The Epidemiology of Treatment-Resistant Depression in Manitoba: A Retrospective Cohort Study Using Administrative Health Data

By

Michael Paillé

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Community Health Sciences

University of Manitoba

Winnipeg

Abstract

Major depressive disorder (MDD) is a condition that is difficult to treat, as many individuals do not experience complete remission, and many eventually relapse.[1] Individuals who do not respond to antidepressant therapy are defined as having treatment-resistant depression (TRD). The epidemiology of TRD has been studied using commercial insurance claims data in the United States, but no such studies exist for a general population or in Canada.

This thesis describes the epidemiology of TRD in Manitoba between 1996 and 2016 and compares the risks for ambulatory visits, emergency department (ED) visits, hospitalizations, and all-cause mortality for individuals with TRD to those with non-treatment-resistant MDD. I used the Anderson-Gill generalization of the extended Cox proportional hazards model to analyze these risks. TRD was defined using the Massachusetts General Hospital Staging Method (MGH-s), where scores of 2.5 and higher correspond to TRD.

I identified 169,511 adults living in Manitoba between 1996 and 2016 who were diagnosed with MDD and dispensed at least six weeks of antidepressants from the Manitoba Population Research Data Repository at the Manitoba Centre for Health Policy. By 2016, 18,663 individuals (11.0% of the cohort) met the criteria for TRD. Compared to MGH-s scores of 1, the hazard ratios (HR) for mental health related ED visits for MGH-2.5 and MGH-3 were 3.9 (95%CI:3.3–4.3) and 5.6 (95%CI:5.1–6.0), respectively. The HR for hospitalizations with a primary diagnosis of MDD were 4.9 (95%CI:4.3–5.4) and 8.3 (95%CI:7.7–9.0) for MGH-2.5 and MGH-3, respectively. For individuals 10 years below the average age of the cohort, TRD was associated with a three-fold increased risk for all-cause mortality. For individuals 10 years above the average age of the cohort, TRD was associated with a two-fold increased risk for all-cause mortality.

These findings show that TRD is an important risk factor for requiring more intensive medical care for mood and anxiety disorders. Increasing hazards associated with increases in MGH-s scores support the hypothesis that TRD follows a severity continuum. As the risk for all-cause mortality was higher for individuals with TRD, further research is needed to determine whether TRD is associated with a higher risk for suicide.

Acknowledgments

The completion of this thesis would not have been possible without the support of a number of people, the first being my advisor, Dr. Nathan Nickel. Dr. Nickel was a constant source of encouragement who offered valuable mentorship beyond what I could have even asked for.

The members of my advisory committee, Dr. Jennifer Hensel and Dr. James Bolton have continuously offered invaluable clinical insight for which I am most grateful. Attempting this thesis without their support would have been impossible.

A major component of this thesis involved learning a new language, SAS. For this, I must thank Mr. Charles Burchill for his continued guidance, utmost patience, and excellence in teaching.

Dr. Malcolm Doupe, along with the generous financial support of the Manitoba Training Program for Health Services Research, offered me the chance to question the importance of this thesis and ensure that it would improve the health of Manitobans. For this, I am very thankful.

Dr. Dan Chateau, who encouraged me to advance my expertise in pharmacoepidemiology, along with the financial support of the Canadian Network for Observation Drug Effect Studies, ensured that I was able to approach this thesis with a diligent and thoughtful mind. Thank you.

The entire staff of the Manitoba Centre for Health Policy ensured that this centre was a place I could call home and ask for help. Receiving the generous funding offered to me through the Evelyn Shapiro Award for Health Services Research was a tremendous honour. I cannot thank you all enough.

I am also incredibly thankful for the financial support awarded to me through Research Manitoba, from the Government of Manitoba. Supporting the advancement of science must continue to be celebrated.

Finally, I must thank my incredibly supportive partner, Alix, my dear friends, and my loving family. I would not have managed without you all.

The authors acknowledge the Manitoba Centre for Health Policy for use of data contained in the Manitoba Population Research Data Repository under project #2019-003 (HIPC#2018/2019 - 47). The results and conclusions are those of the authors and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, or other data providers is intended or should be inferred. Data used in this study are from the Manitoba Population Research Data Repository housed at the Manitoba Centre for Health Policy, University of Manitoba and were derived from data provided by Manitoba Health and the Winnipeg Regional Health Authority.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIC.....Akaike information criterion
ATC.....Anatomical therapeutic chemical
ATHF .....Antidepressant Treatment History Form
BAD .....Bipolar affective disorder
CI.....Confidence Interval
DDD.....Defined daily dose
DPIN .....Drug Product Information Network
ECT .....Electroconvulsive therapy
ED .....Emergency department
EDIS.....Emergency Department Information System
ESM .....European Staging Model
HR .....Hazard Ratio
ICD.....International Classification of Diseases
MAOI..... Monoamine oxidase inhibitor
MCHP ..... Manitoba Centre for Health Policy
MDD .....Major depressive disorder
MGH-s .....Massachusetts General Hospital Staging Model
MH .....Mental health
MSM ..... Maudsley Staging Model
OECD..... Organisation for Economic Cooperation and Development
OR .....Odds Ratio
PTD .....Pharmaceutically treated depression
PYPV .....Past year psychiatrist visit
SD .....Standard deviation
SEFI .....Socioeconomic factor index
STAR*D ..... Sequenced Treatment Alternatives to Relieve Depression
TRD.....Treatment-resistant depression
TRSM..... Thase Rush Staging Model

Introduction

Major depressive disorder (MDD) is a complex psychiatric condition characterized by low moods, feelings of guilt, worthlessness or helplessness, sleep and cognitive changes, and other debilitating symptoms.[2] In turn, individuals with MDD experience higher rates of health care use[3,4] and premature mortality[5,6] compared to individuals without MDD. In Canada, the lifetime prevalence of MDD is approximately 10%, with an annual prevalence of 4%, a value that is consistent across Canadian provinces.[7] Of survey respondents with MDD, 63% reported having seen a mental health care professional in the past year, while 33% reported taking an antidepressant in the two days before the survey.[7] However, these values must be interpreted with some caution as they are based on self-reported survey data, which may result in misclassification bias and recall bias for both MDD status and pharmaceutical use.

A more reliable source of data when studying health services use is administrative data. One study from Quebec found that 75% of those who were recorded as having used mental health services in the province's administrative data registry did not report using these services in the Canadian Community Health Survey.[8] Many factors may explain this, such as a social desirability bias to under-report mental illness; therefore, one could argue that survey data may not provide an adequate description of the epidemiology of MDD. Linked administrative health data provide an alternative means for identifying individuals with MDD for the purposes of epidemiological research.

The Manitoba Population Research Data Repository (the Repository) at the Manitoba Centre for Health Policy (MCHP) includes administrative data on virtually all health services provided through the provincial healthcare system.[9] The Repository has been used extensively to study a variety of health topics in Manitoba, including mental health. One recent study, *Mental*

Illness Among Adult Manitobans, found that the five-year diagnostic prevalence of mood and anxiety disorders in Manitoba, including MDD, was 23%.[10] With such a high proportion of Manitobans experiencing mental health problems, improving our understanding of how mental health conditions like MDD impact health services use and mortality must be a high priority for this province.

As many individuals experience MDD, understanding the effectiveness of the treatments for this condition is essential. Similarly, understanding the impacts of a failure to respond adequately to treatments is of the utmost importance in improving outcomes associated with MDD. Treatments for MDD can include pharmaceuticals, psychosocial interventions, non-pharmaceutical medical treatments, and alternative therapies. According to annual reports from the Organization for Economic Cooperation and Development (OECD), Canadians have long been one of the highest users of antidepressants compared to other OECD countries.[11,12] Among all OECD countries, the average number of dispensed defined daily doses (DDD) for antidepressants in 2015 was 60 DDD per 1,000 people per day, where a DDD is the assumed average maintance dose per day for a medication.[12] In the same year, Canadians were ranked as the third highest user of antidepressants globally, with 90 DDD dispensed per 1,000 people per day, up from 86 DDD per 1,000 people per day four years prior.[11,12] Overall, Canadians are highly reliant on antidepressants for their treatment of MDD, which underlines the importance of studying the impacts of failing to respond to antidepressant therapies.

There is no cure for MDD, and the effectiveness of the treatments – that is, pharmaceutical treatments – for MDD remains a research priority. One seminal study on the treatment of MDD is the Sequenced Treatment Alternatives to Relieve Depression (STAR*D)

study,[1] which investigated the remission[a] and response[b] rates of 3,671 patients using 22 different treatments options, administered through five sequential rounds of treatment. After two rounds of treatment, half of the patients still exhibited symptoms of depression.[1] After a 12-month follow-up of those who entered remission in the first and second rounds of treatment, 40% and 55% experienced a relapse of depression symptoms, respectively.[1] For individuals who did not respond after two rounds of treatment, rates of response in subsequent rounds were much lower; only 16% of individuals who entered these rounds responded to treatment.[1] Overall, STAR*D showed that MDD is a disorder that is difficult to treat, with low rates of remission and response, and high rates of relapse.

Patients with MDD who do not respond to treatment are identified as having treatment-resistant depression (TRD). While not a formal diagnosis in and of itself, studies have investigated the treatment of TRD for over 40 years. However, research into the epidemiology of TRD, including discussions on how best to define TRD, only started in the early 2000s. There are, therefore, still many knowledge gaps in the current literature. For example, the epidemiology of TRD has yet to be studied in Canada, nor has any study investigated the impact of socioeconomic status or urban status on TRD. Some reasons for these gaps include the difficulties in accurately measuring lifetime pharmaceutical use, the lack of data from a general population, as well as the difficulties in conducting large scale, prospective clinical trials like STAR*D.

The primary limitation in the research on TRD, though, is a lack of consensus on its definition, which in turn leads to studies that are not comparable.[13] In 2017, Conway and

[a] Remission was defined as having a Quick Inventory of Depression Symptomology Self-Report 16 score less than or equal to five.
[b] Response was defined as a 50% reduction in depression severity score.

colleagues proposed that researchers must agree on a definition of treatment-resistance and that it should be one that emphasizes a staging method approach.[13] With these limitations and knowledge gaps in mind, the following review of the literature focuses primarily on how the definition of TRD has evolved and offers insights into the advantages and disadvantages of the various definitions that have been used.

Review of the Literature

Defining Treatment-Resistant Depression

In 1991, Jan Scott was the first to review the literature regarding the definitions used for treatment-resistant depression.[14] Scott identified that there was no standard definition being used to differentiate the many sub-types of depression that researchers and clinicians described at that time. He also found the term "treatment-resistant" was problematic, as new treatments for depression were being developed and had not yet been adequately tested. Instead, he concluded that the term "chronic depression" was more appropriate. He argued that these "chronic depression" patients either received inadequate or inappropriate treatment, had an illness that could not be treated by contemporary therapies, or had a form of a depression that was "characterological in nature", which Scott described as being akin to the DSM-III diagnosis of dysthymic disorder or chronic minor depressive disorder.[14]

Scott's study was followed up in 1996 by Fava and Davidson, who were the first to propose an operational definition for TRD.[15] This first definition for TRD was a failure "to respond to standard doses of antidepressants administered continuously for a minimum duration of six weeks."[15] Fava acknowledged that this definition might be too "liberal" for other researchers, but suggested that a timeframe longer than six weeks would pose ethical challenges, such as leaving non-responding patients without adequate treatment for too long. It must be

noted that this definition was not created to inform research using administrative data, but instead to inform smaller-scale, clinical research trials for individuals with prior non-response. At the time, accurately defining TRD was crucial in ensuring that clinical trials could account for a patient's prior treatment exposure, which is a significant confounder in measuring rates of treatment response.[1] This is especially true in trials that focused on individuals with some history of non-response to treatment. This concern for accurately measuring prior non-response to antidepressants encouraged the publication of other TRD definitions, and eventually TRD staging methods.

In 1999, Souery and colleagues reviewed the literature of clinical trials regarding TRD and summarized the differences.[16] The authors identified six factors that caused difficulties in achieving consensus on a definition of TRD: (1) necessary exclusion, (2) adequacy of treatment (i.e. dosage and duration of treatment), (3) treatment response measures and cut-offs, (4) number and type of failed trials (i.e. changing medication classes or not), (5) terminology and staging of resistance (i.e. accounting for different severities of treatment-resistance), and (6) accounting for predictive factors and confounding variables.[16] Discrepancies regarding these six factors were – and still are – ubiquitous in TRD literature. No consensus has been reached on a definition for TRD, which poses a challenge in comparing the findings from research on the subject.[13] Still, Souery and colleagues described their definition of TRD, referred to as the European Staging Model, as: a "failure to respond to two adequate trials of different classes of antidepressants (i.e. consecutive treatments with two different antidepressants, each given separately in an adequate dosage for a period of six to eight weeks)."[16]

When researchers like Fava and Souery first presented these *binary* definitions of TRD (i.e. either a patient has TRD or they do not), other research focused on determining valid staging

methods for TRD. The idea of staging TRD is akin to the staging system used in cancer diagnoses, though is less developed and validated. The goal of a staging method is to create a severity-hierarchy for TRD that would accurately predict rates of response, remission and relapse. Similar to the binary definitions of TRD, *many* different staging methods were developed. Each staging method uses a unique system to rate a patient's level of treatment-resistance either by the number or type of antidepressants used or by the duration of treatment. These staging methods are described and compared in the following section.

Staging Methods

When research and development of new antidepressants was accelerating the late 1980s and 1990s, researchers recognized a need to control for previous exposure to antidepressants during clinical trials. Researchers found that, for example, a patient who fails to respond to only one antidepressant has a different prognosis than does a patient who fails to respond to three.[1] As the number of failed antidepressant trials for a patient increases, the likelihood of achieving long-term remission generally decreases, and the likelihood of relapse increases.[1] Accurately defining these discrete stages of treatment-resistance among clinical trial participants poses a challenge to researchers who test the efficacy of antidepressants on individuals with various prior histories of treatment-resistance, since misclassification of treatment-resistance could bring bias into the study through improper recruitment.[17]

To address this challenge of confounding by misclassification of treatment-resistance, researchers developed well-defined staging methods. These methods included criteria that attribute points to individuals based on factors such as the number of failed antidepressant trials, dosage increases, augmentation therapies, the antidepressant class used, the duration of treatment, and the use of electroconvulsive therapy (ECT). The overall goal of each staging

model was to develop a system that could accurately and validly predict response, remission, and relapse, and ensure that all clinical trial participants are being evaluated and grouped correctly.[18]

Ruhé and colleagues published a systematic review that compared the five staging methods developed by 2012, described their respective strengths and limitations, and summarized the results of any published validation studies. The five staging methods were: the Antidepressant Treatment History Form (ATHF 1990/1999), the Thase Rush Staging Model (TRSM 1997), the European Staging Model (ESM 1999), the Massachusetts General Hospital Staging Model (MGH-s 2003), and the Maudsley Staging Model (MSM 2009). As only some of these models are relevant to epidemiological research using administrative data, I will not describe them all in detail in this section.

Thase and Rush Staging Model

The TRSM is a five-stage model that includes two factors in TRD staging: the number of failed therapies and the hierarchy of antidepressants. The TRSM does not define an adequate dose or duration of treatment and does not include augmentation or combination therapies, which are defined as either using non-antidepressant medications concurrently with antidepressants or using multiple antidepressants concurrently, respectively. Also, the TRSM requires that different classes of antidepressants be tried in order to advance to the next stage, and does not recognize in-class switches to be relevant to TRD staging. The five stages of the TRSM are summarized in Table 1.

Table 1. The Five Stages of the Thase Rush Staging Model

Stage I	Failure of at least 1 adequate trial of 1 major class of antidepressants
Stage II	Failure of at least 2 adequate trials of at least 2 major classes of antidepressants
Stage III	Stage II resistance plus failure of an adequate trial of tricyclic antidepressants
Stage IV	Stage III resistance plus failure of an adequate trial of monoamine oxidase inhibitors
Stage V	Stage IV resistance plus a course of bilateral electroconvulsive therapy

Ruhé described the primary strength of the TRSM as its simplicity of use for clinicians.[18] However, its main limitations are its lack of definition of treatment intensity (i.e. dosage and duration), as well as its use of a hierarchy of antidepressants despite any evidence supporting such a hierarchy.[18] The predictive value of the TRSM had not been systematically assessed, leading to concerns about its value in future research.[18]

European Staging Model

The ESM was published by Souery and colleagues and was developed two years after TRSM, attempting to build on its limitations.[16,18] The ESM can be categorized as a pseudo-binary definition of TRD because the stages are based on the number of weeks that a patient undergoes antidepressant trials, rather than on the number of antidepressants that were tried. For example, a patient who is maintained on their second antidepressant medication for 16 weeks would be deemed to be at the same stage as a patient who took a second and third medication for eight weeks each.

This lack of differentiation between treatment duration and number of failed antidepressants is problematic as it does not account for the true number of medications that were tried by an individual. This would be an issue in studies using administrative data, as it would be impossible to differentiate an individual who is responding to treatment and simply

maintaining an antidepressant treatment compared to an individual who continues taking the same medication for several weeks without any response to treatment. Simply, the TRD definition of the ESM is a failure to respond to two adequate trials of different classes of antidepressants (i.e. consecutive treatments with two different antidepressants, each given separately in an adequate dosage for a period of six to eight weeks), where the stage is increased for every eight weeks on a medication.[16]

Like the TRSM, the ESM has limitations. For one, it requires that more than two *different classes* of antidepressant be trialled in order to meet the definition of TRD, without any evidence to support the validity of this requirement.[18] Additionally, the ESM does not include augmentation therapy as a relevant criterion until more than a year of antidepressant trials has elapsed, despite evidence supporting its use.[18] The ESM's main strengths over the TRSM is that it does not set an arbitrary hierarchy of antidepressants, and has clear definitions of non-response, using the Hamilton Depression Rating Scale or Montgomery Asberg Depression Rating Scale.[18] No studies have been published on the predictive value of the ESM.

Massachusetts General Hospital Staging Model

The MGSH-s was developed by Maurizio Fava in 2003 and was the first staging method to use a points system to determine a patient's TRD stage. Points are given for three reasons, described in Table 2:

Table 2. Massachusetts General Hospital Staging Model Criteria

Criteria	**Points**
Nonresponse to each adequate (six weeks of an adequate dosage of an antidepressant) trial of a marketed antidepressant	1 point per trial
Optimization (i.e. increase) of dose, optimization of duration, and augmentation or combination of each trial	0.5 points per optimization or augmentation
Electroconvulsive therapy	3 points

The MGH-s is built upon the limitations of the previous two staging methods by accounting for dosage, prolonged duration, augmentation therapies, having no hierarchy of treatment, and no preference for between-class or within-class switches.[18] The main limitations of the MGH-s are that the points attributed to each criterion appear to be arbitrary and that points have no maximum score.[18] Petersen and colleagues compared the TRSM and the MGH-s in terms of their ability to predict non-remission.[19] They showed that a higher MGH-s score succeeded in predicting non-remission (OR=3.7; 95% CI=1.5–9.4), while the TRSM did not (OR=1.9; 95% CI=0.7–5.4).[19] This finding suggests that the MGH-s is superior to the TRSM in staging TRD, at least in terms of its ability to predict non-remission, an important characteristic of TRD.

Overall, Ruhé and colleagues were the first to provide a systematic review of the many staging methods that have been used for TRD but provided no conclusion for which staging method was superior.[18] The authors showed how the staging methods for TRD evolved by including improvements in the treatment of depression, both in terms of effectiveness and quantity of treatment options. While researchers have used these staging methods more

frequently in clinical studies than epidemiological ones, they show the possibility and importance of *not* defining TRD as a binary disorder.

Conway et al., 2017

In their 2017 letter published in *JAMA Psychiatry*, Conway, George and Sackeim propose another staging method for TRD that is based on the findings of STAR*D. A total of three stages are included in this definition, non-TRD, Stage I TRD, and Stage II TRD. Stage I TRD is defined as a "failure of two adequate dose-duration antidepressants or psychotherapy from a different class (either in combination or succession) in the current episode," and Stage II is the same, but requires three or more failures.[13] In response to this proposed definition, Fogelson and Leuchter note that its requirement for trying different classes of antidepressants is not based on empirical evidence.[20] In addition to this, the definition proposed by Conway and colleagues includes the use of psychotherapy, which has not been included in any previous research on the epidemiology of TRD using administrative data. There is also no previous research that has used or validated this definition when used with administrative data.

When reviewing the literature of the epidemiology of TRD using administrative and survey data, researchers seldom used these staging methods. Epidemiological studies have instead favoured a binary approach to defining TRD. One reason for this is likely the simplicity of using a binary definition, both for the data management of a study cohort, as well as for the statistical analysis. Another reason is neither the ESM nor the MSM can be used in epidemiological research using administrative data. The ESM cannot be used due to the inability of administrative data to differentiate between antidepressant maintenance therapy where a patient has responded to treatment and ongoing antidepressant therapy where a patient has not responded to therapy. The MSM, which has not described above, cannot be used with some

sources administrative data as it requires detailed knowledge of a patient's depression severity. The following sections describe how research into the epidemiology of TRD has attempted to overcome these challenges, and what other definitions were created.

First binary definition with administrative data

The first definition of TRD used in a cohort study using administrative data was described by Corey-Lisle and colleagues and is characterized by three dimensions of antidepressant treatment: using specified treatments, the number of drug switches, and the number of dose titrations.[21] The first criterion for TRD was to have received either ECT or to have been prescribed a monoamine oxidase inhibitor (MAOI). If a patient did not meet this criterion, they would be categorized as having TRD only if they met both the "TRD-scale criteria" and the "TRD-matrix criteria". The exclusion criteria for Corey-Lisle included: age less than 18 years or greater than 64 years, or a diagnosis of psychotic disorders, bipolar disorder, schizophrenia, or dementia.

The TRD-scale criteria included factors such as receiving augmented antidepressant therapy, switching medications, or increasing the dosage of a medication. Having a score of at least five for the TRD-scale criteria would be categorized as meeting the TRD-scale criteria. The TRD-matrix criteria can be described as a four-by-three matrix with "Number of Switches" (0, 1, 2, 3+) and "Number of Titrations" (0, 1, 2+) as the axes of the matrix. Having a score of 3+ for the number of switches, or 2 switches and 2+ titrations, would be categorized as meeting the TRD-matrix criteria. No minimum amount of time on an antidepressant was required as part of the TRD-likely criteria.

Corey-Lisle and colleagues' approach classified 12% of the cohort as having TRD. Females were more likely to be characterized as having TRD, and no clinically relevant difference in age was found.[21] Over one year, individuals with TRD used 1.9 times as many services (office visits, and inpatient and outpatient care) compared to individuals without TRD. However, there was no statistically significant difference in inpatient care.

Though this study was the first to analyze rates of health service use between individuals with and without TRD, the data came from the commercial health insurance claims from a single workplace. These data are, therefore, not representative of a general population, and the findings cannot be generalized. However, this study paved the way for future researchers to begin studying the epidemiology of TRD using administrative data.

