79% of suicide attempts were diagnosed with a mood disorder) and an American study showing 34% of attempters being diagnosed with a personality disorder.

These discrepancies might be due to individuals primarily assessed and treated for mental disorders with non-medically serious attempts not receiving ICD codes and therefore not being detected in those studies. Conversely, individuals without these codes might be more likely to be

missed by the LCA model in this study – since probability of membership in the class is partially determined by these diagnoses. It is possible the third latent class may contain some instances of self-harm without these codes. These results suggest that the errors in detection are likely to significantly bias the statistics being reported in these papers as long as they rely on inconsistent ICD coding to identify self-harm and suicide attempts.

There are some limitations to this study. Previous research has used LCA to identify disease groups in a fashion similar to the current study,[13,14,25] but this research has identified some limitations with the LCA methods.[26] A fundamental study limitation is in requiring a series of measures that, as a whole, can identify all the instances of self-harm. This is similar to the limitations encountered when using multiple imputation or other related methods of adjusting for missing data. However, whether the current model is sufficient cannot be reliably determined. One potential issue with the LCA approach is a tendency for the method to overestimate the accuracy of measures of identify the latent class of interest.[26] This is not a significant issue with the current study since predictive statistics are not being determined and identifying classes with potentially missed self-harm cases was a study objective. Another limitation of this study was that it focused on the broader category of self-harm rather than attempting to identify different sub-groups of self-injury due to the limitations of the data available in administrative data. Without the C-CASA data it would likely not be possible to reliably discriminate between these groups with administrative data. Another limitation is that this study relies on a probabilistic linkage of records across consecutive dates to create an event. This may result in some records being combined erroneously, including repeated presentations in a short time frame, which could bias the results away from what would have been found if a perfect linkage of records was available.

Conditional dependence is an issue when using LCA, particularly when diagnosing/detecting conditions as in this study.[27] Although the G^2 statistic suggested good model fit and no violation of conditional independence, the tests for model fit used for LCA are not very accurate.[28] However, using BIC to determine the number of classes has been shown as effective and will eliminate any significant conditional dependence by adding additional classes.[21,29] The measures used in the study also had several limitations. Some of the measures are overly broad (e.g., 'Depression/Suicide/Self-harm', and 'Overdose Ingestion' triage complaints) which limits their usefulness. Another limitation is that approximately 15% of the SAFE C-CASA assessments are missing in the analysis – mostly due to missing provincial health insurance numbers or assessment dates preventing them from being linked into the administrative data.

5.7 Conclusions

This LCA model in this study found a class that was likely all presentations due to self-harm but also identified another class that might contain self-harm presentations not consulted to psychiatry. Currently the available administrative data are not capable of reliably identifying ED events involving self-harm. The main implication of this study is that improvements in documentation of self-harm are required. These improvements likely require changes to the current triage structure of the EDIS system. Particularly there is a need for more specific triage groups that do not combine suicidal ideation, and self-harm with depression under a single triage category.

5.8 References

1 Randall JR, Walld R, Finlayson G, Sareen J, Martens PJ, Bolton JM. Acute risk of suicide and suicide attempts associated with recent diagnosis of mental disorders: a population-based, propensity score-matched analysis. *Can J Psychiatry* 2014; **59**: 531–8.

2 Bolton JM, Spiwak R, Sareen J. Predicting suicide attempts with the SAD PERSONS scale: a longitudinal analysis. *J Clin Psychiatry* 2012; **73**: e735-41.

3 Bolton JM, Walld R, Chateau D, Finlayson G, Sareen J. Risk of suicide and suicide attempts associated with physical disorders: a population-based, balancing score-matched analysis. *Psychol Med* 2015; **45**: 495–504.

4 Rosychuk RJ, Johnson DW, Urichuk L, Dong K, Newton AS. Does emergency department use and post-visit physician care cluster geographically and temporally for adolescents who self-harm? A population-based 9-year retrospective cohort study from Alberta, Canada. *BMC Psychiatry* 2016; **16**: 229.

5 Ting SA, Sullivan AF, Boudreaux ED, Miller I, Camargo CA. Trends in US emergency department visits for attempted suicide and self-inflicted injury, 1993–2008. *Gen Hosp Psychiatry* 2012; **34**: 557–65.

