Predicting Resource Use of Community Mental Health Services at the Transition from Inpatient Psychiatry

by

Nam Tran

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Public Health and Health Systems

Waterloo, Ontario, Canada, 2020

Examination Committee Membership

The following served on the Examination Committee for this thesis. The decision of the Examination Committee is by majority vote.

External Examiner: Elizabeth Lin
Associate Professor, Department of Psychiatry,
University of Toronto

Internal-External Member: Joel Dubin
Associate Professor, Department of Statistics and Actuarial Science,
& School of Public Health and Health Systems,
University of Waterloo

Supervisor: John Hirdes
Professor, School of Public Health and Health Systems,
University of Waterloo

Internal Member: Christopher Perlman
Associate Professor, School of Public Health and Health Systems,
University of Waterloo

Internal Member: Jeff Poss
Associate Adjunct Professor, School of Public Health and Health Systems,
University of Waterloo

Internal Member: Brant Fries
Adjunct Professor, School of Public Health and Health Systems,
University of Waterloo
Professor Emeritus, Department of Health Management and Policy,
University of Michigan

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis consists in part of one manuscript that have been submitted or published. Exceptions to sole authorship:

Chapter 2: **Tran N**, Poss JW, Perlman C, Hirdes JP. Case-Mix Classification for Mental Health Care in Community Settings: A Scoping Review. Health services insights. 2019 Jul;12:1178632919862248.

As the lead author, I was responsible for conceptualizing study design, carrying out data collection, analysis, drafting, and submitting the manuscript. My co-authors provided guidance during each step of the research and provided feedback on draft manuscripts.

Abstract

Mental health is a major health problem for many Canadians. Methods to predict expected mental health care resource use are an essential component in balancing the needs of the population and equitable allocation of limited health care resources. This research examined the relationship between the resource use of community mental health services and the characteristics of their clients using a case-mix classification approach.

A scoping review showed that most of the research on this topic focused on inpatient psychiatry settings. The number of identified studies (n=17) and case-mix systems (n=32) reflected the modest level of research activity in this area.

Secondary analyses were done with a sample of adults discharged from a local psychiatric hospital unit in Ontario (n=4,688 discharges) that was tracked to examine the use of community mental health services after discharge. Only about half of the discharges subsequently used publicly funded community mental health services. Further, only n=1,207 discharges had services initiated within 30 days and were not censored by readmission. Clinical characteristics measured at discharge from inpatient psychiatry were associated with observed use and high use (as binary variables) of community mental health services post-discharge. Usage of services specially designed for persons at risk of self-harm and harm to others (as binary variables) were also associated with higher risk of self-harm and harm to others measured at discharge.

A community episode of 90 days from first contact with the community mental health agency post-discharge appeared to be the most practical for implementation. Two high performing case-mix classification systems were examined for their possible predictive utility for post-discharge community mental health service use. The System for Classification of In-Patient Psychiatry (SCIPP) achieved 6% explained variance of community resource use for an episode. When prior contact with the community mental health agency within 30 days prior to the inpatient episode was included, the model with SCIPP explained up to 14.1% of variance in resource use. The Australian Mental Health Classification (AMHCC) was found to be not immediately applicable outside of the Australian context, and most of its explained variance was likely attributed to the "phases of care" that are subjectively determined by clinicians at the beginning of an episode. The remaining components of the AMHCC explained only 1.2% of variance in resource use.

Using machine learning, new classification models using discharge clinical characteristics achieved up to 12% of explained variance in cross-validation. The two simplest decision tree models showed similar performance in cross-validation as more complex models. Although machine learning identified relevant relationships between clinical characteristics and observed resource use, some relationships required human expertise to adjust to align with the goals of the health care system. This was exemplified by a manual decision tree model that achieved 11.1% explained variance on the development data set. These results pointed to the need for additional research to: expand the sample size; include a broader range of community mental health service users; use more contemporaneous clinical assessment data measured at community service initiation; and broaden the participation of community mental health agencies. Although clinical characteristics measured at discharge yielded only modest predictive utility, designing a system that could leverage both inpatient information and community agency assessment information could improve both predictive utility and care integration across the care continuum. Further development of case-mix classification for community mental health will require a broad collaboration across the health care system.

Acknowledgments

First, I would like to thank my supervisor, Dr. Hirdes, who renewed my commitment in finishing the Ph.D. when I did not think it was possible. From our research group, I have also met many wonderful and supportive friends, whose friendships I will cherish forever.