Comparing definitions

The studies of Crown et al. and Russell et al. were associated with the same cohort, but their definitions for TRD were different. Both studies defined TRD as switching pharmaceutical therapies at least twice, or having a hospitalization recorded for depression and at least one antidepressant prescription. Crown and colleagues required that the patient be prescribed an antidepressant for at least eight weeks, while Russel and colleagues only required four weeks of antidepressant treatment. The authors included all drug switches between 1995 and 2000 as counting towards the definition of TRD. Both Crown et al. and Russell et al. included patients with bipolar disorder and all those over the age of 18 but excluded patients with schizophrenia and psychotic disorders. Russell also expanded the exclusion criteria to exclude patients diagnosed with paranoid states, other nonorganic psychoses, psychoses of childhood, Alzheimer's disease, Parkinson's disease, mental retardation, and senility without psychosis.

Neither study reported the proportion of individuals in the cohort who were categorized as having TRD.

In a conference abstract, Gibson and colleagues compared the methods of Corey-Lisle et al., Crown et al., Russel et al., and the MGH-s in identifying cases of TRD in administrative data.[22] Over two-thirds of patients were classified similarly across all four methods.[22] Gibson and colleagues concluded that using a "gradient" for TRD (i.e. a staging method) in epidemiological research would provide clinicians and insurers more useful information.[22] As this work was not published elsewhere, it is impossible to draw any further conclusions.

Three years after this publication, Gibson and colleagues published a study on the cost burden of TRD.[23] In this study, the authors used the MGH-s score of 3 as a cut off for defining TRD, as this corresponds with at least two failed antidepressant trials and requires either a third antidepressant trial failure or an augmentation/combination therapy attempt. This was the first epidemiological study to evaluate the impacts of TRD in a non-binary way and showed the utility of this method by measuring the health care costs associated with each MGH-s point increase. As the costs associated with TRD increased linearly with each point increase in MGH-s score, the authors demonstrated that staging methods could predict the outcome severity and health care utilization of a TRD patient more precisely than a binary definition.

However, there are many limitations in this study by Gibson and colleagues. The researchers defined the dose and duration optimization criteria as *any* increase in dose and *any* treatment regimen with at least three refills for a medication.[23] This liberal definition of dose and duration optimization likely resulted in many individuals experiencing multiple half-point increases in the study when it may not have been appropriate. Additionally, each combination antidepressant therapy increases the score by a half-point for the MGH-s. The researchers did not

include this criterion in their methods but instead treated all antidepressant dispensation as being worth a full-point. Arguably, Fava was not clear in his initial definition of combination therapies in 2003.[17] However, in the book, *Massachusetts General Hospital Comprehensive Clinical Psychiatry*, co-authored by Fava, the combination of any two antidepressants with different mechanisms of actions are treated as combination therapies.[24] This means that an antidepressant dispensed at the same time as a previously dispensed antidepressant should only contribute a half-point, rather than a full-point. Overall, it is likely that Gibson and colleagues applied the MGH-s in such a way that it over-estimated the true proportion of TRD among individuals with pharmaceutically treated depression (PTD).

For their definition of TRD, Ivanova and colleagues used two different approaches, that of Corey-Lisle and colleagues,[21] as well as that of "at least three antidepressant trials of adequate dose and duration (at least six weeks)".[25] Like Corey-Lisle and colleagues, Ivanova and colleagues excluded patients with psychosis, schizophrenia, bipolar disorder, and dementia, and only included patients aged 18-64. The reason for this age restriction was to account for the different insurance programs available to seniors in the United States.[25] The Corey-Lisle approached yielded a TRD prevalence among those with MDD of 8.96%, while the other approach yielded a TRD prevalence of 12.93%. The prevalence of TRD for individuals who met the criteria for either approach was 14.50%.

In terms of health-related outcomes, the TRD group had higher rates of mental-health comorbidities (including anxiety disorders and substance use disorders), higher Charlson Comorbidity Index scores, and higher rates of health care use. Overall, the direct health care cost associated with TRD was nearly twice that of MDD. As these data are from a private insurance database, these findings cannot be generalized to a whole population.

TRD and health services use

Olchanski and colleagues used data from PharMetrics Patient-Centric database between 2001 and 2009, representing 17 million people from more than 95 different health plans in the United States.[26] The majority (90%) of those in this database were insured through private plans.[26] In contrast to other studies, Olchanski and colleagues defined TRD as having at least four distinct medication therapies for depression, compared to the more common criterion of only three distinct therapies. Augmentation therapies included: atypical antipsychotics, dopaminergic medications, stimulants, lithium, or thyroid hormone. Each drug trial was only defined as adequate it if was prescribed and refilled for a minimum of 90 days (12 weeks), the longest trial duration used in TRD literature.

This divergence from standard definitions of TRD may be seen as an improvement to the specificity defining TRD but adds to the concerns of comparability with other research while also artificially lowering the reported prevalence of TRD. There is another concern for including patients who met the standard TRD definition in the comparison group, which would further bias the results towards the null. Also, the authors did not use an exclusion criteria for psychiatric comorbidities and only included patients greater than or equal to 21 years of age. This lack of exclusion criteria raises concerns for the specificity of this cohort since augmentation therapies are more typically prescribed to patients with schizophrenia and bipolar disorder who likely have different health care needs than patients with MDD. This would, in turn, lead to results that are skewed away from the null, limiting their validity.

The prevalence of TRD among patients with a diagnosis of MDD was 29.4%, the highest found in epidemiological research on TRD. The rate of depression-related medical service was 68% higher in the TRD group compared to the MDD group, while all other medical service use

was 44% higher, without adjusting for other factors. The corresponding differences in depression-related medical costs were 88.6% higher in the TRD group, and all other medical service costs were 62.7% higher. These results show, again, that TRD patients have more intensive health care needs than general MDD patients.

The study by Olfson and colleagues used data from the Truven Health Analytics MarketScan Medicaid Database between 2008 and 2014 and included data from 8-12 million Medicaid enrollees.[27] Medicaid is an American program that provides health care funding for low-asset individuals, and so this dataset is not generalizable to an entire population. Patients were excluded from this study if they were not between the ages of 18 and 64, or if they received a diagnosis for bipolar disorder, schizophrenia, psychosis, or dementia. Patients were followed for the 12 months before antidepressant initiation and 24 months after. The authors defined TRD as initiating a third adequate antidepressant trial within 12 months of the index antidepressant prescription. An adequate trial was defined as one that lasted at least six weeks. Augmentation therapy included anticonvulsants, antipsychotics, lithium, psychostimulants, or thyroid hormone.

The prevalence of TRD among PTD patients was 25.9%. The TRD cohort was more likely than the non-TRD cohort to have been diagnosed with a substance use disorder, to have received inpatient mental health treatment, and to have an elevated Elixhauser comorbidity score during the year before the index date.[27] Patients with TRD were significantly more likely than patients without TRD to receive inpatient care for any reason (31.0% vs. 21.6%), as well as for a mental health condition (12.7% vs. 7.6%).[27] Annual health care costs for TRD patients were $6,765 (USD) higher than those without TRD in the first 12 months after index, after controlling for Elixhauser comorbidity.[27] For the 13 to 24 months after index, this value was $5,898 (USD), after controlling for Elixhauser comorbidity.[27] Overall, patients with TRD required more

intensive health care than non-TRD patients in the two years following their first antidepressant. However, without any comparable studies in a Canadian context, there is still no data to compare this to a Canadian setting. Because per capita spending on health care is significantly higher in the United States compared to Canada,[28] it is possible that the difference in spending between TRD and non-TRD individuals is less in Canada.

Defining treatment-resistant episodes

The study by Kubitz and colleagues also used the PharMetrics Integrated Database, a commercial insurance dataset from the United States.[29] Data between 1995 and 2010 were used and included information for over 70 million people. The authors included patients between 18 and 64 years of age and excluded those diagnosed with schizophrenia or bipolar disorder. Rather than defining individuals as having TRD, the authors classified episodes of depression as being TRD or not. This study is the first to use this approach.

The authors defined an episode of depression as a continuous period where the "first relevant date" of the episode was the date of the first MDD diagnosis. The "last relevant date" of the episode was the data the patient received either a diagnosis code for depression or a prescription for an antidepressant followed by a period of 120 days without either of these.[29] The authors defined a TRD episode as a single episode with at least two distinct failed treatment regimens. Antipsychotics and "antimanic" drugs were the only drugs included as part of augmentation therapy. Prevalence of TRD in the population was not defined, though 6.6% of depression episodes were categorized as TRD.[29]

The study by Fife and colleagues marked the first epidemiological study on TRD that uses population-level data from a nationally representative dataset. Data from a sample of 1

million people from Taiwan's National Health Insurance Program for the year 2005 were used for this study.[30] The authors excluded individuals with a diagnosis of schizophrenia, bipolar disorder or dementia. Rather than measure the prevalence of TRD among the sample, this study replicated and amended the methods used by Kubitz and colleagues that categorized episodes of PTD as being TRD or not. In this study, an episode of PTD was characterized as TRD if at least two antidepressant therapies failed. The authors defined a failed therapy as one that was stopped at least 15 days after it began and replaced with another therapy.

It is important to note, however, that the authors mention that the pharmaceutical treatment of depression is very uncommon in Taiwan.[30] Of the 704,265 patients included in this sample database, only 2,751 (0.39%) had an episode of PTD, compared to the 6% annual estimates of antidepressant use in Canada.[31] As such, the results of this study cannot be generalized to jurisdictions like Manitoba and are not explored further.

The study by Mahlich and colleagues sought to replicate the studies of Kubitz et al. and Fife et al., and used a private insurance claims database from Japan.[32] The definition of PTD and TRD were the same as those used by Fife and colleagues. As the rate of PTD was also very low in this population (1.16%), the results of this study cannot be generalized to jurisdictions like Manitoba and are not explored further.

Overall, it is difficult to draw any conclusions from these three studies, as the use and antidepressants in the study populations is dissimilar to the experience in Canada. One limitation to the studies by Fife et al. and Mahlich et al. are their criterion of replacing a therapy within 15 days of its start. The 2016 guidelines from the Canadian Network for Mood and Anxiety Treatments (CANMAT) recommend that physicians increase the dosage for medications if there

is no response within two to four weeks if the medication is tolerated.[33] Therefore, using Fife et al.'s approach would likely not be appropriate for a Canadian study.

TRD and mortality

Mrazek and colleagues reviewed 62 articles, including 59,462 patients, and summarized the clinical, economic and societal burden of TRD between 1996 and 2013. The authors defined TRD as a failure to respond to one or more adequate trials of drug therapy. This liberal definition of TRD was used to account for the heterogeneity of definitions used in the literature.[34] While this review did not report the prevalence of TRD in each study, it does report some associated outcomes. In terms of the risk for mortality, the authors found that suicidal ideation was reported for 15±8% of TRD patients, 6% of MDD patients, and 1% of the general population.[34]

Pfeiffer and colleagues used data from the Veterans Health Administration's National Registry of Depression to study the prevalence of suicide among cases of TRD. This was the first study to examine the association between TRD and mortality. Rather than use a binary definition of TRD, the authors used the MGH-s staging method and compared patients with scores of 1 or less, 1.5 to 2.5, and 3 or higher using a case-control design. The authors found that those with an MGH-s score of 3 or higher experienced nearly a 50% increase in odds of completed suicide (OR=1.52; 95% CI=0.98 – 2.37) and scores of 1.5–2.5 experienced nearly a 20% increase in odds of completed suicide (OR=1.19; 95% CI:=0.91 – 1.55) compared to stage 1 or less. These results are not statistically significant at $\alpha = 0.05$. However, only three years of data were evaluated, and only suicides that occurred within two years of the index antidepressant prescription were included in the study. Additionally, the population from the Veterans Health Administration are not representative of a general population, as it is 97% male, with a mean age of 56 years, and many have been exposed to traumatic military experiences.[35] Further

investigation into the relationship between TRD and suicide is warranted given these preliminary results.

Reutfors and colleagues were the first to use administrative data from a national dataset to compare all-cause mortality among individuals with TRD and those with treatment-responsive MDD. The sample included 118,774 individuals between the ages of 18 and 69 years who had received pharmaceutical treatment for their depression.[36] The authors defined TRD as receiving at least three different antidepressant trials of at least four weeks within the same episode of depression, which could include augmentation or combination therapy, ECT, or repetitive transcranial magnetic stimulation. Combination therapy was categorized as another antidepressant medication dispensed within 14 days of the previous medication. Individuals were excluded if they ever received a diagnosis for psychosis, mania, bipolar disorder, or dementia. The authors found that 12.6% of the cohort met the definition for TRD.

The adjusted hazard ratio (HR) for all-cause mortality for the entire sample was 1.35 (95% CI=1.21 – 1.50), while the adjusted excess mortality rate ratio was 1.52 (95% CI=1.1-1.76).[36] When stratified by age, those 18-29 years of age had the highest HR of 2.03 (95% CI 1.55 – 2.64), while those 50-69 years of age had the lowest HR of 1.19 (95% CI 1.04 – 1.36).[36] These results are expected as the risk of all-cause mortality increases as people age. There was no difference in risk between males and females.[36] Overall, individuals with TRD experienced higher rates of mortality when compared to individuals with treatment-responsive MDD.

Summary

The definitions used by prior researchers for defining TRD have been varied. With this assortment of definitions has followed a wide range of results that describe the epidemiology of

TRD and its associated health services use and mortality. Some studies have found that 11.6% of individuals with PTD are categorized as TRD,[21] while others have found this number to be as high as 28.1%.[23] The primary limitation in the previous literature, though, is that TRD has consistently been defined using a binary definition rather than a staged one. Conway and colleagues have raised this concern in 2017,[13] but no study has yet to implement this approach for examining the epidemiology of TRD.

Despite a lack of a consistent definition for TRD, previous research has found that TRD is associated with higher costs and rates of health services use,[21,25,27] and increased risk for death.[35,36] Studies into the rates of health services use have, however, been limited to specific populations that cannot be generalized to an entire population. Additionally, the methods used in these studies have not investigated how recurrent health services use, such as recurrent hospitalizations for specific individuals, may be an important variable to include through appropriate data analysis.

Overall, to best understand the epidemiology of TRD, as well as its association with health services use and death, better research is needed. An optimal way of studying TRD would be to define it using a staging method and would use data from a generalizable population. Additionally, using appropriate methods of data analysis that include recurrent health service use is an important consideration in determining the impact of TRD on the health of individuals with MDD.

Research Questions

In order to address the gaps that I have identified in the current literature for TRD, this thesis was completed to answer the following research questions:

1. What percentage of individuals in Manitoba with PTD meet the criteria for TRD?
2. Between 1996 and 2016, what is the incidence and prevalence of TRD in Manitoba?
3. What differences exist between Manitobans with PTD and those with TRD in terms of sociodemographic characteristics, i.e. age, sex, urban residence, socioeconomic status, and Elixhauser comorbidity index?
4. How do the risks for ambulatory visits, emergency department visits, and hospitalization compare between Manitobans with PTD to those with TRD? Do any differences remain when stratifying by the primary reason for health care use, i.e. mental health related visits or hospitalizations compared to all other visits and hospitalizations?
5. How does the risk of all-cause mortality compare between Manitobans with PTD to those with TRD?

I hypothesized that Manitobans with TRD would experience higher risks for health services use and death and that these risks would increase as MGH-s scores increased.[c]

[c] See section below for the rationale in defining TRD using the Massachusetts General Hospial Staging Method

Methods

This retrospective cohort study used administrative health data from the Manitoba Population Research Data Repository (the Repository) at the Manitoba Centre for Health Policy (MCHP). In this section, I first describe the datasets that I used in creating the study cohort. After describing these datasets and the information they provided, I describe the criteria used in creating the cohort, as well as how TRD was defined using these datasets. Next, I provide details on the outcome variables of interest and how I extracted the necessary data from the datasets. Finally, I describe the analytic methods used in answering my research questions and in testing the associated hypotheses.

Study Design

Study period

The study period was January 1, 1996, to November 30, 2016. Individuals were followed from the time they entered the cohort until either: the end of the study period (November 30, 2016), death, or they were lost to follow-up. If an individual left the province and later returned, they re-entered the cohort upon receiving Manitoba Health insurance coverage. An additional washout period, described below, was used from April 1, 1995, to December 31, 1995.

Data sources

The Drug Product Information Network (DPIN) dataset contains data on all pharmaceutical dispensations from community pharmacies in Manitoba from April 1, 1995, until March 31, 2018. The DPIN dataset includes information on the dispensed drug's Anatomical Therapeutic Chemical (ATC) code, the date of dispensation, and the length of the prescription. I used this information to determine an individual's TRD status, MGH-s score, and the inclusion

and washout criterion. This thesis included DPIN data from April 1, 1995, until December 31, 2016.

MCHP defines ambulatory visits as "visits to a licensed physician in an outpatient setting in Manitoba."[37] The date and primary diagnosis for an ambulatory visit, as well as the specialty of the physician associated with the health record, were extracted from the Medical Services dataset from April 1, 1995, until November 30, 2016. The washout period, the inclusion criteria and exclusion criteria used the primary diagnosis listed in the Medical Services dataset as required.

Visits to emergency departments (ED) from January 1, 2007, to March 31, 2015, were extracted from the Emergency Department Information System (EDIS) dataset. I could only determine the chief complaint of an individual visiting an ED for visits extracted from the EDIS dataset, and not from the Admissions, Discharge and Transfer dataset. The latter dataset was, therefore, not used for this thesis. I used the EDIS dataset for determining the date and reason for ED visits for the relevant analysis.

Hospital admissions data were extracted from the Hospital Abstracts dataset from April 1, 1995, until November 30, 2016. The hospital episode macro for SAS, developed at MCHP, was used to include only single episodes of hospitalizations and to ensure that a patient transfer did not contribute to additional records.[38] All macros and formats used in SAS are listed in Appendix 3. I used this dataset for determining whether an individual was included in the washout period, as well as for determining the inclusion and exclusion criteria, and the Elixhauser comorbidity index. Dates and primary diagnoses for hospitalizations were also used for the relevant analysis.

The Manitoba Health Insurance Registry contains information on when an individual is covered by provincial health insurance, as well as the reason for a person's coverage ending, which includes emigration and death. These data were used to determine when an individual was eligible to be included in the cohort, and when they were either censored or deceased.

An individual's place of residence is based on their six-digit postal code and is updated bi-annually in the Registry. I used these data to determine urban status as it changed during the study. An individuals' socioeconomic factor index (SEFI) was calculated based on their postal code, using data from census dataset. Lastly, this dataset was also used to determine an individual's sex and age at cohort entry. Table 3 summarizes the datasets used in this study and the reason for their use.

Table 3. Datasets Used and Reason for Use

Dataset	**Reason for Use**
Manitoba Health Insurance Registry	To determine coverage status, age, sex, population of Manitoba, and date of death.
Drug Product Information Network	To determine TRD status and MGH scores, inclusion criteria, and washout criteria.
Medical Services	To determine inclusion, exclusion and washout criteria, as well as the date and reason for ambulatory visits.
Emergency Department Information System	For emergency room visit dates and reason for visits.
Hospital Abstracts	For inclusion, exclusion and washout criteria, as well as the date and reason for hospitalizations.
Canadian Census	To determine income quintile, and SEFI.

Study Population

Inclusion Criteria

All Manitoba residents 18 years of age or older, with at least one year of continuous insurance coverage between January 1, 1996, and November 31, 2016, were eligible for the study. In order to be retained in the cohort, individuals had to meet the following inclusions criteria: (1) they needed to have at least one MDD diagnosis recorded in either the Medical Services or the Hospital Abstract dataset, (2) they needed to have a dispensation for an antidepressant during the period from 30 days before a diagnosis for MDD to 30 days after the MDD diagnosis, and (3) they needed to have a second dispensation for an antidepressant within 14 days of the first, (4) they needed to be dispensed at least 42 days worth of doses of one antidepressant. The third inclusion criterion increases the likelihood that the individual had taken their medication after it was dispensed. The fourth inclusion criterion ensures that an individual had taken their antidepressants for enough time to achieve a response. I defined individuals as having PTD if they met these criteria. See Appendices 1 and 2 for details on the ICD-9 and ICD-10 codes used for the MDD diagnoses and the ATC codes associated with antidepressants.

Washout and Exclusion Criteria

Data from the DPIN begin on April 1, 1995. I excluded all individuals with either an antidepressant dispensation or a diagnosis of MDD if these occurred between April 1, 1995, and December 31, 1995. This criterion ensured that prevalent users of antidepressants and individuals with previous diagnoses for MDD were more likely to be excluded, which increased the likelihood of knowing an individual's complete pharmaceutical history.

For this thesis, I used a 275-day washout period for individuals whose health insurance coverage started after December 31, 1995.[d] Therefore, if an individual was dispensed an antidepressant or received a diagnosis of MDD within 275 days of their coverage starting, they were also washed out of the study. This additional washout criterion ensured that the same amount of eligible washout time was applied for all individuals, and to be more confident that an individual was not receiving treatment for their MDD before receiving Manitoba Health insurance coverage.

Augmentation therapy for MDD may include atypical antipsychotics. In order to ensure that these medications were used as augmentation therapies for MDD, I excluded individuals with a diagnosis for dementia, schizophrenia, or a psychotic disorder not-otherwise-specified (i.e. disorders where atypical antipsychotics may be prescribed). Because the Medical Services dataset only includes three ICD-9 digits, it was not possible to differentiate individuals with a diagnosis of MDD from those with a diagnosis of bipolar affective disorder (BAD) in this dataset. Therefore, I excluded individuals with a diagnosis of BAD in the Hospital Abstracts dataset only. Diagnoses of dementia, schizophrenia, and psychotic disorder not-otherwise-specified were identified from both the Medical Services and Hospital Abstract dataset. See Appendix 1 for the ICD-9 and ICD-10 codes used for these diagnoses.

An individual required 42 days of continuous antidepressant dispensations to be included in the cohort. Therefore, an individual is essentially immortal for the 42 days following the start of their first adequate antidepressant dispensation, which would result in immortal time bias. I excluded individuals who entered the cohort between October 17, 2016, and November 30,

[d] There are 275 days between April 1 and December 31.

2016,[e] from the survival analysis in order to account for this bias. I also excluded these individuals from all other analyses because they would not have experienced a complete 42 days of antidepressant therapy during this time, meaning they had not had enough time to respond to their medication.

Definition TRD in Administrative Data

For this thesis, only pharmaceutical treatments for depression were investigated. I examined drug dispensation data from the DPIN to distinguish TRD from non-treatment-resistant PTD (referred to as PTD) individuals. This dataset includes all drug dispensations from community pharmacies and includes the ATC code, the date of dispensation, and the number of days worth of medication dispensed. Dispensations for drugs with ATC codes for antidepressant medications and augmentation therapies were examined to determine t he order that drugs were dispensed. See Appendix 2 for the included ATC codes.

If more than 120 days elapsed since the end of a drug dispensation and the start of a dispensation for the same drug, I defined these as separate episodes of dispensation, following the methods of Fife and colleagues.[30] Otherwise, subsequent dispensations for the same drug were grouped as a single episode of dispensation. For a drug dispensation episode to be included in defining TRD status, the drug must have been dispensed continuously for at least 42 days. The dispensation date for the first drug to be dispensed for 42 days was the date of cohort entry.