6 Canner JK, Giuliano K, Selvarajah S, Hammond ER, Schneider EB. Emergency department visits for attempted suicide and self harm in the USA: 2006-2013. *Epidemiol Psychiatr Sci* 2016; : 1–9.

7 Doshi A, Boudreaux ED, Wang N, Pelletier AJ, Camargo CA. National Study of US Emergency Department Visits for Attempted Suicide and Self-Inflicted Injury, 1997-2001. *Ann Emerg Med* 2005; **46**: 369–75.

8 Larkin GL, Smith RP, Beautrais AL. Trends in US Emergency Department visits for suicide attempts, 1992-2001. *Crisis* 2008; **29**: 73–80.

9 Walkup JT, Townsend L, Crystal S, Olfson M. A systematic review of validated methods for identifying suicide or suicidal ideation using administrative or claims data. *Pharmacoepidemiol Drug Saf* 2012; **21 Suppl 1**: 174–82.

10 Randall JR, Roos LL, Lix LM, Katz LY, Bolton JM. Emergency department and inpatient coding for self-harm and suicide attempts: valdiation using clinician assessment data. *Int J Methods Psychiatr Res* 2017. doi:10.1002/mpr.1559.

11 Martens PJ, Burchill C, Fransoo R, De Coster C, McKeen N, Ekuma O, *et al.* Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study. , 2004.

12 Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the

FDA's pediatric suicidal risk analysis of antidepressants. *Am J Psychiatry* 2007; **164**: 1035–43.

13 Rutjes a WS, Reitsma JB, Coomarasamy a, Khan KS, Bossuyt PMM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007; **11**: iii, ix-51.

14 van Smeden M, Naaktgeboren C a, Reitsma JB, Moons KGM, de Groot J a H. Latent class models in diagnostic studies when there is no reference standard--a systematic review. *Am J Epidemiol* 2014; **179**: 423–31.

15 Nickel NC, Chateau DG, Martens PJ, Brownell MD, Katz A, Burland EMJ, *et al.* Data resource profile: Pathways to Health and Social Equity for Children (PATHS Equity for Children). *Int J Epidemiol* 2014; **43**: 1438–49.

16 CIHI. The Canadian Enhancement of ICD-10. , 2001.

17 National Center for Health Statistics. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). , 2005.

18 Uebersax J. A Practical Guide to Conditional Dependence in Latent Class Models. 2009. (http://www.john-uebersax.com/stat/condep.htm#intro).

19 Vermunt JK. LEM: A general program for the analysis of categorical data. 1997.

20 Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS Procedure for Latent Class Analysis. *Struct Equ Model A Multidiscip J* 2007; **14**: 671–94.

21 Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Struct Equ Model A Multidiscip J* 2007; **14**: 535–69.

22 Ramaswamy V, DeSarbo W, Reibstein D, WT R. An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Mark Sci* 1993; **12**: 103–24.

23 SAS Institute Inc. SAS software. 2012. SAS Institute; Cary, NC.

24 Kawashima Y, Yonemoto N, Inagaki M, Yamada M. Prevalence of suicide attempters in emergency departments in Japan: A systematic review and meta-analysis. *J Affect Disord* 2014; **163**: 33–9.

25 Sewitch MJ, Jiang M, Joseph L, Hilsden RJ, Bitton A. Developing model-based algorithms to identify screening colonoscopies using administrative health databases. *BMC Med Inform Decis Mak* 2013; **13**: 45.

26 Spencer BD. When do latent class models overstate accuracy for diagnostic and other classifiers in the absence of a gold standard? *Biometrics* 2012; **68**: 559–66.

27 Van Smeden M, Oberski DL, Reitsma JB, Vermunt JK, Moons KGM, De Groot JAH. Problems in detecting misfit of latent class models in diagnostic research without a gold standard were shown. *J Clin Epidemiol* 2016; **74**: 158–66.

28 Subtil A, de Oliveira MR, Gonçalves L. Conditional dependence diagnostic in the latent class model: A simulation study. *Stat Probab Lett* 2012; **82**: 1407–12.