Many people working in mental health care across the Waterloo region came together to make this research possible. Most notably are Ben Windling, Dr. Tina Mah, and Sarah Laferriere from the Grand River Hospital; Anna Tersigni, Lana Monakhova, Karim Fathi, Meredith Gardiner from the Canadian Mental Health Association - Waterloo Wellington.

Many members of the my research group also provided assistance, guidance, and wisdom that allowed me to complete this research. I want to thank Micaela Jantzi, Julie Koreck, Jonathen Chen, Sophie Hogeveen, Joanna Sinn, as well as Trevor Bain and Brent Clerk from the information technology department.

I would also want to thank Drs. Perlman, Poss, Dubin, Fries, and Lin for serving on my thesis committee and providing valuable guidance for my research.

Lastly, my doctoral study was supported by the Ontario Trillium Scholarship by the Ontario Ministry of Advanced Education and Skills Development and the University of Waterloo.

Table of Contents

List of Figures

List of Tables

Abbreviations

Acronym	**Description**
ABF	Activity-Based Funding
ABS	Aggressive Behavior Scale
ACG	Ambulatory Care Groups
ACT	Assertive Community Treatment
ADG	Ambulatory Diagnosis Groups
ADL	Activities of Daily Living
ADRG	Alternative Diagnostic-Related Groups
AUROC	Area Under the Receiver Operating Curve
CAGE	Cut down, Anger, Guilt, Eye-opener
CART	Classification and Regression Tree
CIHI	Canadian Institute for Health Information
CMG	Case Mix Groups
CMI	Case-mix Index
CMHA-WW	Canadian Mental Health Association - Waterloo Wellington
CPS	Cognitive Performance Scale
CPSI	Computerized Psychiatric Severity Index
CRG	Clinically Related Groups
DALY	Disability-adjusted Life Years
DCC	Diagnostic Cost Groups
DRG	Diagnostic-Related Groups
DRS	Depression Rating Scale
DSI	Depressive Severity Scale
DSM	Diagnostic and Statistical Manual of Mental Disorders
DSM-III	Diagnostic and Statistical Manual of Mental Disorders III
GAF	Global Assessment of Functioning

GRH	Grand River Hospital
HCC	Hierarchical Condition Categories
HoNOS	Health of the Nation Outcome Scales
IADL	Instrumental Activities of Daily Living
ICD	International Classification of Diseases and Related Health Problems
IMPACT	Integrated Mobile Police and Crisis Team
interRAI CMH	Resident Assessment Instrument - Community Mental Health
LHIN	Local Health Integration Network
LPPC	Long-Stay Psychiatric Patient Classification
LSP-16	Life Skills Profile-16
MDC	Major Diagnosis Categories
MH-CASC	Mental Health Classification and Service Costs
mhcCTRLINT	Control Intervention Clinical Assessment Protocol
mhcCRIM	Criminal Activity Clinical Assessment Protocol
mhcEDEMP	Education and Employment Clinical Assessment Protocol
mhcEXER	Exercise Clinical Assessment Protocol
mhcFALLS	Falls Clinical Assessment Protocol
mhcFINAN	Personal Finance Clinical Assessment Protocol
mhcHARMOTH	Harm to Others Clinical Assessment Protocol
mhcINFSUPP	Informal Support Clinical Assessment Protocol
mhcIPCON	Interpersonal Conflict Clinical Assessment Protocol
mhcMEDMGT	Medication Management and Adherence Clinical Assessment Protocol
mhcPAIN	Pain Clinical Assessment Protocol
mhcREHOSP	Rehospitalization Clinical Assessment Protocol
mhcSELFCR	Self-Care Clinical Assessment Protocol
mhcSELFHARM	Self-Harm Clinical Assessment Protocol
mhcSLEEP	Sleep Disturbance Clinical Assessment Protocol
mhcSOREL	Social Relationships Clinical Assessment Protocol
mhcSSDIS	Support System for Discharge Clinical Assessment Protocol
mhcSUBUSE	Substance Use Clinical Assessment Protocol
mhcTOBUSE	Smoking Clinical Assessment Protocol
mhcTRAUMA	Traumatic Life Events Clinical Assessment Protocol
mhcWTMGT	Weight Management Clinical Assessment Protocol
MHSU	Mental Health and Substance Use