Choosing the MGH-s

I opted to use a staging method to define TRD to test the hypothesis that the outcomes associated with TRD follow a severity gradient, meaning that as the stage of treatment-resistance

[e] There are 42 days between October 17 and November 30.

increases, so to would the parameter estimates associated with the outcomes. This hypothesis was supported by previous findings using the MGH-s,[23,35] as well as by the commentary of Conway and colleagues.[13]

In deciding which staging method to use, I based this decision on the limitations of using administrative data for this thesis and the use of evidence-based criteria. In terms of data limitations, MCHP does not have any information on an individual's depression severity, which is required for using the Maudsley Staging Method. MCHP also does not have information on the reason for which an individual is continued on the same medication for many months. Because of European Staging Method considers the amount of time that an individual is kept on the same medication for its TRD staging, even CANMAT guidelines recommend prolonged antidepressant use for many individuals,[33] using this staging method would likely overestimate levels of treatment-resistance for people who do respond to treatment.

The Thase-Rush Staging Model, and the staging method proposed by Conway and colleagues in 2017, both require that different classes of antidepressant be trialled in order to increase a person's TRD stage. However, there is no evidence to support the fact that individuals who are dispensed different classes of antidepressants after failing to respond have different rates of response to those who are dispensed the same class of antidepressants.[33] Therefore, the use of such a criterion is not only unnecessary, but it also fails to accurately reflect how antidepressants may be prescribed in Manitoba, where in-class changes may be used. It is given these limitations that the MGH-s staging method was used for this thesis.

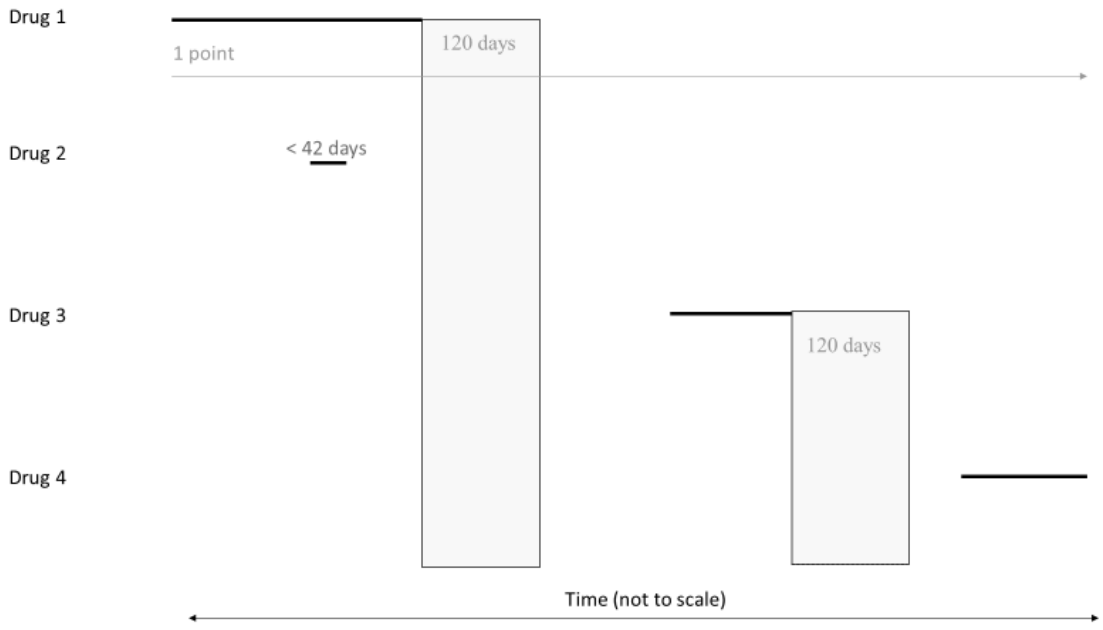
Determining MGH-s score

When assigning MGH-s scores, it was necessary to investigate how drug dispensation episodes overlapped with each other. Each drug an individual was dispensed was evaluated separately as a *reference drug*. For each reference drug, all drugs that were dispensed after that for at least 42 days were assessed for their respective overlap with the reference drug. Overlaps were assigned one of four types based on: (1) how long the dispensation episode overlapped with the reference-drug, if at all, (2) when the overlap occurred in reference to the start and stop dates of each dispensation episode, and (3) whether the overlap drug was an antidepressant or an augmentation therapy. Overlaps of an antidepressant that occurred entirely within the dispensation episode of the reference drug and did not surpass it, or which ended within 42 days of the reference drug's end date, were defined as a combination therapy and assigned a point value of 0.5. Overlaps of an antidepressant that occurred for any part of the reference drug dispensation episode and continued for more than 42 days after this were defined as independent therapies and assigned a point value of 1. Overlaps for a drug defined as an augmentation therapy were always categorized as an augmentation therapy and assigned a point value of 0.5. Finally, if a drug was dispensed within 120 days of the end of a reference drug, it was defined as an independent therapy and assigned a point value of 1.

For every reference drug, the point values of the associated drug overlaps were determined in order to calculate the MGH-s score for the respective reference-drug dispensation episode. The first time a reference drug was assigned a point value greater than 1, I defined the date when an individual's MGH-s score increased as the start date of the overlapping drug responsible for the increase in point value. I repeated this process if another reference drug

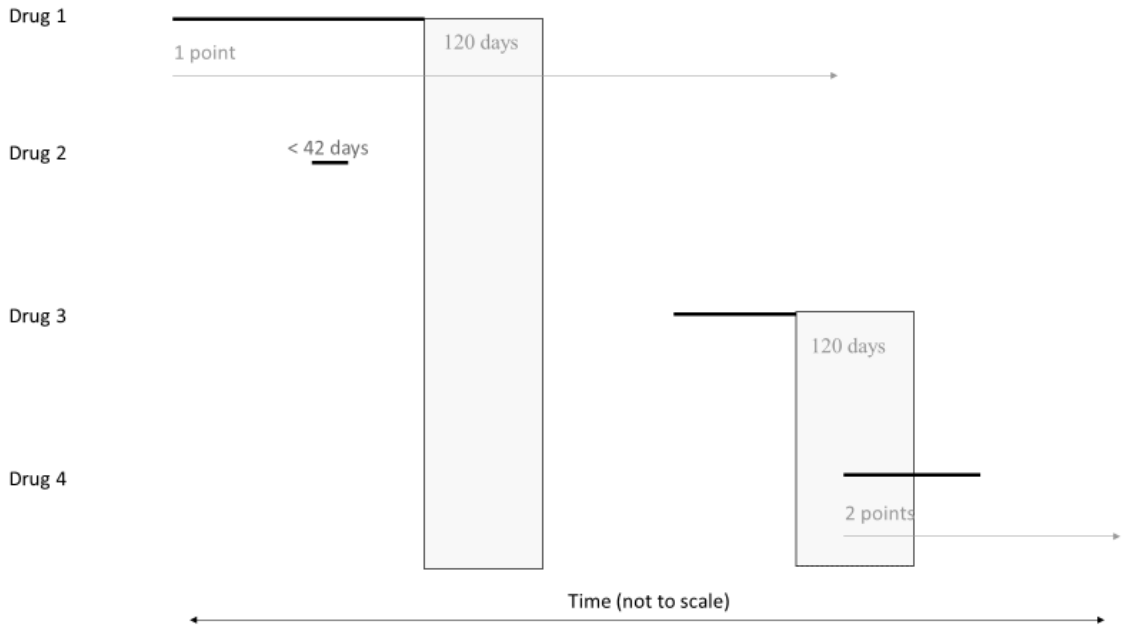
dispensation episode was assigned a point value greater than a preceding reference drug. The following three figures provide examples of how scores were determined.

Figure 1. Massachusetts General Hospital Staging Model Scoring Example - MGH-s score of 1



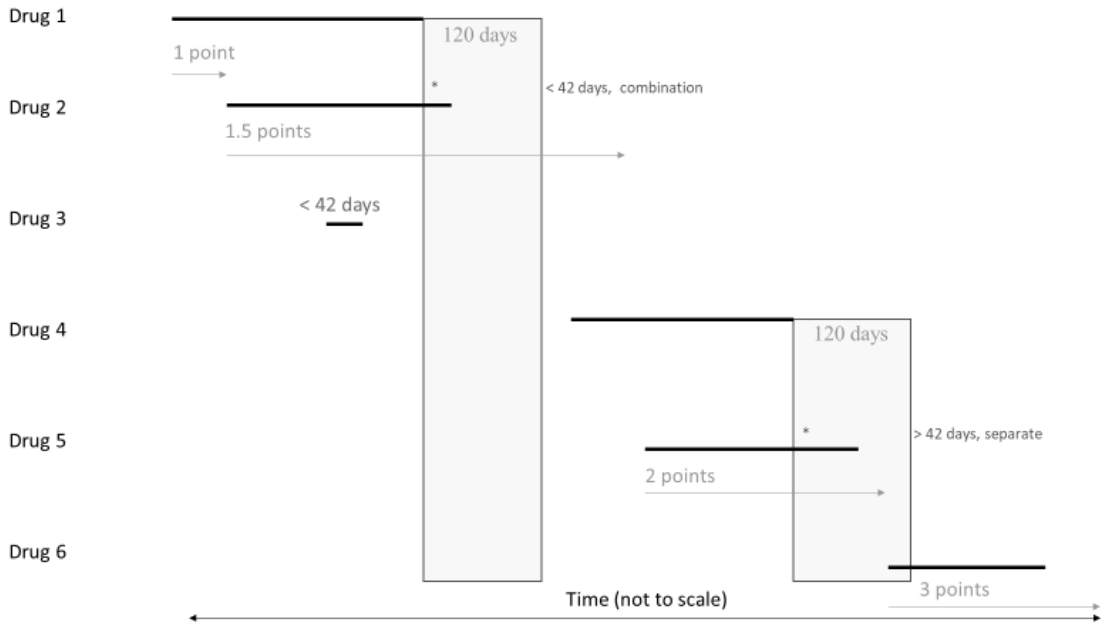
In Figures 1 to 3, time is denoted as progression along the x-axis, while the thick black lines represent the times when an individual was being dispensed a particular drug. For simplicity, all drugs are treated as antidepressants. The example illustrated in Figure 1 shows an individual who, despite being dispensed four different medications over the study period, had a score of one. Though Drug 2 was dispensed concurrently with Drug 1, thus meeting the MGH-s criterion for a combination therapy, it was dispensed for less than 42 days, and was therefore not included. No other drug was dispensed with 120 days of Drug 1 being discontinued, so no point increased would be recorded for Drug 1's dispensation episode. The same is true for both Drug 3 and Drug 4.

Figure 2. Massachusetts General Hospital Staging Model Scoring Example - MGH-s score of 2



The example illustrated in Figure 2 shows a slight variation to the example provided in Figure 2. The difference is that Drug 4 is dispensed within 120 days of Drug 3 being discontinued. Because of this, Drug 4 is coded as a being a second, independent treatment that is part of the dispensation episode that began when Drug 3 was first dispensed, which is associated with a full-point increase. Therefore, this person's MGH-s score increases for 1 to 2 one the date that Drug 4 was first dispensed.

Figure 3. Massachusetts General Hospital Staging Model Scoring Example - MGH-s score of 3



The example illustrated in Figure 3 shows an individual who was dispensed 6 different drugs over the study period, and eventually reached an MGH-s score of 3 after two different dispensation episodes. The first dispensation episode includes Drug 1 and Drug 2, which were dispensed concurrently. As such, Drug 2 was coded as a combination therapy and was associated with a half-point increase in MGH-s score. This beings this individual's MGH-s score to 1.5. Because no other drugs were dispensed within 120 days of Drug 1 being discontinued, no other drugs were included in this first dispensation episode.

The second dispensation episode begins when Drug 4 is first dispensed. In this case, the individual continues to have an MGH-s score of 1.5 until their score increases again. However, when counting the scores for this second dispensation episode, the relative point values reset to one. In this example, Drug 5 was dispensed concurrently with Drug 4, but also continued to be dispensed for greater than 42 days after Drug 4 was discontinued. Because individuals may have some leftover drugs from a previous dispensation when they begin a new drug, and because

individuals may be slowly titrated off a drug as they begin a subsequent one, it is possible for drugs to overlap without necessarily being combination therapy. For this reason, Drug 5 was coded as an independent therapy, rather than a combination therapy, as it is believed that the intention was to change medications, rather than for Drug 5 to supplement Drug 4. This would result in a full-point increase from an MGH-s score of 1 (given the new dispensation episode) to and MGH-s score of 2. If Drug 5 were instead an augmentation therapy, Drug 5 would have been associated with a half-point increase, and MGH-s score for the dispensation episode would be increased to 1.5. Finally, because Drug 6 was dispensed within 120 days of Drug 4 being discontinued, it is associated with a full-point increase for this second dispensation episode, and which would increase this individual's MGH-s score from 2 to 3.

Determining binary TRD-status

Determining an individual's binary TRD-status was based on the methods used by Olfson et al.[27] I examined the number of drugs with start dates that were within 365 days of the start date of each reference drug. If at least two subsequent drugs met this criterion, an individual met the binary definition of TRD. The start date of the second drug that was dispensed for at least 42 days that met the within-365-day criterion was the date when an individual was defined as meeting the binary definition of TRD.

TRD as a time-dependent variable

When all individuals enter the cohort, they are assigned an MGH-s score of 1. This score can increase at any time after that, once the appropriate criteria are met. For some individuals, increases in MGH-s score may happen very quickly. For others, they may not experience an episode of treatment-resistance until many years after first entering the cohort. Because the

analyses used for this study measure specific time-at-risk, it was essential to ensure that immortal time was accounted for with regards to MGH-s scores.

For example, an individual who enters the cohort and then goes on to reach an MGH-s score of 3 could not have died between cohort entry and achieving this score. One solution to this immortal time bias could have been to ignore the time-at-risk between when an individual entered the cohort and when they reached their highest MGH-s score. This approach would have resulted in a considerable amount of loss of data. To minimize this loss of data, and to account for some individuals having considerable amounts of time between score increases, MGH-s score was included as a time-dependent variable. This means that the time between an individual entering the cohort and first reaching a higher MGH-s score was time attributed to having an MGH-s score of 1. Then, the time between this higher score and any subsequent score increase was attributed to this next higher score. This approach ensured that all time after an individual has entered the cohort was maintained, while also accounting for immortal time.

An additional benefit to this approach was how it adjusted for the variance in help-seeking behaviour of the cohort. For an individual to reach a higher MGH-s score, they were required to be prescribed multiple antidepressants, which required additional follow-up care compared to an individual whose score never increased. Therefore, it was possible that individuals with a higher MGH-s score exhibited more help-seeking behaviour than individuals with a lower MGH-s score. Directly adjusting for help-seeking behaviour was not possible when using administrative data. Because many of the outcomes of this study (e.g. ambulatory visits, ED visits) were likely confounded by help-seeking behaviour, using an approach to minimize this confounding was helpful.

By including MGH-s score as a time-dependent variable, an individual who went on to reach a higher MGH-s score contributed some of their time-at-risk to scores lower than their highest score. If help-seeking behaviour is constant over time for an individual, then the time-at-risk at lower scores for individuals who eventually reach greater MGH-scores would equally be subject to this greater help-seeking behaviour. This meant that while help-seeking behaviour for individuals with different MGH-s scores may have been a possible source of bias in the health services use analyses of this study, a viable solution to this was treating MGH-scores as time-dependent.

Outcome variables

Ambulatory visits

Ambulatory visits between January 1, 1996, and November 30, 2016, were extracted from the Medical Services dataset. These included all contacts with physicians, excluding contacts in hospitals as inpatients or in emergency departments. I categorized visits as mental health related visits if the primary diagnosis associated with a visit was for a mood or anxiety disorder. See Appendix 1 for specific ICD-9 codes.

Emergency department visits

Visits to Winnipeg emergency departments between January 1, 2007, and March 31, 2015, were extracted from the EDIS dataset. If the chief complaint was coded was “Mental Health” then I categorized the visits as such. Additionally, if the chief complaint was coded as “ETRIAGE”, meaning the chief complaint was missing, and the reason for the visit was coded as “Mental health assessment”, I coded the visit as a mental health related visit.

Hospitalizations

Inpatient hospitalizations between January 1, 1996, to November 30, 2016, were extracted from the Hospital Abstracts dataset. Hospitalizations, where the primary diagnosis was associated with a mood or anxiety disorder, were categorized as such. See Appendix 1for the ICD-9 and ICD-10 codes used to make this distinction.

All-cause mortality

I identified the date an individual died or was censored from the Health Insurance Registry dataset. Additionally, another source of immortal time bias comes from the 42 days of antidepressant dispensation required for an individual to enter the cohort. Because an individual must survive these 42 days in order to enter the cohort, this is a time where an individual is immortal. To address this, I removed the first 42 days after cohort entry from the analysis for all-cause mortality.

Statistical Analyses

Data management and statistical analyses for this thesis were performed using SAS® version 9.4. When required, macro programs and formats under the authorship of data analysts from MCHP were used, with permission. See Appendix 4 for the macros and formats that I used. Data extraction from the Repository was performed by Mr. Charles Burchill of MCHP. I performed all other data management and statistical analyses.

The majority of the analyses for this study examined individuals based on their MGH-s score. When individuals first entered the cohort, they were assigned an MGH-s score of 1, but all individuals could be assigned scores of 1, 1.5, 2, 2.5, and 3 if the criteria were met. Once an individual's score increased, it remained at this level for the remainder of their time in the study,

regardless of whether they stopped receiving treatment for depression. This was due to there being no way to be certain of the reason for which an individual discontinued seeking medical care for their depression. For example, being dispensed antidepressants continuously is not indicative of a continuation of depressive symptoms. Individuals being dispensed antidepressants may be in remission of their depression but can continue taking their medication for months or years after symptoms have resolved. There was, therefore, no accurate way to be sure that an individual should be removed from the cohort, based solely on administrative data. Unless otherwise specified, individuals with an MGH-s score of 1 were the reference group for all analyses in this study.

Descriptive Models

Demographics of individuals in the cohort were compared by their highest-ever MGH-s score. These variables included age, sex, income quintile, urban status (i.e. living in Winnipeg or Brandon), and whether an individual ever received ambulatory care from a psychiatrist. I determined age, income quintile, and urban status at cohort entry. Chi-square tests were used to determine the statistical significance of differences in demographics across MGH-s groups. Direct standardized prevalence and incidence of MGH-s and PTD were calculated annually using *proc stdrate*, with the Manitoba population of 2016 as the reference population.

Outcome Models

I modelled the hazards for ambulatory visits, ED visits, and hospitalizations for the entire study period using the counting process model of Andersen-Gill, a generalization of the Cox proportional hazards model.[39] This model allows for the analysis of recurrent events, for time-dependent variables to be included with the ability to change at random intervals, and for its

ability to accurately count time-at-risk. The primary assumptions of the model are proportional hazards, which were tested using Schoenfeld residuals, and independence of events.

Time-dependent variables included MGH-s status, SEFI, urban status, Elixhauser Comorbidity Index, and having a past-year psychiatrist visits. Time-independent variables included age, sex, and year of cohort entry. I included each variable in the model in a stepwise way and evaluated the effect of each new variable on the fit of the model using the Akaike information criterion (AIC). A decrease in the value of the AIC was associated with a model with a better fit. For the model with the lowest AIC, I tested the interactions of each covariate with MGH-s score separately, in order to test whether each covariate modified the relationship between MGH-s score and the outcome. I also did this for the interactions between each covariate and year of cohort entry, but no significant results were found.

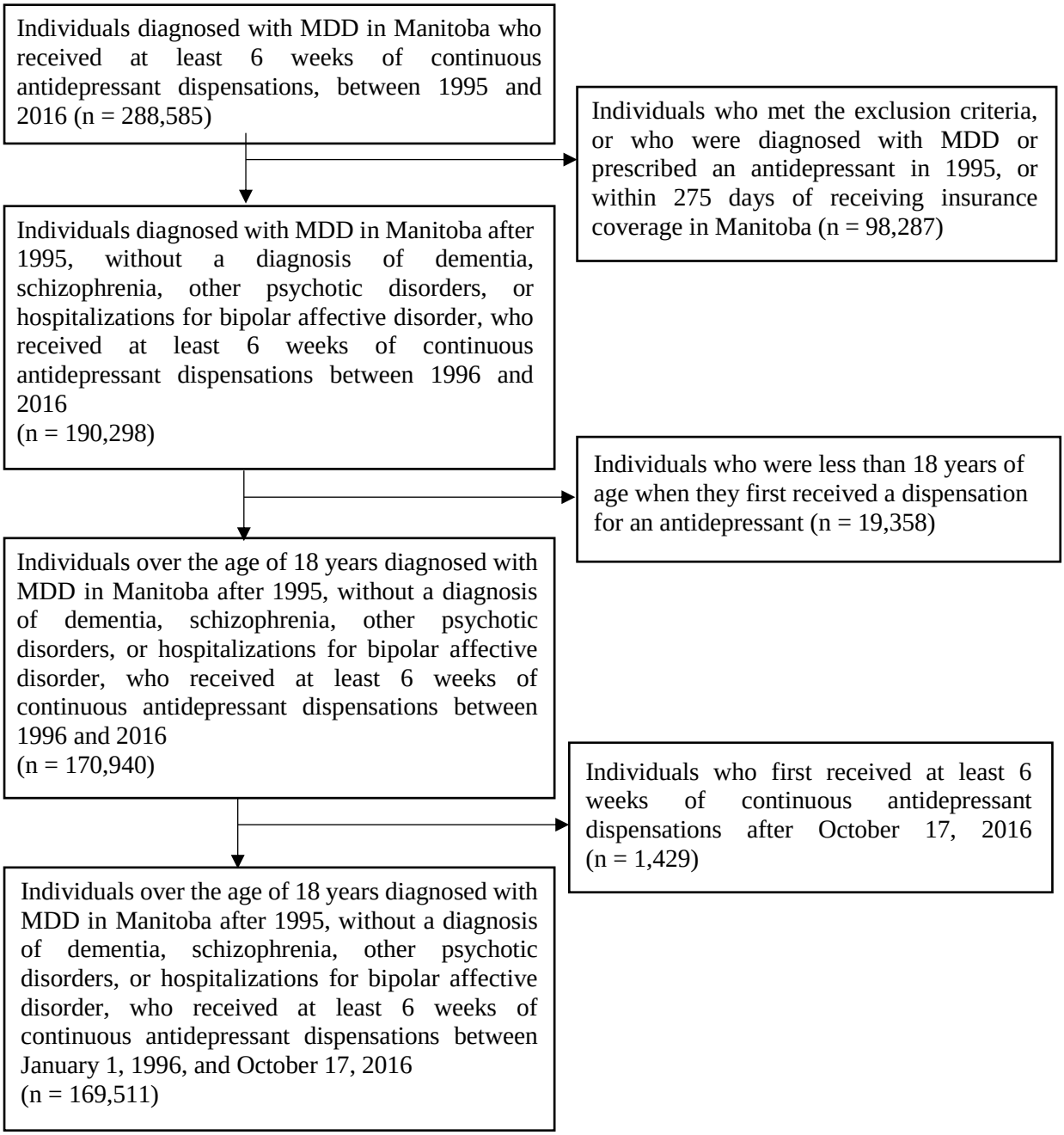
Because this thesis included a large cohort, many of the effect modifiers were statistically significant, but few covariates modified the association between MGH-s score and the outcome in a clinically significant way. In the following chapter, I present the results of the covariate with the effect modifier that was associated with the lowest AIC, and where the value of the parameter estimates for the interaction variables were clinically significant.