29 Oberski DL. Beyond the number of classes: separating substantive from non-substantive dependence in latent class analysis. *Adv Data Anal Classif* 2016; **10**: 171–82.

Table 5-1: ICD codes used to detect mental disorder and injuries

Measure	ICD-9-CM codes	ICD-10-CA
Anxiety:	300	F40, F41.0, F41.1, F41.3, F41.8, F41.9, F42
Mood and adjustment disorders:	296, 309, 311	F31, F32, F33, F34.1, F38.0, F38.1, F41.2, F43.1, F43.2, F43.8, F53.0, F93.0
Personality disorders:	301	F60
Substance use:	291	F10–F19, F55
Schizophrenia:	295	F20, F21, F23.2, F25
Intentional self-harm:	-	X60-X84
Injury of unknown intent:	-	Y10-Y34
Accidental poisoning:	-	X40-X49
Other injury:	800-999	-

Note: Hospital discharge abstracts records diagnosis using ICD-10-CA codes. Medical claims record diagnosis using ICD-9-CM.

Table 5-2: Descriptive statistics for the study data

		n	%*
All events		313057	
Events with at least 1 indicator		29951	
Analysis sample**		19363	100.0
Sex	Male	9386	48.5
	Female	9977	51.5
SAFE assessment		4537	23.4
	Self-harm	883	4.6
	Ideation	1279	6.6
Triage complaints			
	Overdose Ingestion	1522	7.9
	Depression/Suicidal/Self-harm	3340	17.3
	Other Mental Health	3840	19.8
	Attempted Suicide, clear plan	618	3.2
	Suicidal Ideation	586	3.0
Injury code			
	Intentional Self-harm	307	1.6
	Unknown Intent	128	0.7
	Accidental Poisoning	327	1.7
	Other injuries	1904	9.8
ICD-9/10 Mood/adjustment codes		5455	28.2
ICD-9/10 Substance use codes		2877	14.9
ICD-9/10 Personality disorder codes		780	4.0
ICD-9/10 Anxiety codes		3567	18.4
ICD-9/10 Schizophrenia codes		609	3.14
Psychiatry assessment/treatment		6438	33.3
Age	Mean	46.3 (SD = 20.2)	
	Median	44	

*percent of analysis sample

**analysis sample consists of events with at least 1 indicator after restricting to only one event per individual

Table 5-3: Latent class analysis model fit statistics

Classes	4	5	6	7	8
G^2 (DF)	4999 (22940)	3886 (22915)	3384 (22890)	3086 (22865)	2870 (22840)
BIC	5908	5024	4751	4683	4697
Entropy	0.93	0.93	0.95	0.85	0.86
Agreement	51%	75%	16%	41%	10%

DF = degrees of freedom
BIC = Bayesian Information Criterion

Table 5-4: Latent class model with 7 classes

	Percent of analysis sample within class						
	1	2	3	4	5	6	7
	6.26	6.78	10.28	36.46	13.70	13.64	12.89
C-CASA Self-harm/Suicide attempt	73.7	0.0	0.7	0.0	0.0	0.0	0.0
C-CASA Suicidal Ideation	1.6	96.4	0.6	0.0	0.0	0.0	0.0
Triage complaints							
Overdose Ingestion	31.4	0.3	31.6	0.7	16.3	0.9	0.3
Depression/Suicidal/Self-harm	40.5	73.6	68.3	0.1	6.1	12.8	0.2
Other Mental Health	5.8	13.6	0.0	0.8	77.6	55.4	0.4
Attempted Suicide, clear plan	19.7	9.1	12.7	0.0	0.0	0.2	0.0
Suicidal Ideation	4.5	15.4	14.7	0.0	0.0	1.4	0.0
Injury code							
Intentional Self-harm	19.5	0.0	0.0	1.0	0.0	0.3	0.0
Unknown Intent	2.7	0.1	0.0	0.9	0.0	0.4	0.0
Accidental Poisoning	3.7	0.2	5.5	2.2	0.0	0.3	0.0
Other injuries	8.5	3.1	1.9	16.2	1.7	8.5	12.4
ICD-9/10 Mood/Adjustment codes	65.9	68.0	0.7	36.1	0.2	44.8	0.5
ICD-9/10 Substance use codes	33.7	27.7	0.6	23.4	0.0	16.3	0.5
ICD-9/10 Personality disorder codes	19.9	16.4	0.0	1.8	0.1	7.4	0.0
ICD-9/10 Anxiety codes	17.7	16.8	0.5	4.6	0.9	10.7	100.0
ICD-9/10 Schizophrenia codes	6.5	9.0	0.1	3.5	0.1	29.3	0.0
Psychiatry assessment/treatment	98.25	99.19	0	16.7	0	99.53	5.5
Sex/Age Groups							
Males	46.9	49.5	45.2	47.9	50.8	56.1	39.6
<25	14.4	11.4	14.1	3.6	10.9	11.8	3.3
25 to 49	21.8	25.4	25.6	17.4	26.3	23.9	18.8
50+	10.7	12.8	5.5	27.0	13.6	20.5	17.5
Females	53.1	50.5	54.8	52.1	49.2	43.9	60.4
<25	15.0	9.4	20.1	3.3	10.5	4.1	6.4
25 to 49	28.4	28.1	27.0	18.4	26.2	19.1	27.5
50+	9.8	13.0	7.7	30.4	12.5	20.7	26.5