OASIS	Outcome and Assessment Information Set
OHT	Ontario Health Team
OMHRS	Ontario Mental Health Reporting System
PDG	Psychiatric Diagnosis Groupings
PSI	Psychiatric Severity of Illness Index
PSSL	Positive Symptoms Scale Long-form
PSSS	Positive Symptoms Scale Short-form
PPC	Psychiatric Patient Classes
RAI-MH	Resident Assessment Instrument - Mental Health
RHO	Risk of Harm to Others
RUG-III	Resource Utilization Group Version III
SCI	Self-care Index
SCIPP	Systems for Classification of In-Patient Psychiatry
SOS	Severity of Self-harm
SWS	Social Withdrawal Scale
U.K.	United Kingdom
U.S.	United States
USD	United States Dollars
WHO	World Health Organization
xgboost	Extreme Gradient Boosted Trees
YLDs	Years Lived with Disability

Chapter 1
Introduction

Globally, mental health illness could account for as much as 32% of years lived with disability (YLDs) and 13% of disability-adjusted life-years (DALYs) [1], as well as direct costs (such as treatment, prevention, rehabilitation) and indirect costs (such as lost wages and lost productivity) of about USD 2.5 trillion in 2010 [2]. It is estimated that one in five people in Canada lives with mental health illness or substance use problems [3]. The total economic burden, which includes health care costs and indirect costs attributable of mental health and addiction, in Canada is estimated to be about $50 billion per year in 2011, rising to about $88 billion by 2021 [4]. This places mental health and substance use among the costliest health problems in Canada [3, 4], in comparison to the economic burden of cardiovascular diseases of about $12 billion and cancer of about $5 billion in 2008 [5].

1.1 Mental Health

The World Health Organization defined mental health as "a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community" [6].

This definition is a major step forward in recognizing that mental health well-being is not simply the absence of mental health illness [7]. Additionally, it also recognizes the importance of an individual's environment and community in one's mental health.

Although the absence or presence of mental health illness alone does not define an individual overall mental health, defining the mental health diagnostic categories is essential is differentiating the type of case and potential care plans. Currently, there are two main mental health diagnostic classification systems, the International Classification of Diseases and Related Health Problems (ICD) and the Diagnostic and Statistical Manual of Mental Disorders (DSM). The ICD included diagnoses for all of medicine, and DSM focuses only on mental health. These two classifications were developed in coordination with each other and shared a compatible coding scheme [8]. In addition to being used as part of the delivery of care, the diagnostic codes can also be used for the purpose of financing in a health care system.

1.2 Mental Health Care

Specialty psychiatric hospitals used to be the primary way of delivering mental health care [9]. In the past few decades, in Canada and in many countries, mental health care has undergone a transformation that shifted care from primarily inpatient settings within a hospital towards care that is delivered in community settings, often referred to as deinstitutionalization [9–11]. Although the number of specialty psychiatric hospitals have decreased, more psychiatric units were opened within general hospitals as well as residential care facilities, which contributed to the decrease in stigma and increase in care coordination across clinical professions [2, 8, 12].

Inpatient psychiatry is still an important component of the health care system for crisis and intensive care [8].

Community mental health care is described by the World Health Organization as a secondary care provided in a community settings that often do not require overnight stay, such as at home, outpatient clinic, or office, that are provisioned to assess and treat mental health illness by mental health professionals but not by primary care physicians [13]. While the exact services offered and their availability vary widely across jurisdictions, community mental health services generally include collaborative care models that bring together many professionals, peer support, and residential care [8]. Community mental health services can range from intensive or urgent care (such as assertive community treatment) to less intensive care (such as peer support group or therapies)

The trend in deinstitutionalization is expected to continue. In addition to the reduction in stigma of seeking mental health services, the volume and demand of community mental health services, therefore, are expected to also increase in the future. In 2012, the Mental Health Commission of Canada recommended an increase of two percent in funding for mental health services and social programs, in addition to increasing the proportion of health spending devoted to mental health to nine percent by 2020 [14]. While the amount of the total funding is always crucial, an equally important line of inquiry is how community mental health services are financed and whether the health system has the tools to make appropriate funding allocation decisions [2, 15].

As of 2019, mental health services for both inpatient and community settings in Ontario are organized and funded by the local health integration networks (LHINs) [16]. The LHINs typically enter an agreement with the providers to provide a set of services for residents of a geographical area in return for funding on an annual basis [17]. Funding for children mental health services may also come directly from the provincial Ministry of Children, Community and Social Services [16]. Likewise, targeted intervention programs may also be funded by local or provincial governments [16].