As described above, I categorized health services use events as either mood and anxiety disorder related or not and performed separate analyses for both event types. I used p*roc phreg* for this analysis and used the Efron approximation for handling tied events. The Andersen-Gill model was not required for modelling all-cause mortality. Instead, I used an extended Cox proportional hazards model for this outcome. All other methods described above were also used for this model, where appropriate.

Results

The cohort development for this project is presented in Figure 1.

Figure 4. Cohort Development



Demographics

Descriptive data for the entire cohort are summarized in Table 4, separated by the highest MGH-s score for which each individual met the criteria. The cohort included 169,511 individuals and a total of 1.7 million person-years (mean (SD) = 10.0 (5.9) person-years ; median = 9.9 person-years). A total of 18,663 (11.0%) individuals had met the criteria for MGH-s scores of 2.5 or higher. Individuals who reached higher MGH-s scores tended to be younger and more likely to live in Winnipeg or Brandon. Females comprised the majority of the cohort, and this difference in sexes was slightly higher as MGH-s scores increased. Individuals with higher MGH-s scores were also more likely to have ever had an ambulatory visit with a psychiatrist, compared to individuals with lower MGH-s scores. Figures 2 and 3 further show the trends in income quintiles in the cohort, by MGH-s score. For individuals living in Winnipeg or Brandon, those with the higher MGH-s scores were more likely to be living in areas associated with lower income. This trend was not apparent for individuals living in rural areas, where individuals with pharmaceutically treated depression were more likely to be living in areas associated with higher income, across all MGH-s score.

Table 4. Characteristics of Manitobans with Pharmaceutically Treated Depression, 1996-2016

	MGH 1 n = 102,367 60.4 %	MGH 1.5 n = 15,992 9.4%	MGH 2 n = 32,489 19.2%	MGH 2.5 n = 6,271 3.7%	MGH 3 n = 12,392 7.3%
Age					
Average (SD)	42.8 (16.4)	42.3 (14.8)	40.3 (14.8)	39.2 (13.8)	37.6 (12.5)
Median (IQR)	41.0 (23.0)	41.0 (22.0)	39.0 (20.0)	38.0 (20.0)	36.0 (17.0)
Sex (%)					
Female	63.7	62.8	66.4	65.6	66.4
Living in Winnipeg or Brandon (%)					
	62.9	62.6	62.6	64.6	65.1
Income Quintile (%)					
(Lowest) Urban 1	12.4	14.6	13.1	15.2	15.9
Urban 2	13.1	13.5	13.0	14.3	13.8
Urban 3	13.1	13.0	12.6	12.7	12.8
Urban 4	12.5	11.1	12.3	11.6	12.0
(Highest) Urban 5	11.9	10.2	11.7	10.9	10.7
(Lowest) Rural 1	6.3	7.3	6.3	6.9	6.0
Rural 2	7.2	6.8	7.0	6.5	6.1
Rural 3	7.5	7.4	7.6	6.8	7.0
Rural 4	7.4	7.5	7.9	7.4	7.3
(Highest) Rural 5	8.1	7.4	8.0	7.0	7.7
Ever had an Ambulatory Visit with a Psychiatrist (%)					
	14.9	11.2	25.6	46.8	56.6

Figure 5. Income Quintile Distribution Among Individuals with Pharmaceutically Treated Depression, Urban Quintiles, in Manitoba, 1996 – 2016

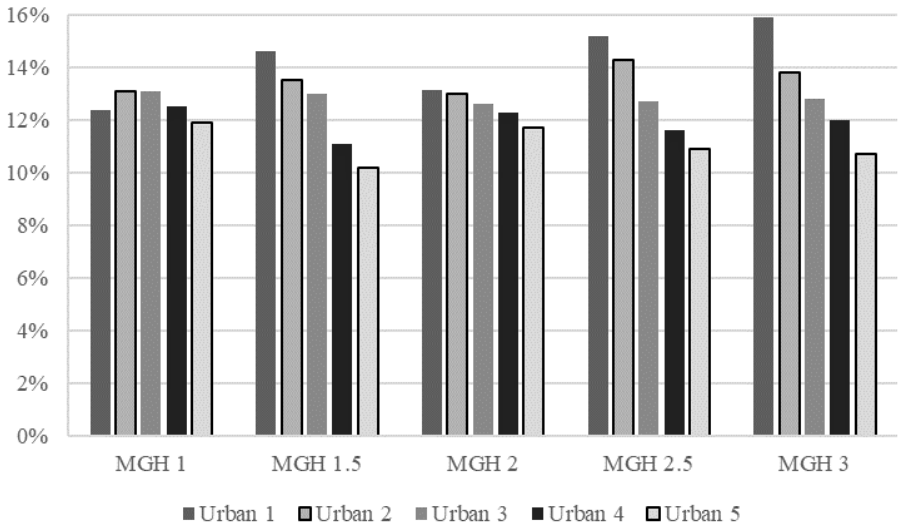
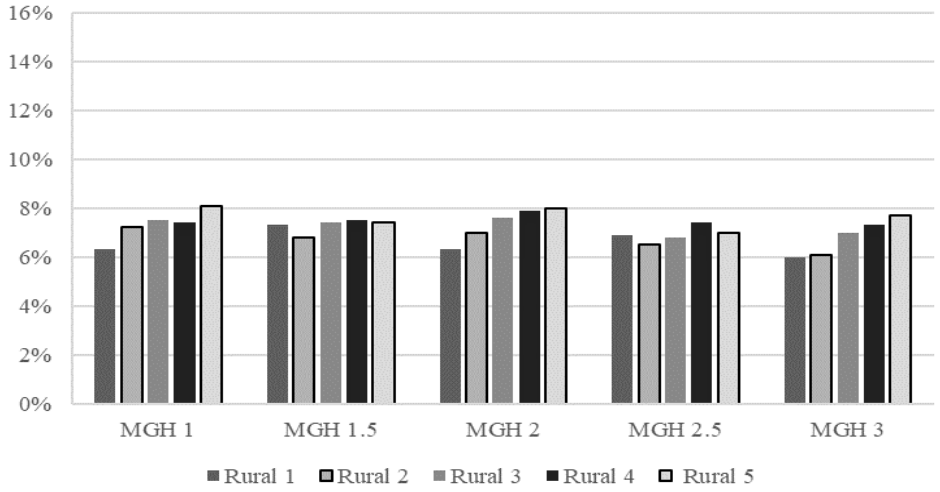


Figure 6. Income Quintile Distribution Among Individuals with Pharmaceutically Treated Depression, Rural Quintiles, in Manitoba, 1996 – 2016



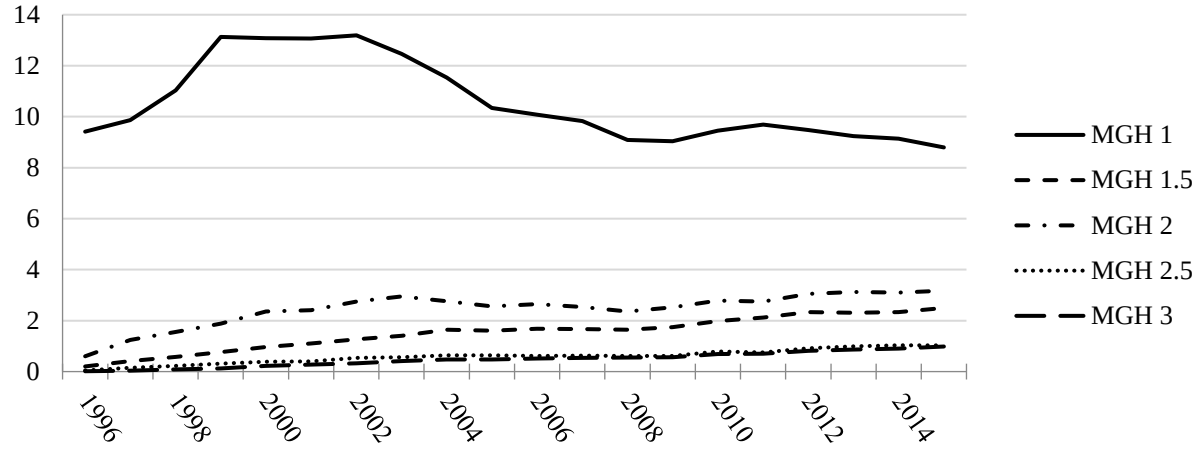
Trends in Treatment-Resistant Depression

Incidence

Between 1996 and 2002, the annual incidence rate for individuals with MGH-1 increased from 9.4 per 1,000 person-years to 13.2 per 1,000 person-years. After that, the incidence rate decreased to an all-time low of 8.8 per 1,000 person-years in 2015. The incidence rate of 2016 was not calculated, as the study period ended on November 30, 2016. The incidence rates are presented in Figure 4, below.

Figure 7. Age- and Sex-Standardized Incidence of Pharmaceutically Treated Depression, by MGH-s score, in Manitoba, 1996 to 2015

Per 1,000 person-years at risk, standardized to the 2016 population of Manitoba

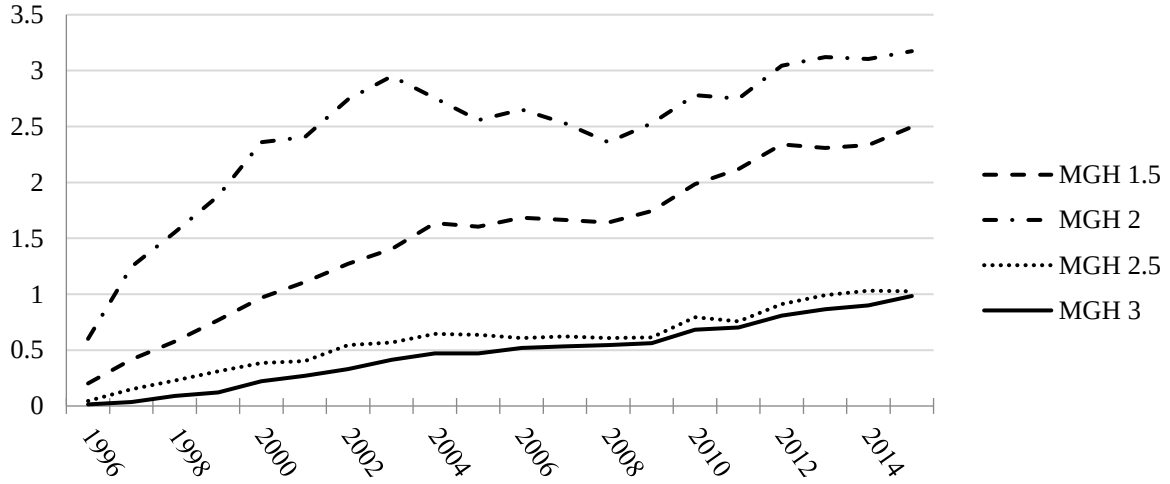


From 1996 to 2015, the annual incidence rate for all MGH-s scores higher than 1 increased linearly. Note that these are not the rates of individuals entering the cohort who eventually reached these specific MGH-s scores. Instead, the new cases of an MGH-s score higher than 1 were defined as when individuals already in the cohort reached a higher score than the one they had previously.

Among these higher MGH-s scores, the incidence rate for individuals with MGH-2 was the greatest, followed by MGH-1.5. The incidence rates for MGH-2.5 and MGH-3 were not statistically different. Note that Figure 5, shown below, presents the same data as shown in Figure 4 but focuses specifically on individuals with MGH-s scores of 1.5 to 3.

Figure 8. Age- and Sex-Standardized Incidence of Pharmaceutically Treated Depression, by MGH-s scores 1.5 - 3, in Manitoba, 1996 to 2015

Per 1,000 person-years at risk, standardized to the 2016 population of Manitoba



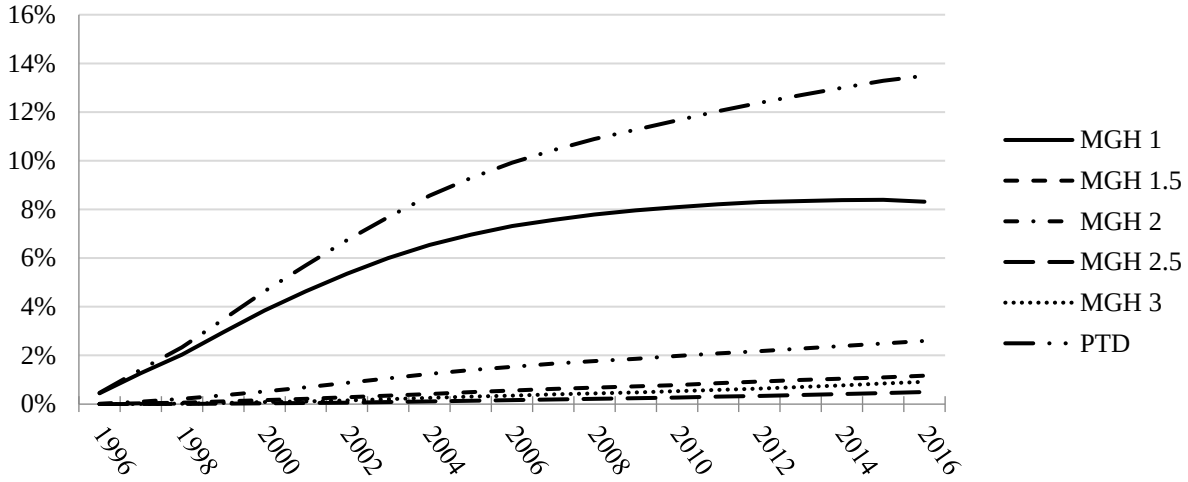
Prevalence

Figures 6 and 7 show the annual prevalence of individuals with PTD in the cohort. Note that Figure 7 presents the same data as shown in Figure 6 but focuses specifically on individuals with MGH-s scores of 1.5 to 3. Because MGH-s scores can change within a given year, the prevalence was counted from the score of all individuals on July 1 of each year. Prevalence of PTD continued to increase throughout the study period, while the prevalence of MGH- 1 increased until 2011, after which the prevalence stayed consistently close to 8.3% until 2016.

When evaluating the prevalence of MGH-s scores in this study, note that TRD was treated as a chronic, incurable illness. Once an individual had reached a certain MGH-s score, they maintained this score until it increased, they lost coverage, died, or the study ended. Therefore, the reported prevalence may not accurately reflect the percentage of Manitobans living with active cases of pharmaceutically treated depression in a given year.

Figure 9. Age- and Sex-Standardized Prevalence of Pharmaceutically Treated Depression, by MGH-s score, in Manitoba, 1996 to 2016

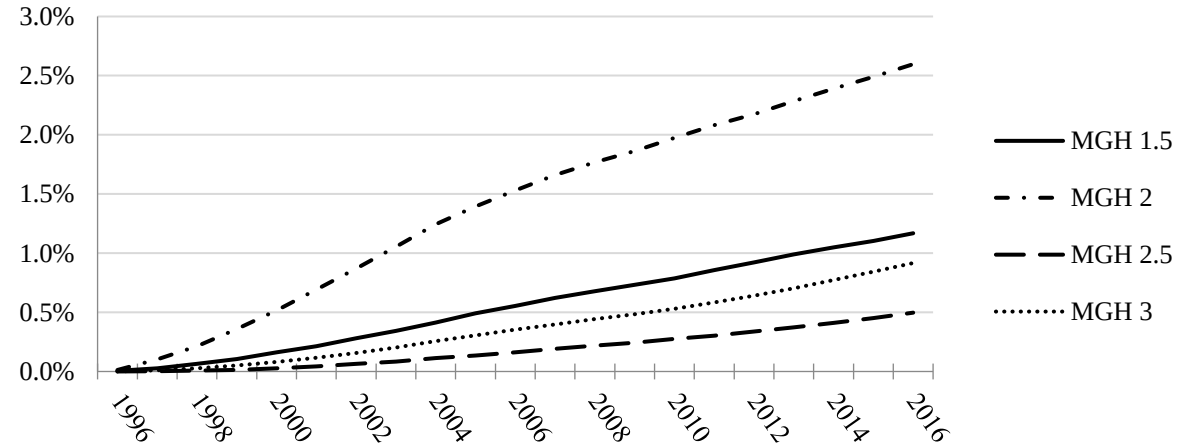
Standardized to the 2016 population of Manitoba



Though the prevalence of MGH-1 reached a peak value in 2011, the prevalence of all other MGH-s scores continued to increase through the entire study period. In 2016, 5.2% of Manitobans had reached MGH-s scores greater than 1, and 1.4% had reached MGH-s scores of 2.5 or greater.

Figure 10. Age- and Sex-Standardized Prevalence of Pharmaceutically Treated Depression, by MGH-s scores 1.5 to 3, in Manitoba, 1996 to 2016

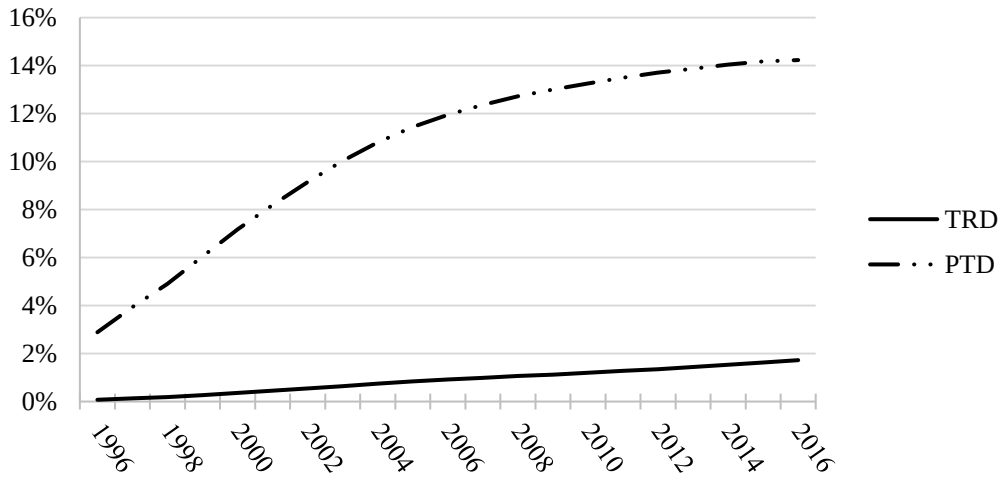
Standardized to the 2016 population of Manitoba



Though all analyses were performed using MGH-s scores for defining TRD, Figure 8 presents the age- and sex-standardized prevalence of TRD using the binary definition compared to all other individuals in the cohort. In 2016, 1.7% of Manitobans 18 years of age and older had, at some point, met the binary definition of TRD, while 14.2% of Manitobans 18 years of age and older had experienced PTD. It is important to note that these values do not necessarily reflect actively treated cases of PTD or TRD.

Figure 11. Age- and Sex-Standardized Prevalence of Pharmaceutically Treated Depression, Using the Binary Definition, in Manitoba, 1996 to 2016

Standardized to the 2016 population of Manitoba



In order to briefly compare the two definitions of TRD proposed for this study, Table 5 presents the percentage of individuals in the cohort who met the binary definition of TRD, stratified by the highest MGH-s score that had reached. Only a small percentage of individuals with MGH-s scores less than 2.5 were categorized as having TRD with the one-year binary definition. In comparison, nearly half of individuals with MGH-2.5 and three-quarters of individuals with MGH-3 were categorized as having TRD with the binary definition.

Table 5. Percentage of Individuals who met the Binary Criteria for Treatment-Resistant Depression, by MGH-s score, Manitoba, 1996 – 2016

MGH-s Score	Binary TRD Status: Non-TRD	Binary TRD Status: TRD
1	103,696 (99.98%)	20 (0.02%)
1.5	15,566 (97.18%)	451 (2.82%)
2	28,757 (88.72)	3,658 (11.28%)
2.5	3,382 (54.45%)	2,829 (45.55%)
3	3,017 (24.58%)	9,258 (75.42%)

Health Services Use

The following section summarizes the hazard ratios and 95% confidence intervals for ambulatory visits, emergency department visits, and hospitalizations. The variables included in the tables are the variables included in the model with the lowest AIC. See Appendix 5 for the AIC values associated with each model. When an interaction term was found to be statistically significant, a separate table was created to summarize how this interaction term affected the hazard ratios associated with MGH-s scores for the relevant outcome. For all outcomes with a significant interaction term, I present the results only for the models with the lowest AIC and with clinically significant parameter estimates for the interaction term.

For all models, age was determined at cohort entry. Additionally, SEFI, urban status, Elixhauser comorbidity index, and past year psychiatrist visits were included in the models as time-dependent covariates. MGH-s scores were also included as a time-dependent variable. Urban status (urban or rural), sex (male or female), past year psychiatrist visit (yes or no), and MGH-s scores were all included as categorical variables. The category listed in the tables

represent the categories associated with the hazard ratio. For urban status, sex, and past year psychiatrist visit, the reference categories are rural, female and no, respectively.

Ambulatory Visits

Table 6 summarizes the hazard ratios and 95% confidence intervals for ambulatory visits where the primary diagnoses was not a mood and anxiety disorder. The hazard ratios for ambulatory visits increased significantly as MGH-s scores increased. Compared to individuals with MGH-1, those with scores of 2.5 and 3 experienced 38% increased hazards for ambulatory visits, controlling for all other covariates. Every one-point increase in the Elixhauser comorbidity index was associated with a 22% increased hazards for ambulatory visits. Additionally, increases in SEFI (i.e. decreasing socioeconomic index), living in an urban area, and having a past year ambulatory visit with a psychiatrist were associated with approximately 12% higher hazards of ambulatory visits.

Table 6. Adjusted* Hazard Ratios for Ambulatory Visits Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For reasons other than mood and anxiety disorders

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	1.22	1.22	1.23
MGH 2	1.18	1.17	1.18
MGH 2.5	1.38	1.37	1.38
MGH 3	1.38	1.38	1.39
Age			
	1.02	1.02	1.02
Sex			
Male	0.87	0.87	0.87
SEFI			
	1.12	1.12	1.12
Urban Status			
Urban	1.11	1.11	1.11
Elixhauser			
	1.22	1.22	1.23
Past Year Psychiatrist Visit			
Yes	1.13	1.12	1.13

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visit, and year of cohort entry*

When the interaction between age and MGH-s scores is included in this model, the hazard ratios for MGH-s scores change significantly. Because the interquartile range of age for individuals in all five MGH-s score groups was approximately 20, hazard ratios were calculated for MGH-s score ten years below and above the cohort average, based on the parameter estimate for the interaction term. Table 7 summarizes these results. For individuals ten years below the average cohort age (i.e. approximately 30 years of age), the hazard ratios for MGH-s scores of 2.5 and 3 increased statistically more than for MGH-s scores of 1.5 and 2. For individuals ten years above the average cohort age (i.e. approximately 50 years of age), the hazard ratios for MGH-s scores of 2.5 and 3 were much more similar to MGH-s scores of 1.5 and 2.