Numbers indicated percent with indicator within each latent class
C-CASA: Columbia Classification Algorithm of Suicide Assessment
ICD: International Statistical Classification of Diseases and Related Health Problems

CHAPTER 6: SUMMARY

6.1 Summary of findings

The goal of this research was to examine the occurrence of suicide attempts and associated mental health disorders using administrative health data held in The Repository at the Manitoba Centre for Health Policy (MCHP). The first two papers in the thesis examined the occurrence of mental disorders and suicide attempts in a cohort of individuals born in Manitoba that were living in Winnipeg on their tenth birthday. The thesis also assessed the limitations of relying on administrative data to study the occurrence of suicide attempts in two papers. Specifically it examined the validity of inpatient coding using the Canadian enhancement of the 10th edition of the International Classification of Diseases (ICD-10-CA)[1] for intentional self-harm. Historically a lack of ICD codes from emergency department (ED) visits has prevented research using the Repository from examining the occurrence of attempts presenting to the ED. The following paragraphs briefly describe the main findings from the four projects forming this thesis.

The first paper (**Chapter 2**) estimated the cumulative incidence of diagnosis for 6 mental health disorders after the age of 10; anxiety, mood and adjustment disorders, personality disorders, schizophrenia, and substance use disorders. These incidences were estimated at ages 15, 20 and 25 to show the increase in incidence over adolescence. This paper found that approximately a third of individuals in the sample had a diagnosis of at least one of the six disorders. Mood and adjustment disorders, and anxiety disorders were significantly more common occurring approximately a quarter and a fifth of the sample by age 25. Compared to previous estimates for lifetime incidence of these disorders based on subjective recall of being diagnosed the rates found in **Chapter 2** were high. However, the rates in **Chapter 2** tended to

more closely agree with studies that had used prospectively collected data, and data collected with shorter recall windows. This study underscored the potential dangers of relying on long-term recall and suggests administrative records as an alternative method of epidemiological study.

However, using administrative data for epidemiological research also has limitations. The issue of misdiagnosis of individuals in the data could be causing inflated estimates of incidence. This is potentially aggravated by the occurrence of provisional diagnoses being recorded by physicians. The coding methods employed in this study are based on the detection algorithms used previously by the Manitoba Centre for Health Policy to attempt to remain consistent with prior research. However, prior research has used shorter time periods to detect the occurrence of mental disorders and there might be a higher risk of inflating the incidence of disorders when longer follow-up periods are employed. Comparison between the estimates provided by this study and survey-based estimates should bear in mind that the difference in estimates is also caused by the difference between largely self-report survey prevalence and the treated prevalence that is estimated using administrative data diagnosis. For example, individuals that may have been diagnosed with a mental disorder might not accept that diagnosis and not report this during a survey. Conversely the potential for self-diagnosis of a variety of disorders may also affect survey estimates, regardless of whether these individuals meet any accepted criteria for a mental disorder.