1.3 Health Care Funding

There are many methods to pay health care providers for the health care services they provide. This section outlined the most common methods, their strengths, and weaknesses. It is worth noting that a funding formula may use more than one funding methodologies, with each method contributing a percentage of the total funding, or different methodologies are used at different levels of the health care system (such as one method is used for payment to regional health authority, and another for payment to individual physicians).

1.3.1 Fee-for-Service

The simplest method of reimbursing is to pay the provider for each service [18, 19]. Typically, the payer, either a public organization or an insurer, has a pre-determined price list for all possible service items that they cover. In Ontario, most physicians are reimbursed under the fee-for-service schedule established by the provincial government [20].

Fee-for-service has been blamed for driving up the health care costs and inefficiency because it incentivizes volume regardless of clinical appropriateness and discourages care coordination [21]. There is a growing number of health professionals who advocate for the elimination of fee-for-service or, at a minimum, condition the fee-for-service payments on quality benchmarks to increase accountability [21].

1.3.2 Population-Based Funding

This method, also known as capitation funding, funds the provider, or the LHINs in the case of Ontario, proportionally based on the characteristics and size of the population. The underlying assumption is that populations of similar size or prevalence of diseases should have similar health expenditures [19].

This method allows the regional health authority to have more autonomy and flexibility in allocating resources in their own jurisdictions [19]. However, a population-based funding formula constructed using historical data may be influenced by biases, such as inappropriate historical usage of services or barriers to access services. If not actively corrected for, using observed historical data for allocation may continue to reinforce inequities. Although the characteristics and prevalence of diseases of a population may change over time, these changes may be gradual and need to be continuously monitored in order to ensure that the funding responses to the clinical characteristics of the population.

1.3.3 Global Budget

Global budget has traditionally been used to fund hospitals or regional health authorities by making an once-a-year payment for all services that occurred within the fiscal year [19]. The advantage of global budget is its predictability during the fiscal year. On the other hand, it could require lengthy negotiations every year if the providers wish to increase their funding. The disadvantage of global budget is that the funding does not response to clinical characteristics of the patients in a timely manner, because funding is usually fixed at the start of a fiscal year. Additionally, since the providers only have control over the cost side of the budget, there is potential for restricting access due to limited resources in an effort to lower expenditures, which could result in long wait time or lowered accessibility [19].

1.3.4 Case-Mix Funding

Case-mix funding is also often referred to as activity-based funding, which is a broader concept that funds the providers based on their activities. A case-mix funding scheme has two main components: describing the activities using a classification system and pricing for the activities [18, 22]. The case-mix classification system serves as a link between the clinical characteristics driving the need of health care of an individual to the expected resources required to provide care [18]. The underlying assumption is that individuals with similar clinical severity or complexity should consume similar amount of health care resources [23].

The providers are reimbursed according to the case-mix classification of the clinical needs. Because the providers are able to keep the surplus and responsible for the loss, there is an incentive to be cost-efficient in providing care. The pricing component also needs to be responsive to changes in non-clinical factors within the health care systems, such as: inflation, changes in practice, or observed attempts to game the system [22]. Therefore, case-mix funding requires a robust pricing component to ensure that equitable allocations across the health care system and over time.

1.3.5 Bundled Payment

Bundled payment reimburses the providers based on a defined “bundle” of care that covers all aspects of care of a person during a defined period of time regardless of care settings [19]. A bundle is typically well-defined and standardized across a wide-range of patients. For example, a knee replacement surgery bundle can cover pre-surgery therapies, surgery, and post-surgery rehabilitation to restore function [19]. The rationale is that bundled health care services are similar to other products and services purchased by consumers, in which consumers can make a single payment to get what they need instead of sourcing individual components from many different suppliers.

This method also incentivizes providers to control costs of the services because they are responsible for the profit or loss. Since the bundle may be designed to be indifferent to care settings, it can also incentivize care coordination, and reducing poor outcomes, such as rehospitalization [19]. On the other hand, bundled payment may not be suitable for every of health care case. For example, the outcomes of mental health care can depend on many factors beyond the interaction between providers and patient, such as: environmental triggers, traumas, social relations, employment, and housing [8].

1.3.6 Pay-for-Performance

Pay-for-performance is also known as value-based payment, or pay-for-quality, which attaches funding to a performance measure [19]. There are many possible performance measures that can be used. Some common measures include: unplanned readmission, hospital acquired conditions, surgical site infections [19]. Payment-for-performance is also susceptible to gaming if the criteria are not designed well. For example, a study found that a program designed to increase outreach to population with severely mental illness only increased the documentation of such population but not their treatments [24].