Table 7. Adjusted* Hazard Ratios for Ambulatory Visits, Including the Effect Modification of Age on MGH-s scores, Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For reasons other than mood and anxiety disorders

	10 years below average age			10 years above average age		
	Hazard Ratio	95% Confidence Interval		Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression						
MGH 1.5	1.22	1.22	1.23	1.22	1.21	1.22
MGH 2	1.20	1.20	1.21	1.15	1.15	1.16
MGH 2.5	1.43	1.42	1.43	1.32	1.31	1.32
MGH 3	1.46	1.46	1.47	1.24	1.24	1.25

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visit, and year of cohort entry*

Table 8 summarizes the results for ambulatory visits where the primary diagnosis was a mood and anxiety disorder. Compared to the results in the previous tables, individuals with pharmaceutically treated depression and higher MGH-s scores experienced much higher hazards of ambulatory care, compared to individuals with lower MGH-s scores. Note that for this model, past year psychiatrist visit was not included, as this variable would be highly associated with the outcome. This model also shows that individuals living in urban areas experience a 46% higher risk for mood and anxiety disorder related ambulatory visits, compared to those living in rural areas. In the previous models, urban status was associated with an 11% increase in hazards. Elixhauser comorbidity index was not associated with an increased hazard for these ambulatory visits, contrasted by the 23% increased risk in the previous model. There were no interaction terms with clinically significant results for this model.

Table 8. Adjusted* Hazard Ratios for Ambulatory Visits Among Individuals Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For mood and anxiety disorders

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	2.05	2.04	2.06
MGH 2	2.37	2.36	2.38
MGH 2.5	3.87	3.84	3.89
MGH 3	5.21	5.19	5.23
Age			
	0.99	0.99	0.99
Sex			
Male	0.98	0.98	0.98
SEFI			
	0.95	0.94	0.95
Urban Status			
Urban	1.37	1.34	1.40
Elixhauser			
	1.00	0.99	1.00

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, and year of cohort entry*

Emergency Department Visits

Table 9 summarizes the hazard ratios and 95% confidence intervals for ED visits where the chief complaint was not associated with a mental health concern. In this model without any interaction terms, MGH-s score, SEFI, urban status, and past year psychiatrist visits were associated with increased hazards for ED visits. As MGH-s scores increased, so too did the hazard ratios for ED visits, although there was no statistically significant difference between individuals with MGH-s scores of 2.5 and scores of 3.

Table 9. Adjusted* Hazard Ratios for Emergency Department in Winnipeg Visits Among Individuals with Pharmaceutically Treated Depression in Manitoba, January 1, 2007, to March 31, 2015

Where a mental health concern was not the chief complaint

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	1.34	1.32	1.37
MGH 2	1.25	1.23	1.26
MGH 2.5	1.57	1.53	1.60
MGH 3	1.55	1.52	1.57
Age			
	1.00	1.00	1.00
Sex			
Male	1.04	1.03	1.05
SEFI			
	1.49	1.49	1.50
Urban Status			
Urban	3.81	3.77	3.86
Past Year Psychiatrist Visit			
Yes	1.19	1.18	1.21

**Adjusted for age, sex, SEFI, urban status, past year psychiatrist visit, and year of cohort entry*

Similar to non-mood and anxiety disorder related to ambulatory visits, the relationship between MGH-s scores and ED visits was modified by age. As age increased, the magnitude of the relationship between MGH-s scores of 2.5 and 3 and ED visits decreased. As age decreased, however, the magnitude of the relationship between MGH-s scores of 2.5 and 3 and ED visits became much greater. Table 10 summarizes the hazard ratios for MGH-s scores, when including the interaction between MGH-s score and age, for individuals 10 years below and above the average cohort age. There was no statistically significant difference between MGH-s scores of 2.5 and 3.

Table 10. Adjusted* Hazard Ratios for Emergency Department Visits, Including the Effect Modification of Age on MGH-s scores, Among Individuals with Pharmaceutically Treated Depression in Manitoba, January 1, 2007, to March 31, 2015

Where a mental health concern was not the chief complaint

	10 years below average age			10 years above average age		
Treatment-Resistant Depression	Hazard Ratio	95% Confidence Interval		Hazard Ratio	95% Confidence Interval	
MGH 1.5	1.43	1.41	1.45	1.32	1.30	1.34
MGH 2	1.32	1.30	1.33	1.21	1.20	1.23
MGH 2.5	1.81	1.77	1.85	1.42	1.39	1.45
MGH 3	1.82	1.70	1.85	1.36	1.34	1.39

**Adjusted for age, sex, SEFI, urban status, past year psychiatrist visit, and year of cohort entry*

When examining ED visits where the chief complaint is a mental health concern, the effect of MGH-s scores increased significantly, as does the difference in magnitude of hazard ratios between MGH-s scores. Additionally, SEFI, urban status, past year psychiatrist visit and sex were all associated with increased risks for mental health related ED visits. The magnitude of the relationship of past year psychiatrist visit with ED visits more than doubled when examining only ED visits for mental health concerns, compared to the previous model. These results are presented in Table 11, below.

Table 11. Adjusted* Hazard Ratios for Emergency Department Visits in Winnipeg Among Individuals with Pharmaceutically Treated Depression in Manitoba, January 1, 2007, to March 31, 2015

Where a mental health concern was the chief complaint

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	2.09	1.95	2.24
MGH 2	1.89	1.78	2.00
MGH 2.5	2.70	2.44	2.97
MGH 3	3.53	3.30	3.78
Age			
	0.96	0.96	0.97
Sex			
Male	1.37	1.30	1.43
SEFI			
	1.50	1.47	1.53
Urban Status			
Urban	3.10	2.92	3.29
Past Year Psychiatrist Visit			
Yes	2.71	2.58	2.86

**Adjusted for age, sex, SEFI, urban status, past year psychiatrist visit, and year of cohort entry*

In contrast to ED visits for non-mental health concerns, the interaction term with the lowest AIC and largest parameter estimates for this model was that of sex and MGH-s score. However, this interaction was only significant for MGH-s scores of 2.5 and 3. Table 12 summarizes the results of including this interaction term in the model and separates the results for females and males. This analysis shows that sex moderated the association between MGH-s score and ED, where females experienced higher risks than males at equivalent MGH-s scores.

Table 12. Adjusted* Hazard Ratios for Emergency Department Visits, Including the Effect Modification of Sex on MGH-s scores, Among Individuals with Pharmaceutically Treated Depression in Manitoba, January 1, 2007, to March 31, 2015

	Females			Males		
Treatment-Resistant Depression	Hazard Ratio	95% Confidence Interval		Hazard Ratio	95% Confidence Interval	
MGH 1.5	2.33	2.13	2.57	2.39	2.15	2.65
MGH 2	2.08	1.93	2.24	2.19	2.01	2.39
MGH 2.5	3.78	3.32	4.26	3.17	2.72	3.70
MGH 3	5.56	5.13	6.02	4.08	3.66	4.54

**Adjusted for age, sex, SEFI, urban status, and year of cohort entry*

Hospitalization

Comparable to ambulatory and ED visits, there was a statistically significant relationship between the risk for hospitalizations for reasons other than mood and anxiety disorders and MGH-s scores. As MGH-s scores increased, so too did the magnitude of this relationship. Compared to MGH-s scores of 1, individuals with MGH-s scores of 2.5 and 3 experienced a 47% increased risk for hospitalization for reasons other than mood and anxiety disorders. In contrast to previous outcomes, living in Winnipeg or Brandon was associated with a decreased risk for hospitalization. SEFI, Elixhauser comorbidity index and having a past year visit with a psychiatrist were associated with an increased risk for hospitalization. There was no clinically significant effect modifier for this model.

Table 13. Adjusted* Hazard Ratios for Hospitalizations Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For reasons other than mood and anxiety disorders

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	1.28	1.27	1.30
MGH 2	1.22	1.21	1.23
MGH 2.5	1.47	1.44	1.50
MGH 3	1.47	1.45	1.49
Age			
	1.03	1.03	1.03
Sex			
Male	0.95	0.94	0.95
SEFI			
	1.16	1.16	1.17
Urban Status			
Urban	0.80	0.79	0.80
Elixhauser			
	1.35	1.34	1.35
Past Year Psychiatrist Visit			
Yes	1.13	1.11	1.14

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visits, and year of cohort entry*

As was seen in previous outcomes, when examining only hospitalizations for a mood and anxiety disorder, the relationship between MGH-s score and hospitalizations was of a greater magnitude than was the case for all other hospitalizations. Individuals with MGH-s scores of 3 experienced 8.3 times greater risk for mood and anxiety disorder related hospitalizations compared to individuals with MGH-s score of 1. Similar to non-mood and anxiety disorder related hospitalizations, living in Winnipeg or Brandon was associated with a 52% decrease risk for hospitalization, compared to living in rural areas. Having a visit with a psychiatrist within the

past year was associated with nearly four times the risk for hospitalizations for mood anxiety disorders, but only a 13% increased risk for all other hospitalizations.

Table 14. Adjusted* Hazard Ratios for Hospitalizations Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For mood and anxiety disorders

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	2.51	2.19	2.75
MGH 2	2.58	2.41	2.78
MGH 2.5	4.85	4.32	5.44
MGH 3	8.33	7.73	8.98
Age			
	1.00	1.00	1.00
Sex			
Male	1.00	0.94	1.05
SEFI			
	1.21	1.18	1.24
Urban Status			
Urban	0.48	0.45	0.50
Elixhauser			
	1.37	1.33	1.40
Past Year Psychiatrist Visit			
Yes	3.90	3.68	4.14

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visits, and year of cohort entry*

Tables 15 and 16 present the results of the association of two clinically significant interactions with MGH-s with mood and anxiety disorder related hospitalizations, modelled separately. The hazard ratios for MGH-s scores for the models that included the interaction between MGH-s scores and having a past year psychiatrist visit are presented in Table 15. When including this interaction, the risks of hospitalization associated with increased MGH-s scores were higher for individuals with a past year psychiatrist visit than for those without such a visit.

For those without a past year psychiatrist visit, the hazard ratios range from 2.6 to 6.5, compared to a range of 2.7 to 12.4 for individuals with a past year psychiatrist visit.

Table 15. Adjusted* Hazard Ratios for Hospitalizations, Including the Effect Modification of Having a Past-Year Visit with a Psychiatrist on MGH-s scores, Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For mood and anxiety disorders

	Individuals With a Past Year Psychiatrist Visit			Individuals Without a Past Year Psychiatrist Visit		
Treatment-Resistant Depression	Hazard Ratio	95% Confidence Interval		Hazard Ratio	95% Confidence Interval	
MGH 1.5	2.71	2.32	3.30	2.58	2.33	2.87
MGH 2	3.16	2.72	3.68	2.54	2.35	2.75
MGH 2.5	6.10	5.05	7.41	4.74	4.08	5.49
MGH 3	12.35	10.75	14.08	6.45	5.81	7.09

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visits, and year of cohort entry*

Because the interaction between MGH-s score and sex had comparable clinical significance as that of MGH-s score and having a past year psychiatrist visit, the summary of this interaction is presented in Table 16. Females experienced a higher risk for mood and anxiety disorder related hospitalizations than males with the same MGH-s score.

Table 16. Adjusted* Hazard Ratios for Hospitalizations, Including the Effect Modification of Sex on MGH-s scores, Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

For mood and anxiety disorders

	Females			Males		
Treatment-Resistant Depression	Hazard Ratio	95% Confidence Interval		Hazard Ratio	95% Confidence Interval	
MGH 1.5	2.60	2.31	2.92	2.35	2.03	2.72
MGH 2	2.67	2.45	2.91	2.43	2.17	2.72
MGH 2.5	5.49	4.78	6.29	3.80	3.11	4.67
MGH 3	9.26	8.47	10.10	6.85	6.10	7.69

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visits, and year of cohort entry*

All-Cause Mortality

When examining the relationship between MGH-s scores and the hazards for all-cause mortality, Table 17 shows there is no significant trend. Rather, both MGH 1.5 and 3 scores were associated with 89% and 87% increased hazards of all-cause mortality, respectively, while MGH-s scores of 2.5 were associated with a 112% increase.

Table 17. Adjusted* Hazard Ratios for All-Cause Mortality Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

	Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression			
MGH 1.5	1.89	1.79	2.00
MGH 2	1.47	1.40	1.53
MGH 2.5	2.15	1.96	2.36
MGH 3	1.87	1.73	2.02
Age			
	1.09	1.09	1.09
Sex			
Male	1.72	1.66	1.77
SEFI			
	1.28	1.26	1.30
Urban Status			
Urban	1.06	1.02	1.09
Elixhauser			
	1.38	1.37	1.40
Past Year Psychiatrist Visit			
Yes	1.54	1.45	1.64

**Adjusted for age, sex, SEFI, urban status, Elixhauser comorbidity index, past year psychiatrist visits, and year of cohort entry*

When the interaction between age and MGH-s score is included in the model, though, a significant trend is found. Table 18 displays the results of the effect modification of age on MGH-s score for all-cause mortality. For young individuals, MGH-s scores of 2.5 and 3 were associated with significantly higher hazards for all-cause mortality compared to all lower scores. In comparison, for older individuals, there is no significant difference in hazards for all-cause mortality between individuals with MGH-s score of 1.5 and 3

Table 18. Adjusted Hazard Ratios for All-Cause Mortality, Including the Effect Modification of Age on MGH-s scores, Among Individuals with Pharmaceutically Treated Depression in Manitoba, 1996 to 2016

	Ten years below average age			Ten years above average age		
	Hazard Ratio	95% Confidence Interval		Hazard Ratio	95% Confidence Interval	
Treatment-Resistant Depression						
MGH 1.5	2.26	2.07	2.48	2.03	1.85	2.22
MGH 2	1.70	1.58	1.83	1.55	1.44	1.67
MGH 2.5	3.11	2.74	3.55	2.40	2.11	2.73
MGH 3	2.86	2.54	3.16	2.02	1.83	2.23

Discussion

In this thesis, my primary goals were to develop a better understanding of the negative outcomes associated with TRD, and to test the hypothesis that TRD follows a severity continuum. Death, in particular, is the most serious outcome that could be investigated. Health services use, however, offer a more nuanced description of the health of individuals. Because healthy individuals are less inclined to visit a physician's office or emergency department, and inpatient hospitalizations are generally limited to the most severely ill, experiencing a higher risk for these events is generally a sign of poorer health. In this section, I first describe the demographics of this cohort and explain the reasons for the trends observed for incidence and prevalence. After that, I explain the importance of the results and compare them to previous findings. Finally, I describe the strengths and limitations of this thesis and offer ways of improving it in the future.

Demographics

Many of the studies that have previously described the demographics of individuals with TRD have had limited generalizability due to their study populations. This thesis is one of the first to describe the demographics of individuals with PTD and TRD for a general population. The population for this thesis includes urban, rural and remote communities, adults of all ages, and individuals from a broad range of socioeconomic backgrounds. In this section, I discuss why the differences in demographics between different MGH-s score may exist, while also noting the limitations of using administrative data for drawing definitive conclusions about these differences. I also compare these findings to those of previous studies and discuss why some differences occur.

Age

Age at cohort entry represents the age at which individuals were first prescribed an antidepressant for at least 42 days. Individuals who eventually reached MGH-3 were on average five years younger when they entered the cohort than individuals who had maintained a score of 1 for the entire study. Though this may reflect a true difference in average age, it is possible that this difference results from having more complete antidepressant histories for younger individuals. It is more likely that the Repository contains the complete antidepressant history of younger individuals because the DPIN dataset may not include earlier antidepressant treatments of older individuals because of its 1995 start date.

For example, an individual who was 38 years old in 1995 could have been prescribed antidepressants at 18 years old in 1975 without this being included in DPIN. In comparison, if an individual was 18 years old in 1995, they were less likely to have had any antidepressant dispensation before this year due to their young age. As such, younger individuals may not be more likely to experience TRD. It may instead be more likely that younger individuals' complete medical history is better captured in DPIN, which increases the likelihood of their MGH-s score increasing.[40]

Because this study was not designed to determine the age of first antidepressant dispensation, we cannot be certain that there is a significant age difference between individuals with different MGH-s scores. A solution could be to change the inclusion criteria to allow only individuals who were 18 years of age or older as of 1995. This change would have dramatically decreased the size of the cohort, however, and would have affected the overall power of the study.

Sex

A government-commissioned study done at MCHP, *Mental Illness Among Adult Manitobans*, found that between 2010/11 and 2014/15, females experienced a higher prevalence for mood and anxiety disorders than males.[10] For females, the prevalence of mood and anxiety disorders ranged from 30% to 36%, compared to males who experienced a range of 18% to 22%.[10] Other survey research in Canada has found that females are twice as likely to be taking antidepressants compared to males.[41,42] These findings are consistent with the results of this thesis, where females were nearly twice as likely to meet the criteria for PTD than males. This difference may be due in part to females being more likely than males to seek care for their mental health.[43] When comparing the percentage of females with different MGH-s scores, the difference between groups is not very large. Females were marginally more likely to have MGH-s scores of 2 or greater than scores of MGH-1. If females are less likely to meet the criteria for TRD than males, there would be a more considerable difference in the percentage of females with MGH-s scores of 2.5 or higher compared to scores of 1. The difference of 2.7 in the percentage of females with MGH-1 versus MGH-3 is statistically significant, though it is not clinically relevant. As females are more likely to seek medical treatment for their mental health compared to males, the difference in the proportion of females among those with higher MGH-s scores may be a reflection of this behavioural difference, rather than a difference in treatment-resistance.

Urban status

Approximately 62.7% of Manitobans lived in Winnipeg or Brandon in 2016. In comparison, 63.0% of the cohort lived in one of these cities when they first met the criteria of PTD. Had urban status been associated with an increased likelihood of meeting the criteria for

PTD to TRD, a significant difference between these values would have been found. Individuals with MGH-s score of 2.5 and 3 were 2.2 and 2.5 percentage point more likely to live in Winnipeg or Brandon. This difference means that Manitobans living in a rural setting are slightly less likely to meet the criteria for TRD than those living in an urban setting.

A 2018 report found that Manitobans living in rural areas struggle to access mental health care.[44] As such, these findings may be due in part to a decreased ability in seeking follow-up care in rural areas after not responding to treatment. If individuals are not seeking follow-up care after not responding to treatment, they are less likely to switch medications and have an increase in MGH-s scores. These findings may also be due to a lower availability to specialized psychiatric care available in rural areas. If more specialized psychiatric care is available, there may be a higher likelihood of a psychiatrist prescribing combination or augmentation therapies, which also increase MGH-s scores. However, the observed difference in percentage is small, and may not offer any clinical relevance.

Urban income quintile

For Manitobans living in Winnipeg or Brandon, there are significant differences between the five MGH-s scores regarding the distribution of individuals living in areas with different incomes quintiles. For MGH-1, there is no significant difference in the distribution of individuals across income quintiles. This distribution means that individuals living in areas associated with lower income were as likely to be included in this group as individuals living in areas associated with higher income. In comparison, for MGH-3, 15.9% were living in an area associated with the lowest income quintile, and 10.7% were living in an area associated with the highest income quintile.

One possible explanation for this difference may be that individuals living in lower-income areas are less able to seek out non-medical treatment for MDD, such as psychological care. Lower-income individuals may, therefore, be more reliant on seeking care from a physician and using pharmaceuticals, because of provincial health insurance coverage. Lower-income individuals may also receive provincial insurance coverage for prescription medication either through provincial subsidies for low-income individuals, whereas insurance coverage for private psychological therapy is not included in provincial insurance, which is a barrier to access.[44] Therefore, individuals living in lower-income areas who do not respond to antidepressants may be less likely to try non-medical treatment, and may instead switch medications, thus having an increased MGH-s score.

It is difficult to explain the variability in the distribution of income quintiles between MGH-1, MGH-1.5, and MGH-2. The distribution observed for MGH-1.5 is more similar to the distributions of MGH-2.5 and MGH-3 than it is of the distributions of MGH-1 and MGH-2. This trend, where the results associated with MGH-1.5 are more similar to MGH-2.5 and MGH-3 is found in other outcomes, too. One possible explanation is that MGH scores do not measure the exact number of drug changes, because some changes have a value of a half-point, while others have a value of a full-point. Therefore, differences between MGH-1.5 and MGH-2 may relate to the differences associated with half-point changes, such as augmentation therapies. MGH-1.5 can only be reached with a single half-point change, whereas MGH-2 can be reached with either two half-point changes or a single full-point change. It is possible that receiving a half-point increase, such as being dispensed an atypical antipsychotic rather than completely changing antidepressants, may be associated with socioeconomic status, though additional analysis would be needed to confirm this. Additionally, assigning income quintiles to individuals is not precise,

as it is an area-level measure. This difference may, therefore, be associated with imprecisions in assigned income quintiles to individuals in the cohort.

Rural income quintile

For Manitobans living in rural areas, a different trend exists. There is no significant difference in the income quintile distribution between different MGH-s scores in rural areas. Instead, the income quintile distributions of all five MGH-s scores show that more individuals in this cohort were living in higher-income areas than in lower-income areas. However, this trend was very small.

Like individuals living in urban settings, this difference may also be related to access to non-medical treatments for MDD. In urban settings, individuals living in higher-income areas who fail to respond to initial antidepressant treatment may be more likely to try non-medical treatments for MDD, and therefore switch antidepressant medications, which would increase their MGH-s score. In comparison, non-medical services for MDD such psychological therapy may be less available in rural areas of Manitoba. As a result, individuals living in rural areas may not have the same opportunity to avoid pharmaceutical treatments for MDD after not responding to an initial treatment. Individuals living in rural higher-income areas may, therefore, be as likely as individuals living in rural lower-income areas to change their antidepressant medications if there are not responding, and thus increase their MGH-s score.

This discussion is only speculative, however. Non-medical treatment for MDD, such as psychological care, is not captured in the Repository. Other methods would be required to test why these differences exist. Additionally, the income quintiles used for this analysis do not represent the actual income of the individuals. They are instead area-level measures associated

with the average income of a large area. It is, therefore, possible for an individual's income to be different from the income quintile to which they are assigned. Finally, though a significant trend exists across some MGH-s scores for the distribution of income quintiles, these differences may not translate to clinically important differences. This analysis offers a first-time look at how income may be associated with TRD, and further research is needed to support or question these findings. Additionally, the difference between urban and rural individuals regarding the association between income and MGH-s score also raises questions into whether income is truly associated with TRD, or whether it represents disparities in access for mental health care. If lower income is associated with higher MGH-s scores, this trend would be expected in both urban and rural areas.