The second paper (**Chapter 3**) examined the occurrence of suicide attempts and death from suicide after the age of 10 in the birth cohort. This study found that approximately 1% of the sample had a recorded admission for a suicide attempt and about 0.1% had died from suicide by age 25. The occurrence of these disorders was considerably higher among those with a

recorded diagnosis of one of the mental disorders. The first year following a new diagnosis with one of these conditions was a period of elevated risk of hospitalization or death after self-harm. The occurrence was particularly high for schizophrenia and personality disorders.

The third paper (**Chapter 4**) examined how well inpatient coding using ICD-10-CA identified individuals admitted to hospital after a suicide attempt or self-harm. Presentations in the clinical dataset known as the Suicide Assessment Form in the Emergency Department (SAFE) were linked to Emergency Department Information System (EDIS) and hospital discharge abstracts. Codes for intentional self-harm, injury of undetermined intent and accidental poisoning were examined. These codes failed to identify the majority of individuals assessed as having made a suicide attempt or present with self-harm. The positive predictive value of the codes was generally high, but not perfect. Attempting to identify only suicide attempts and not self-harm presentations resulted in a lower positive predictive value due to the self-harm presentations also being identified. The EDIS triage system indicated that people presentation with self-harm were most frequently triaged with the chief complaints of 'depression/suicide/self-harm' and 'overdose ingestion'. An optional triage field that specified 'attempted suicide – clear plan' was frequently not used and only a fifth of people with self-harm had this code. Examining the triage complaints of the self-harm group identified by the SAFE data versus the inpatient ICD codes showed that considerable bias might be occurring. One particular issue is the increased occurrence of 'overdose ingestion' complaints in the events identified in the hospital discharge abstract data.

The fourth paper (**Chapter 5**) addressed the issue of a lack of coding methods for suicide attempts that present to the ED but are not admitted to inpatient units. Assessment and treatment of suicidal patients in the ED is an area of research that is growing. It is an important area of

research and treatment since the ED is often the point of contact between at risk patients and medical professionals. This point of contact can potentially direct individuals away from future suicidal behaviour and death from suicide, but due to limitations in our current knowledge of assessment and treatment this is often not the case. This paper combined diagnosis from medical billing claims filled by physicians, inpatient discharge abstracts and triage information from EDIS in a latent class model to attempt to determine whether these data sets provide indicators that can be used to reliable determine which presentations involved self-harm or suicide attempts. Clinical data from SAFE including assessments of suicidality of the presentations were also included to improve the model. The results highlighted the limitations of the newly adopted triaging system. Particularly the triage complaint fields which suffer from insufficiently specific complaint fields, such as depression/suicide/self-harm which inappropriately combines depression and suicidality. Overall the results suggested that the administrative data were not sufficient to reliable identify self-harm presentations without the aid of the psychiatric assessment of presentation status. The SAFE data have limited availability in both time and location which prevents it from being a good source of information to detect self-harm presentations in the province.

6.2 Implications and further research

Previous research has highlighted serious concerns about epidemiological studies that rely on self-reported recall of prior events and diagnosis with mental illnesses.[2] Administrative data are collected regularly as a result of service provision in many health systems and provides an alternative method to obtain epidemiological data. Since these data are routinely collected they do not suffer from recall bias. They also provide a measure of when early clinical interventions can take place for individuals with newly diagnosed mental disorders. The effect of recall bias on

long-term estimates is likely quite significant and many studies used to inform decisions on mental health treatment are likely to underestimate the lifetime incidence.

The data from this thesis maps out the incidence of new diagnosis for six mental disorders. This information is useful for planning health system services that target youth. For instance, these data could prove invaluable in planning early intervention services for youth with newly diagnosed conditions. However, the methods of detecting cases in this thesis were developed to be used over short periods of time and their validity over long-term periods should be examined in future research. In particular, due to the nature of the diagnosis of new disorders often involving provisional diagnoses the validity of using additional claims and hospitalizations should be examined to determine whether this improves the validity of coding cumulative long-term incidence. The risk of suicidal behaviour is high in the period after a new diagnosis (as seen in **Chapter 3**).[3] Given the success of early intervention services for youth diagnosed with psychotic disorders at reducing health service use and suicidal behaviours,[4–6] the development of additional early intervention services for this population is a potentially worthwhile area of clinical program development and research.