1.4 Funding Reform in Ontario

In 2002, a Senate report on the future of the Canadian health care system emphasized the need for a more equitable method of funding than global budget [25]. In 2012, the province of Ontario reformed the way health care services are funded by proposing a new funding formula [26]. The intention is to transition away from entirely global budget towards a funding formula that is a mix of global budget and case-mix [26]. The case-mix portion will use different classification systems appropriate for each health sectors to adjust for the reimbursement. Additionally, for a subsets of well-defined procedures, providers are reimbursed based on a bundled price per procedure [26]. The proportions are adjusted over time to reduce the share of global budget relative to the other two components. The transition was planned in phases, with mental health sector in the later phases and yet to be implemented. As of 2019, inpatient psychiatry and community mental health are funded based on global budget by the LHINs.

To turn the funding reform proposal into reality, case-mix classification systems are required to support case-mix funding. While significant efforts was devoted to the development of a case-mix classification system for the inpatient psychiatry setting, the Systems for Classification of In-Patient Psychiatry (SCIPP) [27], less attention has been paid to a case-mix classification system for the community mental health settings.

1.5 Overview of the Thesis

This research sought to understand the delivery of community mental health services, and the relationship between individual-level clinical characteristics and resource use. This research is expected to provide initial evidence to guide resource allocation for community mental health services using case-mix funding, in Ontario and beyond.

This research started with a scoping review of existing case-mix classification systems applicable to community mental health services in **Chapter 2**. The rest of this research studied a sample of adults who used public-funded services at one of the largest community mental health agencies in Canada after they were discharged from inpatient psychiatry. The overall strategy was to study the relationship between an individual's clinical characteristics (indicated by their discharge assessments from inpatient psychiatry) and their usage of community mental health services post-discharge. Although there is a lack of high quality standardized clinical data in the community mental health settings, this study was made possible by combining standardized clinical data from hospitals at discharge with individual-level resource use data from a community mental health agency. Clients of one of the largest community mental health agencies should be characterized by a broad range of variation that is evident in the population. Lastly, since the clinical data and resource use data were produced by two independent organizations, there is a low risk of inflating the measurements for financial gain.

Chapter 3 described the pattern of community mental health service usage post-discharge and examined the association between the clinical profile measured at discharge and subsequent usage of community mental health. **Chapter 4** examined whether the two existing case-mix classification systems - System for Classification of In-Patient Psychiatry (SCIPP) and Australian Mental Health Case-mix Classification (AMHCC) - can predict community mental health service resource use beyond the context that they were developed for. Lastly, **Chapter 5** leveraged machine learning techniques to build experimental case-mix classification systems using the clinical data measured at discharge from inpatient psychiatry and observed community mental health resource use.

Chapter 2
Case-Mix Classification for Mental Health Care in Community Settings: A Scoping Review

A similar version of this chapter was originally published as:

Tran N, Poss JW, Perlman C, Hirdes JP. Case-Mix Classification for Mental Health Care in Community Settings: A Scoping Review. Health services insights. 2019 Jul;12:1178632919862248.

Abstract

A scoping review was conducted to summarize the nature, extent, and range of research on case-mix classifications used to predict mental health care resource use in community settings. This study identified 17 eligible studies with 32 case-mix classification systems published since the 1980s. Most of these studies came from the USA Veterans Affairs and Medicare systems, and the most recent studies came from Australia. There were a wide variety of choices of input variables and measures of resource use. However, much of the variance in observed resource use was not accounted for by these case-mix systems. The research activity specific to case-mix classification for community mental health care was modest. More consideration should be given to the appropriateness of the input variables, resource use measure, and evaluation of predictive performance. Future research should take advantage of testing case-mix systems developed in other settings for community mental health care settings, if possible.

2.1 Introduction

Although each person in the population is unique, there are shared characteristics that determine the types of treatments or services that individuals receive from the health care system [28]. Recognition of this point led to the idea that there are existing groups of people with similar characteristics that will consume similar amount health care resources and, by extension, incur similar costs of care. These groups represent the mix of cases that are observed in a health care system, or a “case-mix” [23], which can be viewed as a proxy for the types of health care needs of the population.