However, if the findings of these analyses accurately reflect how income is associated with TRD, further action is required. Much of the current literature into treating TRD involves finding physiological reasons for why an individual is not responding to treatment. If income proves to be a predictor in TRD, then it must be investigated as part of a causal pathway. Living with low income is often associated with higher stress, higher rates of chronic disease and disability, and generally results in fewer social opportunities.[45] These are issues that antidepressants alone cannot resolve. It is therefore not a surprise that individuals who experience MDD and who live in poverty may not experience a reduction in their depressive symptoms from antidepressant treatment. Further research into this relationship may be warranted, based on these initial findings.

Summary

It is difficult to make definitive conclusions regarding the differences in age, sex, urban status, and income quintiles between individuals with different MGH-s scores. This is because

these demographic differences may be related to help-seeking behaviour and access to medical services, which can confound the association between demographics and MGH-s scores. Additionally, the magnitude of the differences observed in the baseline demographics of individuals with different MGH-s scores is small. Compared to the Manitoba population, there was a minimal difference in urban status for individuals with PTD. Similarly, the proportion of females in each group is consistent with findings that females are more likely to seek care than males.[43] Comparing the differences in the median age across the five groups is challenging, due to data only being available starting in 1995. Finally, the differences found in the distribution of individuals in different income quintiles may support the need for more research into the impact of income and socioeconomic status on TRD.

Visits with a psychiatrist

The final demographics comparison listed in Table 4 is not related to demographics but describes the percentage of individuals with different MGH-s scores who have ever had a visit with a psychiatrist. This analysis shows that the likelihood that an individual will have ever had an ambulatory visit with a psychiatrist increases as MGH-s scores increase. The likelihood of ever having had an ambulatory visit with a psychiatrist almost doubles from MGH-1 to MGH-2 2. It then doubles again between MGH-2 and MGH-3. Nearly half of all individuals with TRD in Manitoba between 1996 and 2016 were never seen by a psychiatrist, despite their greater need for more specialized care. A valuable next step in researching TRD in Manitoba would be to try and better understand why this is the case.

Psychiatrists are better trained in diagnosing co-occurring psychiatric illnesses and are better suited in treating patients with higher psychiatric needs when compared to family physicians. Psychiatrists may also have different patterns in prescribing antidepressants and

augmentation therapies for MDD compared to family physicians, and may be more likely to offer some form of psychotherapy. Additionally, individuals with more severe forms of depression or with psychiatric comorbidities may be more likely to be referred to a psychiatrist than an individual with less complex needs. The very large difference in the likelihood of seeing a psychiatrist for individuals with different MGH-s scores, therefore, raises many questions about the homogeneity of the cohort in terms of depression severity and psychiatric comorbidities. It was for this reason that all subsequent analyses included having a past-year psychiatrist visit (PYPV) as a time-dependent covariate.

Though data on depression severity are not available in the Repository, describing the psychiatric comorbidities of the cohort would be possible and would be a valuable next step in this research. Furthermore, subgroup analyses by the number of psychiatric visits overtime would be another strong addition to this research. It may be possible that individuals with MGH-2.5 and MGH-3 are not only more likely to ever see a psychiatrist, but they may experience more visits with a psychiatrist than individuals with lower scores.

Trends Over Time

Incidence

Figure 4 showed that the incidence of MGH-s scores higher than 1 increased consistently during the study, whereas the trend in incidence rates of MGH-1 changed over time. Between 1996 and 2003, the incidence of MGH-1 increased; it then decreased between 2002 and 2009, and then remained relatively stable until the end of the study.

One possible explanation for the rise in the incidence of antidepressant dispensations between 1996 and 2002 can be the result of many new antidepressants being approved by Health

Canada during this time. For example, sertraline, venlafaxine, bupropion, and citalopram were approved in 1993, 1995, 1998, and 1999, respectively.[46] As new drugs come to market, there is likely an increase in the marketing of these drugs and a resulting increased likelihood of prescribing these drugs.[47] Similarly, clinical trials for new drugs will often test their superiority to contemporaneously approved drugs. For example, Barbey and Roose found that selective serotonin reuptake inhibitors (SSRI) were "far safer" than tricyclic antidepressants (TCA) in terms of overdose risk.[48] In separate 2000 reviews, Zohar and Westenberg, and Peretti, Judge and Hindmarch independently found that both SSRIs and TCAs were effective treatments for depression, but SSRIs were more tolerable and safer than TCAs.[49,50] As the risk associated with newer antidepressant treatment was lower, physicians may have become more willing to prescribe antidepressants to individuals whom they may not have otherwise treated with pharmaceuticals previously.

In 2002, findings from the National Population Health Survey showed that antidepressant use for individuals with MDD increased in Canada from 18.2% in 1994 to 32.6% in 1998.[51] The author attributes this to a decrease in stigma towards seeking medical care for mental health, increased detection of MDD, and improved treatment of MDD in primary care over this period.[51] Overall, the increased incidence of antidepressant dispensations in Manitoba between 1996 and 2002 can be attributed to an increase in available treatments, improved safety of newer antidepressants, decreased stigma towards seeking care, as well as improved detection for MDD in primary care.

In 2003, researchers in North America and Europe began raising alarms about the safety of many commonly prescribed antidepressants and their risk for suicide and suicide ideation.[52,53] In 2004, the Food and Drug Administration in the United States issued a "black-box" warning

that antidepressants may increase these risks for young people.[54] Health Canada followed suit and issued a warning the same year, adding that this risk may extend to individuals of all ages.[55] The fact that these warnings occurred at the same time that the incidence of MGH-1 began decreasing is likely not a coincidence. Physicians and their patients may have become more apprehensive about prescribing and taking antidepressants following these warnings, which resulted in the observed decrease in incidence.

The trends found in this thesis for MGH-1, which represents the incidence rate of first-time antidepressant use, closely match those found in previous Canadian research.[31,42,47,51,56] The increase in antidepressant use in the 1990s and early 2000s was common across Canada.[31,51,56] In a 2006 to 2012 study of antidepressant prescribing patterns in Canada, Morkem and colleagues found that the incidence of antidepressant prescriptions decreased from 3.54% in 2006 to 2.75% in 2010, and then increased to 3.07% in 2012.[42] While these rates are higher than the results found in this thesis, Morkem and colleagues measured the rates of prescribing antidepressants, rather than the rates of being dispensed antidepressants.[42] Additionally, Morkem and colleagues' study was not limited to prescriptions given to individuals with MDD but included off-label prescriptions as well.[42] These differences in methods likely explain much of the variation in the results of their study and this one.

In comparison, Raymond and colleagues measured the incidence of antidepressant dispensations in British Columbia between 1998 and 2004.[47] The incidence rates reported in that study were nearly identical to the results of this thesis. Incidence rates for antidepressant dispensations reached a peak of approximately 14 per 1,000 population 45 to 64 years of age in 1999, followed by a decrease to approximately 7 per 1,000 by 2001. Incidence rates continued to decrease until a low of approximately 6 per 1,000 population in 2004. Overall, these findings

confirm that the results of this study are not unique and that Manitoba was not alone in experiencing an increase in antidepressant dispensations in the 1990s and early 2000s, followed by a decrease in dispensations starting in the mid-2000s.

Finally, the results for MGH-s scores greater than 1 were not surprising. As the number of individuals in the cohort increased, and as time passed, more individuals were eligible to experience a score increase. Notable, as the incidence of MGH-1 began decreasing in 2003, so too did the incidence for MGH-2. At the same time, the incidence rates for MGH-s score of 1.5, 2.5 and 3 were constant from 2002 to 2009, after which time the incidence rates for all scores greater than 1 began increasing again. This increase coincides with the time where the incidence of MGH-1 stopped decreasing and became more constant. Overall, these trends in incidence rates of scores greater than 1 show that they depend primarily on the number of individuals entering the cohort who may then go on to experience changes to their medications (i.e. the incidence of MGH-1).

Prevalence

In this thesis, I defined PTD as a chronic condition. This means that once an individual reached a particular MGH-s score, including a score of 1, they kept this score for the duration of the study period, or until it increased. As would be expected with this criterion, the prevalence of PTD increased consistently during the study period. As such, interpreting the results for the prevalence of PTD in this study must be done with caution, as it likely does not reflect the prevalence of active cases of PTD or TRD.

An alternative approach would be to include only individuals who received an antidepressant dispensation in a given year or who are actively seeing medical treatment for

MDD in the annual prevalence counts. I did not choose this approach, as I believed that being dispensed an antidepressant does not necessarily mean an individual is experiencing an active episode of MDD. Instead, individuals may continue being dispensed antidepressants as prophylaxis after symptoms have improved to ensure that symptoms of depression do not worsen.[47] Similarly, individuals who have stopped being dispensed antidepressants may not necessarily be in remission. These individuals may have instead discontinued treatment in favour of other therapies, or no treatment at all. Further analysis that includes a criterion for MDD remission in administrative data can be pursued in future follow-up research, but this may never be completely accurate.

It is important to consider why the trend line for the prevalence of MGH-1 is different from that of MGH-s scores greater than 1. In particular, the trend line for MGH-1 begins to experience a slower increase as of 2003, and eventually becomes nearly flat as of 2008, while the trend lines for all other scores continue to increase during the study period. The continuous increase in prevalence for MGH-1 between 1996 and 2002 coincides with the period where the incidence of MGH-1 was also increasing. Similarly, the gradual decrease in the upwards trend for MGH-1 prevalence beginning in 2003 coincides with the time when warnings about the safety of antidepressants began, and incidence rates decreased.

This does not explain why the prevalence of MGH-s scores greater than 1 continued to increase during this same time, however. To explain this, we must understand that the prevalence of MGH-1.5 to MGH-3 does not depend on individuals taking antidepressants for the first time. The prevalence for these scores instead depends on the number of individuals with lower scores. As more individuals experience treatment failures and change medications, the prevalence of scores higher than 1 increases and the prevalence of MGH-1 decreases. The incidence of MGH-

1.5 and 2, therefore, attenuate the effect of the incidence of MGH-1 on its prevalence. More generally, an increase in the prevalence of a score higher than 1 translates into a decrease in the prevalence in the scores less it. However, as there are fewer people with higher scores, there is less likelihood of medication changes occurring, which means that the prevalence of higher scores does not experience as much of a "draining" effect as was experienced by MGH-1.

Overall, the incidence of all MGH-s scores offers a useful insight into the patterns of antidepressant dispensations between 1996 and 2015. The reputation of antidepressants experienced many social changes over this time, which helps to explain some of the observed trends. However, because of limitations in accurately defining active cases of MDD using administrative data, measuring the prevalence of PTD over time was more challenging. These results, therefore, offer less insight into the burden of PTD and TRD over time. Further research is needed to refine these methods.

Comparing TRD Definitions

In the introductory section of this thesis, I presented a thorough comparison of the definitions of TRD in order to highlight the importance and the difficulty in choosing a valid operational definition of TRD. The benefits in using a binary definition for TRD are that they used more frequently in published literature, they offer a less labour-intensive approach to categorizing individuals, and interpreting the results of analyses that used these definitions is less complicated. In comparison, the MGH-s staging method definition has been used less frequently, required more complicated data management, but offer a more nuanced approach in categorizing treatment-resistance. As such, one of the goals of this thesis was to test the hypothesis that TRD follows a severity continuum, where higher MGH-s scores are associated with more negative outcomes.

I did not perform a formal comparison of these two definitions for this thesis. However, Table 5 was presented to compare how the cohort of this study would have been categorized using either definition. Not surprising, as MGH-s score increased, so too did the likelihood that individuals would meet the binary definition of TRD. The likely reason for why not all individuals with MGH-s scores of 2.5 and 3 met the binary definition is that these individuals experienced the changes to their medication over a period longer than one year. Similarly, the reason why individuals with MGH-s scores of 1.5 and 2 met the binary definition is that they experienced more than two drug changes in a given year, but that more than 120 days separated these drug changes.

Percentage of TRD for all PTD

This thesis found that 11.0% of individuals with PTD in Manitoba met the criteria for TRD between 1996 and 2016. Previous studies that calculated the proportion of individuals with PTD who experienced TRD found varied results. For example, Corey-Lisle and colleagues, and Reutfors and colleagues found that 11.6% and 12.6% of individuals with PTD experienced TRD, respective,[21,36] while Olfson and colleagues found that 25.9% of individuals with PTD experienced TRD. [27]

The approach used by Corey-Lisle and colleagues to identify individuals with TRD was one of the most complex approaches used in the literature. This is because it relied on six different variables related to antidepressant therapy, including specification of augmentation therapies, the total number of drug switches, the quartile of drug switches compared to the cohort, the number of titrations of drugs, the quartile of titrations compared to the cohort, and being prescribed monoamine oxidase inhibitors. This sophisticated approach likely increases the

validity of the definition of TRD by including many factors associated with treatment-resistance and drug changes.

The definition of TRD used by Reutfors and colleagues included only whether or not a certain number of drug switches occurred but differed from other studies that used this simpler approach of defining TRD by using a more strict criterion to define combination therapy. Combination and augmentation therapies were not counted as independent drug changes if they were prescribed within 14 days of the start of an antidepressant. This approach would decrease the number of apparent drug changes in a given year, compared to other studies that counted all drugs dispensed in a given year. This study also used data from a national registry, which provides a more generalizable description of PTD compared to studies with more specific populations.

In comparison, Olfson and colleagues found that as many as 25.9% of individuals with PTD experienced TRD.[27] This study used a more liberal definition of TRD than those described previously. Rather than limit how a combination or augmentation therapy can be defined, this study counted the total number of different drugs prescribed in a given year. This approach can overestimate the proportion of TRD in the study by counting combination and augmentation therapies as being separate therapies from the drug they are being added to. Additionally, Olfson and colleagues only included Medicaid recipients, who are, therefore, more likely to have lower income.[27] As was seen in this thesis, individuals with lower income living in urban areas were more likely to meet the criteria for TRD than individuals with higher income. These reasons may, therefore, explain why the proportion of individuals with PTD who experienced TRD was higher in this study compared to the results of this thesis.

The only other study to use the MGH-s to identify TRD was that of Gibson and colleagues, which found that 28.1% of individuals experienced MGH-s scores greater than 3.[23] The approach that Gibson and colleagues used in applying the MGH-s to their cohort was different from the approach used in this thesis. In his original description on the MGH-s, Fava specifies that optimization of dose or duration of medications would mean a half-point increase in score.[17] However, Fava does not specify what optimization means. For this reason, I did not include these criteria in this thesis. In comparison, Gibson and colleague defined these criteria as *any* increase in dose and *any* treatment regimen with at least three refills for a medication.[23] This generous definition of dose and duration optimization likely resulted in many individuals experiencing multiple half-point increases in their study when it may not have been appropriate.

Additionally, each combination antidepressant therapy increases the score by a half-point for the MGH-s. Gibson and colleagues did not include this criterion in their methods but instead treated all antidepressant dispensation as being worth a full-point. Fava was not clear in his initial definition of combination therapies in 2003.[17] However, in the book, *Massachusetts General Hospital Comprehensive Clinical Psychiatry*, co-authored by Fava, the combination of any two antidepressants with different mechanisms of actions are treated as combination therapies.[24] This means that an antidepressant dispensed at the same time as a previously dispensed antidepressant should only contribute a half-point, rather than a full-point. Overall, it likely that Gibson and colleagues used defined TRD in such a way that it over-estimated the true proportion of TRD among individuals with PTD.

Overall, the definition of TRD can greatly affect what proportion of individuals with PTD are categorized as having TRD. In studies with more complex and stringent definitions of TRD, the proportion of TRD among individuals with PTD is approximately 12.0%, which is very

similar to the results of this thesis.[21,36] Studies that used more liberal definitions of TRD, the proportion is much higher.[23,27] I would, therefore, conclude that the results of this thesis are a more accurate representation of the proportion of TRD among individuals with PTD, compared to some previously published studies.

Another important comparison that must be discussed is the difference in the percentage of individuals meeting the criteria for TRD in clinical studies such as STAR*D and the findings of this thesis. In STAR*D, close to half of all study participants had not entered remission after receiving two different treatments for depression. In comparison, I found that only 11% of individuals met the criteria for TRD in this thesis. One reason for this difference is that participants of STAR*D were chosen on more stringent inclusion criteria where depression severity was evaluated by a psychiatrist and verified by a clinical research coordinator using the DSM-IV.[1] In comparison, individuals in this cohort were only required to have been diagnosed with MDD by any physician and received six weeks of antidepressant therapy to be included. Participants in STAR*D were, therefore, more likely to have been correctly diagnosed with a clinically relevant case of MDD compared to individuals in this cohort. This means that a high number of individuals in this cohort with MGH-1 may not have been included in STAR*D, which inflates the total cohort size of this thesis, and decreasing the percentage of individuals in the cohort who meet the criteria for TRD.

In addition to this difference in study participants, individuals in STAR*D were being monitored using a specific clinical trial protocol, which is more rigorous than what a patient would experience in a typical ambulatory visit. Because the majority of individuals in this cohort did not receive care from a psychiatrist, they likely experience less organized treatment for their MDD than did participants in STAR*D. Therefore, individuals in STAR*D may be more likely

to have proceeded from one step to the next, whereas individuals in this cohort may not have been prescribed additional medications after an initial non-response to treatment. This would decrease the percentage of individuals who would go on to meet the criteria for TRD, compared to STAR*D.

Health Services Use

In this section, I briefly describe the main results for ambulatory visits, ED visits, and hospitalizations, and explain some of the reasoning for the results. This includes describing sources of confounding and the effects of any confounding, and also describing some of the limitations of this study. I also describe the significance of the interactions that were analyzed. Finally, I then compare the results of this thesis to previous studies and offer explanations for why differences may exist.

Ambulatory visits

MGH-s scores of 2.5 and 3 were associated with a 38% increased risk for non-mental health related (non-MH) ambulatory visits, compared to MGH-1. Scores of 1.5 and 2 were associated with nearly half this increase, with a 22% and 18% increased risk for non-MH ambulatory visits, respectively. The risks associated with mental health related (MH) ambulatory visits were much higher. MGH-2.5 and MGH-3 were associated with a 287% and 421% increased risk for MH ambulatory visits, respectively, and the risks associated with MGH-1.5 and MGH-2 were 105% and 135% higher, respectively. For both MH and non-MH of ambulatory visits, MGH-s scores are positively associated with the risk for ambulatory visits.

A higher risk of ambulatory visits may be indicative of poorer health and greater needs, as healthy individuals may be less inclined to visit their physician compared to individuals in

poor health. Conversely, these higher risks may also represent increased access to health services and psychotherapy, and improved continuity of care. Although not performed in this thesis, it would be possible to examine the tariff codes associated with the ambulatory visits and investigate how individuals with different MGH-s scores may receive different treatments, such as psychotherapy, during their ambulatory visits. This approach would be an important next step into comparing health services use for individuals with PTD.

Age was an important effect modifier for the associated between MGH-s score and non-MH ambulatory visits. Generally, older individuals are more likely to seek ambulatory visits compared to younger individuals.[57] When the interaction between age and MGH-s scores was analyzed, MGH-2.5 and MGH-3 were more highly associated with the risk of non-MH ambulatory visits for younger individuals, and less highly associated for older individuals. This association means that the effect of TRD on ambulatory care is more important for younger individuals, and this association decreases with age. Because individuals are more likely to seek care for chronic health conditions as they age, it is expected that the effect of TRD becomes less important as age increases. This is supported by the fact that every one-point increase in Elixhauser comorbidity was associated with a 22% increased hazard for ambulatory visits, which shows the association between increased comorbidity – often associated with age – and increased ambulatory visits. This effect modification was not observed for MH ambulatory visits.

There are many important differences in the models for MH and non-MH related ambulatory visits. While female sex was associated with a 13% increased hazard for non-MH ambulatory visits, I did not find a clinically significant difference for the risk of MH ambulatory visits between males and females. This is important to highlight because males generally

experience lower rates of health services use than females, despite controlling for levels of morbidity.[43]

A one-point increase in SEFI (i.e. lower socioeconomic status) was associated with 12% increased risk for non-MH ambulatory visits, while a one-point decrease in SEFI (i.e. increased socioeconomic status) was associated with a 5% increased risk for MH ambulatory visits. Similarly, higher Elixhauser comorbidity was associated with a higher risk for non-MH ambulatory visits, while there was no associated between Elixhauser comorbidity and MH ambulatory visits. Overall, the covariates in the model for MH ambulatory visits were not highly associated with the risk for these visits. Instead, MGH-s scores were the most strongly associated variables in the model for MH visits, but were comparable to other variables in their association for the risk for non-MH visits.

Previous research has found similar findings, where TRD is associated with increased health services use. Compared to individuals with MDD, those with TRD experienced nearly twice as many outpatient visits.[21] Ivanova et al. found that individuals with TRD experienced a 1.7-times greater number of outpatient visits for all causes, and 2.4-times as many for mental health related care in a two-year period.[25] Though no other study has compared the hazards for ambulatory visits between individuals with TRD and PTD, the findings of this thesis are still consistent with the findings from previous literature.

This thesis is the first time to examine the association between MGH-s scores and ambulatory visits in a publicly-funded healthcare system. What is important to note is that even small increases in MGH-s score, which are less likely to meet the binary definition of TRD, are still associated with increased hazards for ambulatory care. Because previous studies did not evaluate the severity continuum of TRD based on MGH-s scores, they likely experienced some

form of misclassification, as individuals with MGH-s score of 1.5, 2, and 2.5 are more likely to be categorized as non-TRD. This would mean that previous results would be biased towards the null and underestimate the true association between TRD and ambulatory visits.

Finally, the increased risk for ambulatory visits associated with higher MGH-s scores is difficult to interpret because having contacts with the health care system is not necessarily a negative outcome. This increased risk may instead be indicative of individuals with TRD receiving more treatment from their physicians, such as psychotherapy. It may also mean that physicians follow these individuals more closely, and that follow-up care is more routinely scheduled. It would be important to separate the analysis for ambulatory visits where the associated tariff codes can differentiate whether psychotherapy is administered or not in order to test this. This sub-analysis would be an important next step in understanding the type of ambulatory care received by individuals with TRD.

Emergency Department Visits

The results for ED visits followed a similar trend to those for ambulatory visits. TRD was associated with a 55% increased risk for non-MH ED visits, compared to 34% and 25% increased risk for MGH-1.5 and MGH-2, respectively. When examining only MH ED visits, however, the risks were higher. MGH-2.5 and MGH-3 were associated with a 170% and 253% increased risk for ED visits, respectively, compared to 109% and 89% increased risk for MGH-1.5 and MGH-2, respectively. Overall, MGH-s scores were associated with increased risk for all ED visits, but the association was much higher when only examining MH ED visits.

In contrast to ambulatory visits, ED visits are generally thought of as a negative outcome. Increased risk for ambulatory visits may represent better access to primary care, whereas the

increased risk for ED visits may instead be associated with poorer access to primary care and reliance of ED visits as a substitute for primary care.[58,59] ED visits that are highly urgent or emergencies may also represent times of acute mental or physical health crises. The distinction between MH and non-MH ED visits in this analysis was not based on a diagnosis code. Instead, notes regarding a patient's chief complaint triage were used to distinguish MH and non-MH ED visits. As such, reasons for MH ED visits can include acute suicidality, while non-MH visits may include both physical health events (e.g. acute dehydration from diarrhea) and MH events like acute poisoning or trauma due to intentional overdose or self-harm. The differentiation between MH and non-MH ED visits must, therefore, be made with some caution as some non-MH ED visits may be associated with a mental health concern.