However, this study also provided significant implications for research that uses administrative data to detect suicide attempts and deaths. Although the validity of codes may vary across different administrative data repositories the validation results from this study suggest that there is likely undercounting of suicide attempts in these data. The actual occurrence of suicide attempts that results in hospitalization is likely approximately double the estimates provided. The bias this underestimation will have will vary depending on how the variable/outcome being constructed. Rates of hospitalization are likely to be underestimated by half since they require measuring all instances. Measures of cumulative incidence will also be

underestimated, but since individuals are often hospitalized several times during high risk periods some of the missed individuals will be identified on subsequent hospitalization. This will result in less undercounting than rates, but the incidence of attempts estimated based on the data will be shifted later in time (since the second or third attempt is being recorded as the first). Researchers should consider the limitations of using administrative data in their specific study and ensure that these limitations are addressed as best as possible in their design and are also thoroughly detailed as limitations within any written report.

These results also strongly indicate that there are likely systematic differences between those that are identified using ICD codes and those that are not. It is likely that the most medically serious occurrences of suicidal behaviour are more likely to be identified. This does not mean that all research using these data sources to identify suicide attempts is inherently wrong and without merit. This does suggest that this research should be assessed knowing that the cohort of identified attempts is likely the most medically severe instances and that there is another group of patients that are clinically different that will be missed.

Suicidal thoughts and behaviours are serious medical issues and improvements in documentation by physician and medical codes should be undertaken. Standardized assessments such as the Columbia Classification Algorithm of Suicide Assessment (C-CASA) should be performed by emergency physicians and noted in emergency room documentation for all individuals being assessed for mental health issues. The EDIS system is currently inadequate at classifying individuals with mental health and suicidal thoughts and behaviours. The reliance on an optional field to delineate between the occurrence suicidal ideation, suicidal behaviour, and neither of these within the most commonly used triage fields is a serious limitation. The mental health triage fields used by Manitoba's version of EDIS should be reviewed by mental health

professionals and improved classifications should be developed that allow a reasonable degree of precision within the framework of relying on triage nurses that do not have specialized training in diagnosis of mental health conditions. Once these adjustments have been made more research can be conducted to assess the validity of using the new triage system to identify presentations involving self-harm.

Other sources of administrative data are becoming available that might also provide information useful to epidemiological research on self-harm and mental health. One of particular interest in Manitoba is the development of electronic medical records as part of the Physician Integrated Network. Standardized electronic medical records can prove to be vastly superior to physician billing claims since they can potentially provide more thorough diagnosis of disorders. Individuals with high comorbidity might be better described in these electronic medical records than in the single available ICD-9 code from the medical claims record. This is likely to increase the sensitivity of the data when used to detect mental disorders, but increased data availability could also allow more strict coding structures to also improve specificity. However, the move towards electronic health records is a recent trend worldwide and is therefore limited with respect to long-term studies

6.3 References

1 CIHI. The Canadian Enhancement of ICD-10. 2001. Canadian Institute for Health Information; Ottawa, ON.

2 Patten SB. Recall bias and major depression lifetime prevalence. *Soc Psychiatry Psychiatr Epidemiol* 2003; **38**: 290–6.

3 Randall JR, Walld R, Finlayson G, Sareen J, Martens PJ, Bolton JM. Acute risk of suicide and suicide attempts associated with recent diagnosis of mental disorders: a population-based, propensity score-matched analysis. *Can J Psychiatry* 2014; **59**: 531–8.

4 Randall JR, Vokey S, Loewen H, Martens PJ, Brownell M, Katz A, *et al.* A Systematic Review of the Effect of Early Interventions for Psychosis on the Usage of Inpatient Services. *Schizophr Bull* 2015. doi:10.1093/schbul/sbv016.

5 Harris MG, Burgess PM, Chant DC, Pirkis JE, McGorry PD. Impact of a specialized early psychosis treatment programme on suicide. Retrospective cohort study. *Early Interv Psychiatry* 2008; **2**: 11–21.