Case-mix classification systems can be of two types: grouping or index systems [23]. Grouping systems assign cases into relatively homogeneous groups in terms of their expected resource use [23]. Each group has a weight associated to represent its expected resource use relative to the average case in the population, also known as “case-mix index” (CMI) [23]. For example, the Resource Utilization Group Version III (RUG-III) is commonly used in the USA and Canada for nursing homes reimbursement [29–32]. Index systems, instead, combine different characteristics of a case to produce a numerical value for each case that represents the expected level of resource use, then map it to a case-mix index value [23]. An example of such a system is the Outcome and Assessment Information Set (OASIS) used by Medicare to reimburse home care services [33].

Case-mix classification systems are primarily used to reimburse health care providers based on the type of patient [30]. It is worth noting that a funding formula is distinct from a case-mix classification system. A funding formula may work by assigning a monetary amount to the case-mix index, also known as tariff, and further adjusted based on numerous factors such as: available funding, inflation, geographic and provider characteristics, or negotiations between health system administrators and the providers. On the other hand, the case-mix index values are expected to remain constant because the health care needs of one group relative to another should not change drastically from year to year [18]. Case-mix index values can change in rare occasions, such as changes in technologies or clinical practices, that can make a group much more or less expensive to care for compared to others. Other applications of case-mix classifications include: risk-adjustment models for health outcomes or other quality measures, and long-term planning and budgeting tools for policy makers [34].

For mental health, the delivery of care can take place in multiple settings as mental health care has shifted from facility-based inpatient care to community-based care, as a result of deinstitutionalization initiatives [10, 11]. Facility-based inpatient care provides intensive observation, diagnosis, and treatment typically in times of crisis [8], and usually requires a hospital admission with one or more overnight stays [35]. Community-based care typically employs a care team that provides a wider range of services, including both urgent and ongoing care, such as: assertive treatment services, crisis management, outreach, recovery, housing, occupation training, and day programs [8].

Previously, Jones et al. reviewed 16 studies between 1990 and 2005 studying predictors of mental health service utilization and costs [36]. Hermann et al. reviewed 36 studies between 1980 and 2002 focusing on risk adjustment models of psychiatric health outcomes and costs that included some case-mix systems [37]. Mason and Goddard reviewed only 5 international examples of activity-based funding systems for mental health between 2006 and 2008 [38]. Harris et al. reviewed 13 case-mix classification systems for all care settings but only in some Western countries published between 1995 and 2012 [39].

However, to date, most mental health case-mix classification systems have predominantly focused on care in acute or inpatient settings. Given the de-institutionalization shift, it is necessary to examine case-mix classification systems for the community settings. Therefore, this review aimed to summarize the nature, extent, and range of the up-to-date research on mental health care resource prediction using case-mix specifically in community settings, and identify the gaps in the current research.

2.2 Methods

In alignment with scoping review methods by Arksey and O'Malley [40], and PRISMA [41], four academic literature databases were searched: PubMed, Web of Science, PsychInfo, and SCOPUS. Keywords were used to search the title and abstract for the presence of mental health, case-mix, and community settings concepts: ("mental health" OR "mental ill*" OR "mental disorder?" OR psychiatr* OR "behavio* care" OR "behavio* health") AND ("casemix" OR "case mix" OR "case-mix" OR "case type?" OR "diagnosis related group*" OR "patient mix" OR "patient? group*" OR "patient? classification?" OR "patient? cluster*" OR "case? cluster*" OR "risk adjust*" OR "case adjust*") AND ("communit*" OR "outpatient?" OR "out-patient?" OR "ambulatory"). Searches were done in October 2018 and included all date ranges. Duplicates and non-English full-text articles were removed. Database searches were also supplemented by scanning references of the eligible articles, consulting with experts and committee members.

Articles' titles and abstracts were then screened for relevance, followed by a screen of the full-text. Articles were included if a case-mix classification system was used to predict resource use of community mental health care or health care resource use of people with mental health disorders in community settings. This review used the World Health Organization's definition of health care resources as the three main inputs of a health care systems as: human resources, physical capital, or consumable resources [42]. As in similar reviews [39, 43, 44], this review considered studies that predict resource use using case-mix classification, rather than to simply describe the differences in resource use among sub-groups of the study sample, or to explain the variation in resource use by adjusting for different variables. Additionally, a predictive study should provide a quantitative assessment of how well the predicted resource use explains the observed resource consumption, such as the R^2 value [36]. The community settings were defined as care settings that do not require an overnight stay at the facility [35], which may include outpatient treatments or day programs.