Of the covariates included in the model, SEFI was the most strongly associated with increased risk for all ED visits, when ignoring urban status (reason described below). For both MH and non-MH ED visits, increases in SEFI (i.e. lower socioeconomic status) was associated with a 50% increased hazards for ED visits. Therefore, individuals who are living in areas with a one standard deviation decrease in SEFI experienced a 50% increased risk for ED visit compared to individuals with the average SEFI of Manitoba. As lower income has been associated with increased emergency department use, often to substitute access to primary care, it is possible that this difference is not a result of increased health crises, but instead, represent a lack of access to primary care.[58]

In contrast to the ambulatory care analyses, having a past year psychiatrist visit (PYPV) could be included in this analysis. PYPV was associated with a 171% increased risk for MH ED visits. In comparison, PYPV was associated with a 19% increased risk for non-MH ED visits. Therefore, individuals who have been in the care of a psychiatrist are at a higher risk for

experiencing a MH ED visit than those who have not. It is possible that having a PYPV is associated with symptom severity of more complex needs, which would increase the risk for more intensive health care needs. However, individuals with higher MGH-s scores were more likely to have ever had a psychiatrist visit. As such, it is possible that the association between these two variables is the reason why PYPV is associated with this increased risk.

Sex did not have a clinically significant association with non-MH ED visits. In comparison, being male was associated with a 37% increase risk for MH ED visits. However, sex was also an effect modifier for MGH-s scores, as shown in Tables 12. When the results were stratified by sex, TRD was more highly associated with MH ED visits for females than for males. There was no difference, however, for MGH-s scores of 1.5 or 2 when stratifying the analysis. Females with TRD may, therefore, experience mental health related crises requiring emergency care more frequently than males with TRD. Another possibility is that females with TRD are more likely to seek emergency care when experiencing a mental health crisis than males.

Because EDs are not the only available resource for mental health crises, it is not possible to draw a definitive conclusion from this analysis. For example, the Winnipeg Regional Health Authority's Crisis Response Centre opened in May 2013, and the data for visits to this centre are not included in the Repository. Therefore, the data from May 2013 to April 2015 may not reflect the actual amount of mental health emergency needs in Winnipeg. A sub-analysis can be performed to compare risks from 2007 to 2013, and from 2013 to 2015 to address this.

Another limitation in this analysis is that it only includes data from Winnipeg EDs. This limitation is a likely reason that urban status was so strongly associated with increased hazards for ED visits. A separate analysis was performed to only include data from individuals living in

Winnipeg. The results from this analysis (not shown) were not significantly different from the analysis for the entire cohort.

Two studies have previously compared the ED visits between individuals with TRD and those with PTD. Olfson and colleagues found TRD was associated with 65% increased odds (95% CI: 1.32 – 2.07) for a MH ED visit over one year, compared to all other PTD individuals.[27] In comparison, TRD was associated with a 28% increased odds (95% CI: 1.12-1.45) for an all-cause ED visits for the same period.[27] Ivanova and colleagues found that, over two years, individuals with TRD experienced a mean of 2.0 (SD=5.7) all-cause ED visits, compared to 1.2 (SD=4.0) for PTD individuals.[25] For MH ED visits, TRD individuals experienced a mean of 0.1 (SD=0.4) visits, compared to 0.03 (SD=0.25) for PTD individuals.[25] Though no study has compared the hazards for ED visits, these results are similar to those found in this thesis, where individuals with TRD experienced a higher likelihood of ED visits for both non-MH and MH ED visits.

Overall, MGH-s scores are associated with increased risk for all ED visits. Even MGH-s scores of 1.5 and 2, which would not meet the binary definition of TRD, were associated with an increased risk for ED visits compared to individuals with the baseline MGH-s score of 1. This further supports the hypothesis that TRD follows a severity continuum, where higher scores are associated with worse health outcomes. ED visits are not necessarily indicative of worse health, however, but may instead indicate decreased access to primary care, so some caution in this interpretation is needed.

Hospitalizations

The results found for the hazards for hospitalizations are similar to those found for ambulatory visits and ED visits. TRD was associated with a 47% higher risk for non-MH hospitalizations, while the increased risk for HM hospitalizations for MHGH-2.5 and MGH-3 were 385% and 733%, respectively. The risks for all hospitalizations increased as MGH-s scores increased, but the risks for MH hospitalizations were much greater than the risks for non-MH hospitalizations.

The hazards associated between age, sex, SEFI, and Elixhauser were nearly identical for MH and non-MH hospitalizations. However, the hazards for urban status and having a past year psychiatrist visit were significantly different between both sets of hospitalizations. Living in Winnipeg or Brandon was associated with a 20% decreased risk for non-MH hospitalizations, compared to a 52% decreased risk for MH hospitalizations. This increased risk for MH hospitalizations for individuals living in rural Manitoba may be explained in part to decreased access to ambulatory psychiatric care in rural Manitoba,[44] and a resulting higher reliance on inpatient hospitalizations for mental health services.

The increased risk for MH hospitalizations associated with having a PYPV may be due to the higher likelihood for individuals with more severe MDD being seen by a psychiatrist, which would increase the risk for hospitalizations. This variable also proved to be an important effect modifier for the association between MGH-s scores and MH hospitalizations. For individuals with MGH-3, having a PYPV was associated with more than a 12-fold increased risk for MH hospitalizations, compared to a 6-fold increased risk for individuals without a PYPV. It is also possible that a PYPV is not associated with depression severity for individuals with TRD but is

instead associated with active cases of MDD. Overall, it is difficult to draw a definitive conclusion regarding the significance of this variable using administrative data.

Another important effect modifier for the association between MGH-s scores and MH hospitalization was sex, where females experienced a higher risk for MH hospitalizations than males with the same MGH-s score. A brief analysis (not presented in this thesis) found that there was no significant difference in the percentage of males and females who ever received an ambulatory visit from a psychiatrist. These effect modifiers are therefore not likely to be associated with each other. Similar to the results found for the risk for ED visits, it possible that females are more likely to seek care for their mental health compared to males.

Olfson and colleagues compared the proportion of individuals with and without TRD who experience any inpatient hospital admissions over one year. TRD was associated with a 58% increased odds of experiencing an inpatient admission for any reason, and a 97% increased odds of experiencing an inpatient admission for a mental health reason, compared to individuals without TRD.[27] The odds of experience non-MH hospitalizations were comparable to the hazards found in this study. However, the hazards for MH hospitalizations that I found were significantly higher than the odds found by Olfson and colleagues. One possible explanation is that individuals without TRD in Olfson and colleagues' study may have experienced some for treatment-resistance and would have been categorized as MGH-1.5 or MGH-2 in this thesis. This would bias the results towards the null, resulting in the relationship between TRD and MH hospitalizations to appear smaller. Another explanation is that individuals with higher MGH-s scores experienced multiple, recurrent MH hospitalizations, which would increase the hazards for MH hospitalization, but would not have been captured with the logistic regression performed

by Olfson and colleagues. This highlights the importance of using the Andersen-Gill model for determining the hazards for recurrent events.

Overall, a severity continuum of TRD was also found for hospitalizations. In this case, it is difficult to associated hospitalizations with anything other than more severe forms of depression and worse health outcomes. As such, screening individuals with MDD for their respective MGH-s score can be an important clinical intervention in identifying individuals at high risk for a MH hospitalization.

Potential bias

The increased risk for health services use as MGH-s scores increased may be explained in part by an increase in help-seeking behaviour experienced by individuals with higher MGH-s scores. However, MGH-s scores were treated as time-dependent variables, so all individuals enter the cohort with a score of 1. This means that the time-at-risk for MGH-1 includes individuals who go on to reach higher scores. It may be assumed that an individual with increased help-seeking behaviour, which contributed to having a higher MGH-s score, would have had this help-seeking behaviour while they had lower MGH-s scores. This would, in turn, increase the risks for events for lower scores, similarly to how they would be increased for higher scores. This depends on the relative amount of time-at-risk that individuals contribute to each MGH-s score, though. It is therefore important to understand that the relationship between MGH-s score and health services use may be biased away from the null as individuals with higher MGH-s scores may be more likely to seek care than other individuals, by nature of the definition of MGH-s score increases.

One possible interpretation for the increased hazards for MH health services use for higher MGH-s scores can be described as follows: (1) individuals who experience these events do so in part because of untreated symptoms, (2) having a contact with a health care provider is an opportunity to address untreated mental health symptoms and change drugs, (3) changing drugs leads to an increase in MGH-s score. This would mean that experiencing a MH health care contact is predictive of an increase in MGH-s score, rather than increased MGH-s scores predicting health services use. This was avoided, however, by treating MGH-s scores as a time-dependent variable. This means that changes to MGH-s scores occurred before the events that are associated with each score, meaning these events could not affect the associated MGH-s score. If MGH-s scores were not treated as time-dependent, then all events that occurred throughout an individual's time in the cohort would have attributed to their highest MGH-s score.

Summary

The association between MGH-s scores and health services use followed a consistent pattern for ambulatory visits, ED visits, and hospitalizations. For all non-MH visits, a small increase in the risks for health services use was found as MGH-s scores increased. When examining only MH visits, the association between MGH-s scores and health services use became much greater. As the intensity of care increased – where hospitalizations are more intensive than ambulatory visits – so too did the association between TRD-levels of MGH-s and the hazards for health services use. MGH-2.5 and MGH-3 were associated with a higher risk for ambulatory visits than MGH-1, but the association was much greater for hospitalizations.

Age, sex, and having a PYPV were all important effect modifiers in these analyses. For individuals with MGH-2.5 and MGH-3, younger individuals, females, and those with a PYPV

experienced a higher risk for certain health services than other individuals with the same MGH-s score. When examining the effect modification of age, the observed effect may be due to older individuals having other reasons to require health services use, which decreases the observed effect of MGH-s on this. For sex, it is possible that males with TRD may experience decreased health services use, rather than females experienced an increased risk. Finally, having a PYPV may also be a sign of increased depression severity or an increased likelihood of a case of MDD being active, but these hypotheses could not be tested in this thesis.

All-Cause Mortality

The findings of this thesis show that treatment-resistance is associated with increased risk for all-cause mortality. In contrast to the findings for health services use, there was not a significant trend for MGH-s to predict the risk for all-cause mortality. However, this was only true for older individuals. As age decreased, a noticeable trend emerged where higher MGH-s scores were associated with an increased risk for all-cause mortality. This can be expected because the risk of death from causes other than MDD increases with age, which would decrease the effect of TRD on the risk of death.

In order to better understand the effect of TRD on all-cause mortality, it would have been important to examine the cause of death, and whether it was associated with mental health or not. Because TRD was more highly associated with increased risk for MH health services use, it is likely that TRD would also be more highly associated with deaths associated with suicide than all other causes. This would be an important next step in this field of research and it is possible to perform this analysis at MCHP by including data from the Vital Statistics dataset. The study by Reutfors and colleagues, the only other study that has analyzed the association between TRD and the risk for mortality, used this approach of performing a sub-analysis by cause of death.[36]

Reutfors and colleagues found that TRD was associated with a 35% increase in all-cause mortality compared to PTD. Like the results of this thesis, the risks were higher for younger individuals compared to older ones. In contrast, the results of this thesis show that MGH-2.5 and MGH-3 are associated with 115% and 87% increases in the risk for all-cause mortality, respectively, and MGH-1.5 and MGH-2 were associated with 89% and 47% increased risk. When comparing these results to those of Reutfors and colleagues, it is likely that some individuals defined as not having TRD in the Reutfors study would have had MGH-s scores greater than 1, and thus experienced higher risks for mortality than others in the non-TRD group. By including these individuals in the non-TRD group, the hazard ratios would be biased towards the null. It is therefore possible that using a binary definition for TRD underestimated the effect of TRD on mortality, highlighting the importance of using a staging method in defining TRD.

The study by Reutfors and colleagues also found that the association between TRD and mortality was significantly greater when only "external causes" of death (e.g. suicides and accidents) were analyzed.[36] The study found a hazard ratio of 1.97 (95% CI: 1.69 – 2.29) for death from external causes. Therefore, a vital next step in studying the association between MGH-s scores and mortality is to include data for cause of death and perform a sub-analysis for suicides.

Importance

This study was the first to describe TRD in a Canadian setting, both in terms of its trends and demographics, but also in terms of how TRD is associated with increased risks for health services use and mortality. As described in the introductory section of this thesis, TRD has been defined in many ways, with most studies defining TRD as a binary condition. I instead applied the MGH-s staging method to test the hypothesis that TRD follows a severity continuum. In

comparing the results of this thesis to those from previous research, there are instances where the effect of TRD on the outcome is significantly higher in this study. Part of the reason for this may be that previous studies have included individuals who would otherwise have had MGH-s scores greater than 1 in the non-TRD group, biasing the results towards the null. This thesis confirms the importance of including treatment-resistance as a covariate and using a staging model in defining treatment-resistance for future studies that investigate the burden of depression on health services use, mortality, or other outcomes of interest.

In terms of clinical relevance, this thesis found that treatment-resistance is associated with serious adverse events, including ED visits, hospitalizations, and death. A simple recommendation for healthcare providers that encounter individuals experiencing an episode of MDD would be to take a thorough medical history and determine their history of treatment-resistance. The *Massachusetts General Hospital Antidepressant Treatment Response Questionnaire (ATRQ)* is a validated questionnaire that can help health care providers identify individuals who are at a higher risk for serious adverse events and ensure that the appropriate level of care is provided.

Table 4 shows that only 56.6% of individuals with MGH-3 have ever had an ambulatory visit with a psychiatrist. The findings of this thesis show, however, that these individuals are at a much greater risk of experiencing serious adverse events. Because psychiatrists are better trained in providing pharmaceutical therapy and psychotherapy for individuals with MDD, primary care physicians must be encouraged to screen patients with MDD for treatment-resistance. Individuals who are shown to have a history of treatment-resistance must be referred to a psychiatrist more often in order to ensure they receive optimal care.

Overall, this thesis has shown that treatment-resistance in MDD is a serious condition affecting a non-negligible proportion of the population in Manitoba. Individuals with TRD are at a higher risk for serious outcomes and must be properly screened in primary care. When possible, individuals with TRD should receive specialist care from psychiatrists, which is currently an area for improvement in Manitoba.

Strengths

The primary strength of this study is the number of years of available data, and the generalizability of the study population. No other study on the epidemiology of TRD has included over 20 years of data. Similarly, most other studies have relied on commercial insurance claims data, which is not representative of all individuals in a population. This study is also the first to include a measure of socioeconomic status, which is an important predictor in many of the outcomes of interest.

By using the MGH-s staging method, I used a more rigorous approach in determining an individual's TRD status than has been used in many previous studies. For example, in the study by Olfson and colleagues, the researchers counted any changes in medication as contributing to an individual's potential TRD status.[27] When examining data from the DPIN dataset, however, it was clear why this was not a valid approach to defining treatment resistance. Generally, most individuals in this thesis had tried many different antidepressants over the course of their life. If any change in medication, regardless of how long it was since the previous drug was taken, was included in an episode of treatment-resistance, the prevalence of TRD would have been much higher. Instead, this thesis defined 120 days as the maximum amount of time to categorize a particular episode of dispensation, following from the methods used in Kubitz et al, and Fife et

al.[29,30] Using this approach required more time and effort spent performing data management but ensured that episodes of care more accurately reflect an individual's TRD status.

The choice of statistical models used is also a first in studying the epidemiology of TRD and its associated outcomes. Because TRD should be included as a time-dependent variable, it was essential to use a model that would allow for this. The Andersen-Gill generalization of the Cox proportional hazards model allows for precise time-at-risk to be calculated, and for precise changes in time-dependent variables. Additionally, calculating hazard ratios for recurrent events is a superior choice than calculating odds ratios, as this allows for a more accurate reflection of the effect of multiple events occurring for the same individual. Some limitations do exist for this model and are described below.

Limitations

There are many limitations that must be considered when interpreting the results of this thesis. The two types of limitations that I identified are those that are associated with the methods of analysis and those associated with limitations that come from using administrative data. In this section, I will describe the limitations that I identified, and offer solutions on how these can be improved in future research.

Model limitations

Using the Andersen-Gill method for modelling the hazards of recurrent events over time assumes that the hazards are continuous for the entire study period, which may not be appropriate. Year of cohort entry was, therefore, included in all models to control for its relationships with each outcome. Another solution to this would have been to divide the 21 years of the study period into four 5-year epochs.

Additionally, this model assumes that the hazards are constant over the entire time-at-risk for all individuals, which may also not be appropriate. As a hypothetical example, it is possible that the risk associated with mortality is higher for the first year following an MGH-s score change, since MGH-s score changes may represent times of increased difficulty or crisis for individuals. It would have been appropriate to investigate the hazards associated with different amounts of time-at-risk following each MGH-s score increase in order to study this.

A final limitation with the Andersen-Gill model is that it assumes the risk for each recurrent event is the same. For example, this assumes the risk for an individual's first hospitalization is equivalent to the risk for their third, which may not be accurate. The Prentice-Williams-Peterson generalization of the Cox proportional hazards model is designed to examine the order of recurrent events and stratifies the analysis accordingly.[60] Therefore, separate hazards are calculated for each recurrent outcome. The limitation of the Prentice-Williams-Peterson approach is that data for individuals with a higher number of recurrent events may not be included in the analysis, as there may be an insufficient number of individuals with a higher number of recurrent events for sufficient power to be achieved for these strata. Overall, I chose the Andersen-Gill model for its ability to measure the relationship of time-dependent variables and the outcome more accurately. However, some improvements should be done in future research in order to better represent the hazards for outcomes over time, and over time-at-risk.

Data limitations

When using administrative data, we are limited by the quality and type of data that are collected. In the Repository, there are no data associated with an individual's depression severity. As such, I could not differentiate between individuals with mild, moderate, or severe levels of depression severity, and could not adjust for depression severity. The causal pathway between

depression severity and treatment-resistance is not fully known,[61] although TRD has been associated with mild and severe case of MDD.[62,63] Including depression severity in future analyses would be important in reducing potential confounding, but this is not possible when relying exclusively on the data from the Repository. One solution would be to perform a cohort study where data from the Repository are linked to individual patient files, but this would be very labour-intensive.

Another way of improving this study would be to include psychiatric comorbidities in the models. Other studies have included the diagnosis of substance use disorders and anxiety disorders in their analyses.[27,36] Controlling for these diagnoses would be possible by including them in the models, or by stratifying the analyses for individuals with these diagnoses. This analysis was not performed due to time constraints.

The final limitations that I identified are related to assumptions that must be made when using administrative data. As mentioned in an above section, I could not distinguish active cases of PTD and TRD from inactive cases. Previous studies have used a 120-day cut off to determine active episodes of MDD from inactive ones.[29,30] While this approach was used in separating dispensation episodes, I believe that it would not have been appropriate for distinguishing individuals experiencing symptoms of depression from those who are not. For example, an individual may be dispensed antidepressants long after their symptoms have subsided. In this case, MDD can be thought of as being inactive, even though the individual is dispensed antidepressants. Alternatively, an individual may discontinue their antidepressant therapies and may experience a relapse in depression symptoms by not seek medical treatment. In both these examples, an individual's record of dispensations or in being diagnosed with MDD at ambulatory visits may not accurately reflect whether their symptoms are present or not.

The most important limitation in this thesis is the assumption of treatment-resistance as the reason for which individuals are dispensed different medications. In all studies that have used administrative data in defining TRD, the authors assumed that individuals who are dispensed different antidepressants do so because they did not respond to the previous medication they were taking. Alternative explanations for changing medications can instead be drug allergies, adverse events and side-effects that outweigh the benefit of the medication, changes in insurance coverage, changes in physician and patient preference for treatment, changes in best-practice guidelines, new drugs coming to market with increased safety profiles, drug interactions from polypharmacy, and so on.

However, all previous research on the epidemiology of TRD that used administrative data has had to assume that drug changes are made because of treatment-resistance. This assumption is supported by the fact that physicians and patients may be reluctant to change medications unless a change was necessary due to the possibility of antidepressant discontinuation symptoms. If the assumption that drug changes are due to treatment-resistance is not always valid, then the prevalence of TRD may be overestimated in this thesis and all previous research on the epidemiology of TRD. Conversely, the prevalence of TRD may also be underestimated due to individuals discontinuing any pharmaceutical treatment after failed attempts of improving their symptoms. This underestimation is possible due to the inability of administrative data to follow individuals over time who no longer choose to seek medical treatment for their health. Therefore, it may not be possible to improve upon the limitations that are associated with administrative data, and that the strengths of this thesis likely outweigh its limitations.

Conclusion

The primary objectives of this thesis were to describe the epidemiology of TRD in Manitoba, and then to compare individuals with TRD to those with PTD in terms of the risks for health services use and mortality. The percentage of individuals who met the criteria for TRD in this thesis was much lower than what was previously described in clinical trials like STAR*D but was comparable to the results found in many previous retrospective cohort studies. This was due to differences between rigorous inclusion criteria in place in clinical trials in terms of MDD diagnostics, compared to what is typically the case in an ambulatory care setting. The similarity in the percentage of individuals with TRD in this thesis while using the MGH-s to those of other cohort studies using binary definitions is compelling evidence for the continued use of staging methods like the MGH-s in studying TRD.

The sociodemographic descriptions of individuals with TRD to those with PTD were not substantially different, even if statistically significant differences were found. Many of these differences may have associated with the likelihood for different demographic groups to experience increases in MGH-s scores rather than being indicative of important demographic risk factors for TRD. However, the results of this thesis have shown that younger individuals with PTD were more likely to meet the criteria for TRD and that younger ages had a moderating effect that increased the association between MGH-s scores and outcomes like ED visits and all-cause mortality.

The results of this thesis have shown that TRD is not only associated with higher risks for health services use, including ED visits and hospitalizations, but also for all-cause mortality. Additionally, the association between TRD and these outcomes followed a severity continuum where the association increased as MGH-s scores increased. These results confirm the findings

from the only other study on the association between TRD and mortality,[36] while also showing the importance of describing TRD as a condition with discreet stages of severity.

The methods used in this thesis, including the use of the MGH-s and the use of the Andersen-Gill generalization of the extended Cox proportional hazards model, were novel improvements on many of the previous studies on this topic. Much of the previous literature on the association between TRD and health services use have used health care costs as their preferred metric of analysis. This metric may be useful for for-profit insurance companies and researchers that rely on their data, but describing the risk for ED visits and hospitalizations can instead frame this issue with a patient-centred approach.

One important limitation of this thesis was my inability to measuring the level of depression severity of individuals. Such information does not exist in the Repository, and cannot be reasonably approximated. However, the decision was made to consider how receiving care from a psychiatrist may influence the analyses. It can be argued that individuals receiving care from a psychiatrist may require more specialized treatment for reasons such as additional psychiatric comorbidities or more severe depression. This covariate proved to be an important predictor in many of the models of this thesis, and additional research into its use should be explored.