6 Randall JR, Chateau D, Smith M, Taylor C, Bolton J, Katz L, *et al.* An early intervention for psychosis and its effect on criminal accusations and suicidal behaviour using a matched-cohort design. *Schizophr Res* 2016; **176**: 307–11.

APPENDIX A: ETHICS APPROVAL

UNIVERSITY OF MANITOBA | BANNATYNE CAMPUS Research Ethics Board

P126 - 770 Bannatyne Avenue
Winnipeg, Manitoba
Canada R3E 0W3
Telephone 204-789-3255
Fax 204-789-3414

HEALTH RESEARCH ETHICS BOARD (HREB)
CERTIFICATE OF FINAL APPROVAL FOR NEW STUDIES
Delegated Review

PRINCIPAL INVESTIGATOR:	INSTITUTION/DEPARTMENT:	ETHICS #:
Mr. Jason Randall	U of M and MCHP/Community Health Sciences	H2015:009

APPROVAL DATE:	EXPIRY DATE:
January 8, 2015	January 8, 2016

STUDENT PRINCIPAL INVESTIGATOR SUPERVISOR (If applicable):
Dr. L. Roos

PROTOCOL NUMBER:	PROJECT OR PROTOCOL TITLE;
N/A	Validation Of Administrative Data For Self-Harm Outcomes Using A Clinical Database

SPONSORING AGENCIES AND/OR COORDINATING GROUPS:
Evelyn Shapiro Award

Submission Date of Investigator Documents:	HREB Receipt Date of Documents:
December 17, 2014	December 22, 2014

THE FOLLOWING ARE APPROVED FOR USE:

Document Name	Version(if applicable)	Date
Protocol: Proposal		December 17, 2014
Consent and Assent Form(s):		
Other: Data Fields List to be used from Administrative Databases		submitted December 17, 2014

CERTIFICATION
The above named research study/project has been reviewed in a ***delegated manner*** by the University of Manitoba (UM) Health Research Board (HREB) and was found to be acceptable on ethical grounds for research involving human participants. The study/project and documents listed above was granted final approval by the Chair or Acting Chair, UM HREB.

HREB ATTESTATION
The University of Manitoba (UM) Research Board (HREB) is organized and operates according to Health Canada/ICH Good Clinical Practices, Tri-Council Policy Statement 2, and the applicable laws and regulations of Manitoba. In respect to clinical trials, the HREB complies with the membership requirements for Research Ethics Boards defined in Division 5 of the Food and Drug Regulations of Canada and carries out its functions in a manner consistent with Good Clinical Practices.

QUALITY ASSURANCE
The University of Manitoba Research Quality Management Office may request to review research documentation from this research study/project to demonstrate compliance with this approved protocol and the University of Manitoba Policy on the Ethics of Research Involving Humans.

CONDITIONS OF APPROVAL:

1. The study is acceptable on scientific and ethical grounds for the ethics of human use only. ***For logistics of performing the study, approval must be sought from the relevant institution(s).***
2. This research study/project is to be conducted by the local principal investigator listed on this certificate of approval.
3. The principal investigator has the responsibility for any other administrative or regulatory approvals that may pertain to the research study/project, and for ensuring that the authorized research is carried out according to governing law.
4. **This approval is valid until the expiry date noted on this certificate of approval. A Bannatyne Campus Annual Study Status Report** must be submitted to the HREB within 15-30 days of this expiry date.
5. Any changes of the protocol (including recruitment procedures, etc.), informed consent form(s) or documents must be reported to the HREB for consideration in advance of implementation of such changes on the **Bannatyne Campus Research Amendment Form.**
6. Adverse events and unanticipated problems must be reported to the HREB as per Bannatyne Campus Research Boards Standard Operating procedures.
7. The UM HREB must be notified regarding discontinuation or study/project closure on the **Bannatyne Campus Final Study Status Report**.

Sincerely,

[redacted]

Chair, Health Research Ethics Board
Bannatyne Campus

Please quote the above Human Ethics Number on all correspondence.
Inquiries should be directed to the REB Secretary **Telephone:** (204) 789-3255/ **Fax:** (204) 789-3414