To capture the scope of the case-mix classification systems presented, we collected some main characteristics from each eligible article. Specifically, we collected information regarding the bibliography (authors, year of publication), sample data (geographic jurisdiction, care settings, age groups, sample size), case-mix system (name, input variables, type), resource use measure (definition of measure), and predictive performance (type, reported value). Data were then recorded and reviewed with the committee members.

2.3 Results

This study identified 17 articles matching the criteria (Figure 2.1), which presented 33 case-mix classification models (Table 2.1). Most were from academic sources, except for the technical reports of the case-mix systems developed in Australia and New Zealand [35, 45, 46]. Most studies (11 out of 17) focused only on adult population.

Most of the research came from the USA, and the largest studies came from the USA

Veterans Affairs and Medicare systems [47–50]. However, it is worth noting that the samples from the Veteran Affairs system were mostly adult males, and samples from the Medicare system were adults aged 65 or older, which are not representative of the US population. The most recent major effort came from Australia with their Australian Mental Health Care Classification (AMHCC) [46], which was developed to predict resource use for both inpatient and community settings and all age groups.

The input variables for the case-mix classification systems were varied. Most common variables were: diagnosis, demographics, variables related to severity, comorbidity, or functional status. The majority of the case-mix systems were grouping systems, and index systems were less common.

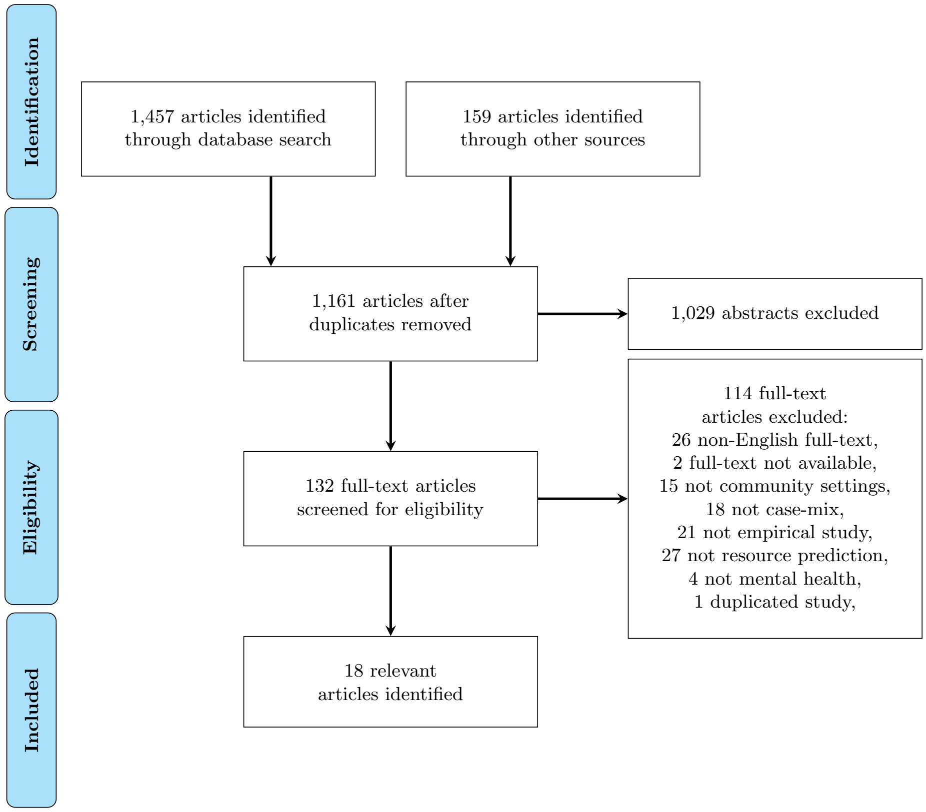

Figure 2.1 Search Procedures for Relevant Articles

There were also a wide range of measures of resource use from the studies identified. These measures can be roughly classified into two types: proxy measures (such as number of visits) (Table 2.2), or direct measures (such as claims data or wage-weighted staff time) (Table 2.3).

For the direct measures, all studies used episodic basis for their resource use measures, which summed all the relevant costs over an episode of care. Only two studies attempted to define episodes of care that were variable based on the group or case [45, 46], while the others pre-defined a fixed episode length for the entire sample. There were also a wide range of follow-up times used for measuring resource use (Table 2.2-2.3), ranging from a few weeks to up to three years. Alternatively, another option is to calculate a direct resource use measure on a per-diem basis, which predicts resource use per day or per visit [23], such as the System for Classification of In-Patient Psychiatry (SCIPP) developed in Canada [27, 51, 52].