Overall, this thesis has provided many valuable contributions to the current literature regarding TRD. First, there was previously a gap in the literature where nearly ever study on this topic used data from American commercial health insurance claims datasets. These datasets are generally not representative of a general population and are from a country without universal access to healthcare. This thesis has filled these gaps by being the first to describe the

epidemiology of TRD among a general population, over a 20-year timeline, in a country with universal access to healthcare.

Another important gap was that all retrospective cohort studies on TRD, except one, have used binary definitions. The most likely reason for this is that the binary definition of a failure to respond to at least two courses of antidepressant therapies has been generally agreed upon, while there is no consensus on the best staging method to be used. Additionally, the data management necessary to study TRD using a staging method is more labour intensive than what is required when defining TRD using a binary method. However, this thesis has shown that using a staging method definition is not only feasible, but it should also be encouraged. This thesis has provided significant evidence that TRD follows a severity continuum that can be model well by using a staging method such as the MGH-s.

Following from this evidence supporting the use of staging models for TRD comes one important recommendation for researchers and clinicians: when studying or treating individuals with depression, their TRD status should be determined using a staging definition. For researchers, this means agreeing upon the most valid staging method, and ensuring that it is used consistently in order to generate comparative results. This is largely missing in the current literature on TRD. Too many different definitions have been used over time, which has resulted in results that are difficult to be compared. For clinicians, especially primary care physicians, it is important to screen patients presenting with depression for their prior history of antidepressant use. Beyond this, however, it is essential to know that individuals who are not responding to antidepressants are at an increased risk for serious events such as ED visits, mental health hospitalizations, and death. Improved mental health care, such as care from a psychiatrist or

clinical psychologist, should be explored for these individuals. However, if patients are not adequately screened for TRD, it is impossible to know who requires this additional care.

Exploring the relationship between TRD and specialized psychiatric care would be an important continuation of this research. Questions that arose during the completion of this thesis were how individuals with TRD experience specialized psychiatric care when compared to other individuals with PTD. For example, how do rates of psychotherapy compare between these two groups? How quickly are individuals with TRD referred to a psychiatrist compared to those with PTD? Does having access to care from a psychiatrist impact the likelihood that an individual will meet the criteria for TRD? All of these questions would require a significant continuation of research but could improve how individuals with TRD receive care in Manitoba.

In conclusion, this thesis has found that individuals with MDD are not a homogenous group. Instead, individuals will experience a continuum of treatment-resistance that impacts their risks for requiring health services use and their risk for death. Though individuals on the higher end of this continuum were more likely to have ever received care from a psychiatrist, half of all individuals with TRD in Manitoba were never seen by a psychiatrist. More needs to be done to effectively screen these individuals while they are in primary care in order to prevent such negative health outcomes.

References

1. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg A a. Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Seve ... 2006;(November):1905-1917. doi:10.1176/appi.ajp.163.11.1905

2. Lam RW, McIntosh D, Wang J, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: Section 1. Disease burden and principles of care. *Can J Psychiatry*. 2016;61(9):510-523. doi:10.1177/0706743716659416

3. Lo C, Calzavara A, Kurdyak P, et al. Depression and use of health care services in patients with advanced cancer. *Can Fam Physician*. 2013;59(3):e168-e174. http://www.cfp.ca/content/59/3/e168.full.pdf+html%0Ahttp://0-ovidsp.ovid.com.lib.exeter.ac.uk/ovidweb.cgi?T=JS&PAGE=reference&D=emed15&NEWS=N&AN=368552567.

4. Egede LE, Zhend D, Simpson K. Comorbid Depression is Associated With Increased Health Care Use and Expenditures in Individuals with Diabetes. *Diabetes Care*. 2002;25(3):464-470. http://www.cfp.ca/content/59/3/e168.full.pdf+html%0Ahttp://0-ovidsp.ovid.com.lib.exeter.ac.uk/ovidweb.cgi?T=JS&PAGE=reference&D=emed15&NEWS=N&AN=368552567.

5. Thomson W. Lifting the shroud on depression and premature mortality: A 49-year follow-up study. *J Affect Disord*. 2011;130(1-2):60-65. doi:10.1016/j.jad.2010.09.028

6. Murphy JM, Gilman SE, Lesage A, et al. Time trends in mortality associated with depression: findings from the Stirling County study. *Can J Psychiatry*. 2010;55(12):776-

783. doi:10.1177/070674371005501205

7. 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(1):23-30. doi:10.1177/070674371506000106

8. Drapeau A, Boyer R, Diallo FB. Discrepancies between survey and administrative data on the use of mental health services in the general population: findings from a study conducted in Québec. *BMC Public Health*. 2011;11(1):837. doi:10.1186/1471-2458-11-837

9. Manitoba Centre for Health Policy. Concept: Manitoba Population Research Data Repository. Concept Dictionary. http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1419. Published 2016.

10. Chartier M, Bolton J, Mota N, et al. *Mental Illness Among Adult Manitobans*. Winnipeg, MB: Manitoba Centre for Health Policy; 2018.

11. Organization for Economic Cooperation and Development. Antidepressants consumption, 2000 and 2011 (or nearest year available). In: *Making Mental Health Count: The Social and Economic Costs of Neglecting Mental Health Care*. Paris: OECD Health Policy Studies, OECD Publishing; 2014. doi:, https://doi.org/10.1787/9789264208445-graph17-en.

12. Organization for Economic Cooperation and Development. Antidepressant drugs consumption, 2000 and 2015 (or nearest year). In: *Health at a Glance 2017*. Paris: OECD Publishing; 2017. doi:https://doi.org/10.1787/health_glance-2017-graph181-en

13. Conway CR, George MS, Sackeim HA. Toward an Evidence-Based , Operational Definition of Treatment-Resistant Depression When Enough Is Enough. *JAMA Psychiatry*. 2017;74(1):9-10. doi:10.1001/jamapsychiatry.2016.2586

14. Scott J. Epidemiology, Demography and Definitions. *Int Clin Psychopharmacol*. 1991;6 (Supp 1):1-12.

15. Fava M, Davidson KG. Definition and epidemiology of treatment-resistant depression. *Psychiatr Clin North Am*. 1996;19(2):179-198. doi:10.1016/S0193-953X(05)70283-5

16. Souery D, Amsterdam J, De Montigny C, et al. Treatment resistant depression: Methodological overview and operational criteria. *Eur Neuropsychopharmacol*. 1999;9(1-2):83-91. doi:10.1016/S0924-977X(98)00004-2

17. Fava M. Diagnosis and Definition of Treatment-Resistant Depression. *Biol Psychiatry*. 2003;3223(03):649-659. doi:10.1016/S0006-3223(03)00231-2

18. Ruhé HG, Van Rooijen G, Spijker J, Peeters FPML, Schene AH. Staging methods for treatment resistant depression. A systematic review. *J Affect Disord*. 2012;137(1-3):35-45. doi:10.1016/j.jad.2011.02.020

19. Petersen T, Papakostas GI, Posternak MA, et al. Empirical Testing of Two Models for Staging Antidepressant Treatment Resistance. *J Clin Psychopharmacol*. 2005;25(4):336-341. doi:10.1097/01.jcp.0000169036.40755.16

20. Fogelson DL, Leuchter A. Defining Treatment-Resistant Depression. *JAMA Psychiatry*. 2017;74(7):758-759. doi:10.1001/jamapsychiatry.2017.0837

21. Corey-Lisle PK, Birnbaum HG, Greenberg PE, Marynchenko MB, Claxton AJ.

Identification of a claims data “signature” and economic consequences for treatment-resistant depression. *J Clin Psychiatry*. 2002;63(8):717-726. doi:10.4088/JCP.v63n0810

22. Gibson T, Jing Y, Bagalman J, et al. Identification of treatment resistant depression in claims data: a comparison of methods. *Value Heal*. 2006;12(3):A187. doi:10.1016/S1098-3015(10)73999-1

23. Gibson T, Jing Y, Carls G, et al. The Cost Burden of Treatment Resistance in Patients With Depression. *Value Heal*. 2009;12(3):A177. doi:10.1016/S1098-3015(10)73947-4

24. Baek JH, Nierenber AA, Fava M. Pharmacological Approaches to Treatment-Resistant Depression. In: *Massachusetts General Hospital Comprehensive Clinical Psychiatry*. 2nd ed. New York: Elsevier, Inc; 2016:506-506.

25. Ivanova JI, Birnbaum HG, Kidolezi Y, Subramanian G, Khan SA, Stensland MD. Direct and indirect costs of employees with treatment-resistant and non-treatment-resistant major depressive disorder. *Curr Med Res Opin*. 2010;26(10):2475-2484. doi:10.1185/03007995.2010.517716

26. Olchanski N, McInnis Myers M, Halseth M, et al. The Economic Burden of Treatment-Resistant Depression. *Clin Ther*. 2013;35(4):512-522. doi:10.1016/j.clinthera.2012.09.001

27. Olfson M, Amos TB, Benson C, McRae J, Marcus SC. Prospective Service Use and Health Care Costs of Medicaid Beneficiaries with Treatment-Resistant Depression. *J Manag Care Spec Pharm*. 2018;24(3):226-236. doi:10.18553/jmcp.2018.24.3.226

28. Canadian Institute for Health Information. How does Canada’s health spending compare internationally? National Health Expenditure Trends. https://www.cihi.ca/en/how-does-

canadas-health-spending-compare-internationally.

29. Kubitz N, Mehra M, Potluri RC, Garg N, Cossrow N. Characterization of Treatment Resistant Depression Episodes in a Cohort of Patients from a US Commercial Claims Database. *PLoS One*. 2013;8(10):1-9. doi:10.1371/journal.pone.0076882

30. Fife D, Feng Y, Wang MYH, et al. Epidemiology of pharmaceutically treated depression and treatment resistant depression in Taiwan. *Psychiatry Res*. 2017;252(March):277-283. doi:10.1016/j.psychres.2017.03.006

31. Patten SB, Williams JVA, Lavorato DH, Fiest KM, Bulloch AGM, Wang J. Antidepressant use in Canada has stopped increasing. *Can J Psychiatry*. 2014;59(11):609-614. doi:10.1177/070674371405901107

32. Mahlich J, Tsukazawa S, Wiegand F. Estimating Prevalence and Healthcare Utilization for Treatment-Resistant Depression in Japan: A Retrospective Claims Database Study. *Drugs - Real World Outcomes*. 2018;5(1):35-43. doi:10.1007/s40801-017-0126-5

33. Kennedy SH, Lam RW, Bhat V, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder : Section 3 . Pharmacological Treatments. *Can J Psychiatry*. 2016:1-21. doi:10.1177/0706743716659417

34. Mrazek DA, Hornberger JC, Altar CA, Degtiar I. A Review of the Clinical, Economic, and Societal Burden of Treatment-Resistant Depression: 1996–2013. *Psychiatr Serv*. 2014;65(8):977-987. doi:10.1176/appi.ps.201300059

35. Pfeiffer P, Kim H, Ganoczy D, Zivin K, Valenstein M. Treatment-Resistant Depression

and Risk of Suicide. *Suicide Life-Threatening Behav*. 2013;43:356-365. doi:10.1111/sltb.12022

36. Reutfors J, Andersson TML, Brenner P, et al. Mortality in treatment-resistant unipolar depression: A register-based cohort study in Sweden. *J Affect Disord*. 2018;238(June):674-679. doi:10.1016/j.jad.2018.06.030

37. Manitoba Centre for Health Policy. Concept: Ambulatory Visits - Physician. Concept Dictionary. http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1069. Published 2015.

38. Manitoba Centre for Health Policy. Concept: Home Care and Hospitalizations. Concept Dictionary. http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1146. Published 2004.

39. Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A large Sample Study. *Ann Stat*. 1982;10(4):1100-1120.

40. Balshaw R. In-person meeting. 2019:April 4.

41. Beck CA, Williams JV, Li Wang J, et al. *Psychotropic Medication Use in Canada*. Vol 50.; 2005. https://journals-sagepub-com.uml.idm.oclc.org/doi/pdf/10.1177/070674370505001006. Accessed April 27, 2019.

42. Morkem R, Barber D, Williamson T, Patten SB. A Canadian primary care sentinel surveillance network study evaluating antidepressant prescribing in Canada from 2006 to 2012. *Can J Psychiatry*. 2015;60(12):564-570. doi:10.1177/070674371506001207

43. Bertakis KD. Gender Differences in the Utilization of Health Care Services. *J Fam Pract*.

2000;49(2). https://www.mdedge.com/familymedicine/article/60747/womens-health/gender-differences-utilization-health-care-services#bib4. Accessed May 7, 2019.

44. Virgo Planning and Evaluation Consultants Inc. Improving Access and Coordination of Mental Health and Addiction Services: A Provincial Strategy for all Manitobans. 2018:1-279. https://www.gov.mb.ca/health/mha/docs/mha_strategic_plan.pdf.

45. Fernandez L, MacKinnon S, Silver J. *The Social Determinants of Health in Manitoba*. 2nd ed. Winnipeg: Canadian Centre for Policy Alternatives, Manitoba office; 2015.

46. Health Canada. Drug Product Database online query. Drugs and Health Products. https://health-products.canada.ca/dpd-bdpp/. Published 2018.

47. Raymond CB, Morgan SG, Caetano PA. Antidepressant Utilization in British Columbia From 1996 to 2004: Increasing Prevalence but Not Incidence. *Psychiatr Serv*. 2014;58(1):79-84. doi:10.1176/ps.2007.58.1.79

48. Barbey JT, Roose SP. SSRI safety in overdose. *J Clin Psychiatry*. 1998;59(SUPPL. 15):42-48.

49. Zohar J, Westenberg M. Anxiety disorders: a review of tricyclic antidepressants and selective serotonin reutake inhibitors. *Acta Psychiatr Scand*. 2000;101(6):39-49. http://onlinelibrary.wiley.com/store/10.1111/j.1600-0447.2000.tb10947.x/asset/j.1600-0447.2000.tb10947.x.pdf?v=1&t=j9yv2c1l&s=16709e93b16b1113cadf6a4fbf9219ee5a0a7f7b.

50. Peretti S, Judge R, Hindmarch I. Safety and tolerability considerations: tricyclic antidepressants vs. selective serotonin reuptake inhibitors. *Acta Psychiatr Scand*.

2000;101(s403):17-25. doi:10.1111/j.1600-0447.2000.tb10944.x

51. Patten SB. Progress against major depression in Canada. *Can J Psychiatry*. 2002;47(8):775-780. doi:10.1177/070674370204700810

52. Nutt DJ. Death and dependence: Current controversies over the selective serotonin reuptake inhibitors. *J Psychopharmacol*. 2003;17(4):355-364. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed6&NEWS=N&AN=2004027271.

53. Kondro W. UK bans, Health Canada warns about antidepressants. *Can Med Assoc J*. 2004;171(1):23.

54. Friedman RA. Antidepressants' Black-Box Warning - 10 Yers Later. *N Engl J Med*. 2014;371(18):1666-1668. doi:10.1056/nejmp1410540

55. Health Canada. Health Canada advises Canadians of stronger warnings for SSRIs and other newer anti-depressants - Recalls and safety alerts. Recalls and Safety Alerts. http://www.healthycanadians.gc.ca/recall-alert-rappel-avis/hc-sc/2004/13708a-eng.php. Published 2004. Accessed April 28, 2019.

56. Hemels MEH, Koren G, Einarson TR. Increased Use of Antidepressants in Canada : 1981 – 2000. *Ann Pharmacother*. 2005;36:1375-1379.

57. Welzel FD, Stein J, Hajek A, König H-H, Riedel-Heller SG. Frequent attenders in late life in primary care: a systematic review of European studies. *BMC Fam Pract*. 2017;18(104). doi:10.1186/s12875-017-0700-7

58. Roberge D, Larouche D, Pineault R, Levesque J, Hamel M, Simard B. Hospital

Emergency Departments: Substitutes for primary care? Results of a survey of the populatoin of Montreal and Monterergie. *Inst natioanl sante publique Quebec*. 2007:1-12.

59. Canadian Institute for Health Information. Sources of Potentially Avoidable Emergency Department Visits. 2014;(November).

60. Ozga A-K, Kieser M, Rauch G. A systematic comparison of recurrent event models for application to composite endpoints. *BMC Med Res Methodol*. 2018;18(2). doi:10.1186/s12874-017-0462-x

61. Johnston KM, Powell LC, Anderson IM, Szabo S, Cline S. The burden of treatment-resistant depression: A systematic review of the economic and quality of life literature. *J Affect Disord*. 2019;242:195-210. doi:10.1016/j.jad.2018.06.045

62. Souery D, Oswald P, Massat I, et al. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *J Clin Psychiatry*. 2007;68(July):1062-1070. doi:10.4088/JCP.v68n0713

63. Serretti A, Fabbri C. Factors That Predispose Patients to Treatment-Resistant Depression. *Pschiatric Times*. 2014;31(9). https://www.psychiatrictimes.com/special-reports/factors-predispose-patients-treatment-resistant-depression. Accessed May 6, 2019.

Appendix 1. ICD-9 and ICD-10 Codes Used for Inclusion, Exclusion and Washout Criteria

Inclusion and Washout criteria

From Medical Services dataset:

- 296 ; 309 ; 311

From Hospital Discharge Abstracts dataset, 1995 to March 31, 2004:

- 296.2 ; 296.3 ; 300.4 ; 309.*x* ; 311.*x*

From Hospital Discharge Abstracts dataset, April 1, 2004 to 2016:

- F32.*x ;* F33.*x ;* F34.1 ; F38.1 ; F41.2

Exclusion criteria

From Medical Services dataset:

- 295 ; 298 ; 290 ; 331 ; 797

From Hospital Discharge Abstracts dataset, 1996 to March 31, 2004:

290.0 to 290.9 ; 291.1 ; 291.2 ; 292.82 ; 292.83 ; 294.0 ; 294.1 ; 294.8 ; 294.9 ; 331.0 ; 331.1 ; 331.3 ; 331.7 ; 331.9 ; 797.*x* ; 799.3 ; 295.*x* ; 296.0 ; 296.1 ; 296.4 ; 296.5 ; 296.6 ; 296.7

From Hospital Discharge Abstracts dataset, April 1, 2004 to 2016:

F00 to F04 ; F05.1 ; F06.5 ; F06.6 ; F06.8 ; F06.9 ; F09.*x* ; F10.7 ; F11.7 ; F12.7 ; F13.7 ; F14.7 ; F15.7 ; F16.7 ; F17.7 ; F18.7 ; F19.7 ; G30.*x* ; G31.0 ; G31.1 ; G31.9 ; G32.8 ; G93.7 ; G94.*x* ; R54.*x* ; F20.*x* ; F21.*x* ; F25.*x* ; F23.2 ; F31.*x*

Appendix 2. Anatomic Therapeutic Chemical Classification Codes of Included Drugs

Antidepressants: N06A*xxx*

Antipsychotics: N05AX12 N05AX08 N05AX16 N05AH03 N05AH04 N05AN01 N05AE04

Triiodothyronine: H03AA02

Psychostimulants: N06BA02 N06BA03 N06BA04 N06BA05 N06BA07

Anticonvulsants: N06AF01 N03AG06 N03AX09 N03AX11 N03AX12 N03AB02

Appendix 3. SAS Macro Statements and Formats Used

SAS Macro Statements:

1. *%assign_iq* – Used to assign an income quartile value to a particular postal code.
2. *%_ElixhauserICD9CM* – Used to assign an Elixhauser comorbidity score to individuals in the cohort in a given year, from 1996 to March 31, 2004.
3. *%_ElixhauserICD10CM* – Used to assign an Elixhauser comorbidity score to individuals in the cohort in a given year, from April 1, 2004 to 2016.
4. *%_epihosp* – Used to ensure that multiple hospitalization observations are distinct observations. If these observations were instead part of a single hospitalization episode, this macro amalgamated the observations.

SAS Formats:

sefi96f. to *sefi99f.* and *sefi00f* to *sefi16f* – Used to assign a SEFI value to a particular postal code.

Appendix 4. AIC Values for All Models

Ambulatory Visits, non-mental health reason

Model	AIC
Without Covariates	303,220,236
MGH	303,075,073
MGH + AGE	302,317,693
MGH + AGE + SEX	302,268,798
MGH + AGE + SEX + SEFI	301,751,539
MGH + AGE + SEX + SEFI + URBAN	301,706,930
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	301,691,668
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX	301,566,531
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV	301,554,468
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEX	301,553,341
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEFI	301,553,115
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*PYPV	301,550,897
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*AGE	301,545,381

Ambulatory Visits, mental health reason

Model	AIC
Without Covariates	57,249,716
MGH	56,537,751
MGH + AGE	56,483,454
MGH + AGE + SEX	56,483,237
MGH + AGE + SEX + SEFI	56,365,097
MGH + AGE + SEX + SEFI + URBAN	56,288,378
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	56,280,593
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX	56,280,590
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + MGH*SEX	56,280,100
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + MGH*SEFI	56,279,506
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + MGH*PYPV	56,279,834

Emergency Department Visits, non-mental health reason

Model	AIC
Without Covariates	5,454,809
MGH	5,446,434
MGH + AGE	5,446,283
MGH + AGE + SEX	5,446,075
MGH + AGE + SEX + SEFI	5,417,032
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	5,342,087
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV	5,341,530
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*SEFI	5,341,415
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*SEX	5,341,395
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*PYPV	5,341,366
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*AGE	5,340,896

Emergency Department Visits, mental health reason

Model	AIC
Without Covariates	202,191
MGH	199,602
MGH + AGE	197,364
MGH + AGE + SEX	197,076
MGH + AGE + SEX + SEFI	195,611
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	192,779
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX	192,781
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV	191,587
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*AGE	191,580
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*PYPV	191,578
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*SEFI	191,572
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + PYPV + MGH*SEX	191,565

Hospitalizations, non-mental health reason

Model	AIC
Without Covariates	8,632,291
MGH	8,627,438
MGH + AGE	8,563,109
MGH + AGE + SEX	8,562,809
MGH + AGE + SEX + SEFI	8,540,936
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	8,530,052
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX	8,514,041
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV	8,513,713
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*PYPV	8,513,672
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEFI	8,513,660
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEX	8,513,652
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*AGE	8,513,367

Hospitalizaiton, mental health reason

Model	AIC
Without Covariates	152,624
MGH	147,975
MGH + AGE	147,953
MGH + AGE + SEX	147,955
MGH + AGE + SEX + SEFI	147,389
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	146,680
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX	146,225
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV	144,490
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEFI	144,476
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEX	144,474
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*AGE	144,467
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*PYPV	144,423

Mortality

Model	AIC
Without Covariates	332,043
MGH	331,292
MGH + AGE	302,343
MGH + AGE + SEX	301,359
MGH + AGE + SEX + SEFI	299,444
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY	298,768
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX	296,944
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV	296,740
Interactions	**AIC**
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEX	296,735
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*SEFI	296,733
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*PYPV	296,715
MGH + AGE + SEX + SEFI + URBAN + YEAR_ENTRY + ELIX + PYPV + MGH*AGE	296,675