The measures of resource use could also be expressed as a continuous variable or a categorical variable. As a result, there were also various performance metrics used to evaluate the case-mix classification systems, but most common was the coefficient of determination (R^2) for the measures of resource use expressed as a continuous variable (Table 2.2-2.3). The R^2 was sometimes referred to as the reduction in variance (RIV), or the amount of variance in resource use explained by the case-mix classification system. Although the R^2 was commonly reported, the differences in the measures of resource use and follow-up duration did not allow for a meaningful comparison.

Since the distribution of the resource use was often positively skewed, some studies attempted to approximate a symmetric distribution with a log transformation [53, 54] (Table 2.3). Some studies also trimmed the outliers to improve their predictive performance [35, 45] (Table 2.3).

There were also other notable case-mix classification systems currently being used, where activity-based funding has been implemented, such as: the Netherland's Zorgzwaartepakketten (ZZP) and the UK's Mental Health Clustering Tool (MHCT) but, to our knowledge, these did not have empirical results regarding their predictive performance. The ZZP has 38 psychosocial care packages which classifies all ages based on psychosocial or cognitive functioning, social skills, mobility, activities of daily living, and behavioral problems [38]. The MHCT has 21 groups which used the Health of the Nation Outcome Scales (HoNOS) [55] as input, then classifies adults using diagnosis, severity, chronicity, and cognitive impairment [56]. An earlier version of the MHCT with 13 groups reported an $R^2 = 10.9$ % [57].

Table 2.1 Eligible studies, ordered by year of publication

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
Wood and Beardmore, 1986 [58]	USA, adult outpatient service at an university affiliated mental hospital	1,000 adults	Diagnostic Related Groups (DRGs): 8 mental health and substance abuse DRGs [59]	Grouping
Wittman and Lerner, 1990 [60]	Israel, mentally ill outpatients	2,118 out-patients, age: 15-65	Chronicity: 6 terminal groups classified by long-term service, age, disability, diagnosis, and prior hospitalizations	Grouping
Barker et al., 1994 [61]	USA, Oregon's local community mental health agencies	240 adults	Multnomah Community Ability Scale (MCAS). 4 domains: interference with functioning, adjustment to living, social competence, behavioral problems	Index

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
Uehara et al., 1994 [62]	USA, Washington's Community Psychiatric Clinic	598 adults	Level of Need-Care Assessment (LONCA): Clients were assessed for 10 key needs, each has 4 levels (none, low, moderate, intense). These needs were then grouped according to physical, psychological, and social functioning	Grouping
Ettner et al., 1997 [63]	USA, New Hampshire Medicaid enrollees	12,218 adults, 17,901 children	Ambulatory Care Groups (ACGs): 51 mutually exclusive (ACGs) [64] based on ICD-9 codes, age, gender, and intermediate Ambulatory Diagnostic Groups (ADGs) of similar expected resource consumption	Grouping
Ettner et al., 1998 [65]	USA, claim records from a private insurer provided plans for employer-sponsored health insurance.	51,621 adults, 14,145 children	Demographics	Grouping
			Demographics and ACG	Grouping
			Demographics and ADG	Grouping

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
			Demographics and HCC	Grouping
			Demographics, diagnosis and comorbidity	Grouping
Trauer et al., 1998 [53]	Australia, Melbourne public psychiatric service registration list	200 adults	Diagnosis (schizophrenia, personality disorder, and social withdrawal)	Grouping
			Life Skills Profile (LSP) functional assessment [66, 67], which contained 5 sub-scales: antisocial, bizarre, compliance, withdrawal, and self-care	Index
			Diagnosis and Life Skills Profile (LSP)	Mixed
Samuels, 1996 [54]	USA, New York's licensed mental health service providers	24,463 adults	High/Medium/Low user groups based on historical usage, diagnosis and insurance type	Grouping
			High/Medium/Low users groups based on historical usage, insurance type, diagnosis, and age	Grouping
			High/Medium/Low user groups based on historical usage, insurance type	Grouping

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
Buckingham et al., 1998 [35]	Australia, 22 sites (inpatient and outpatient)	Adults: 9,806 episodes (outliers trimmed: 9,096), Children/adolescents: 2,098 episodes (outliers trimmed: 1,956)	Mental Health Classification and Service Costs (MH-CASC): 19 community terminal groups (adults: 10, children/adolescents: 9), out of 42 groups for all settings. Adult variables: focus of care, legal status, HoNOS assessment [55], and LSP-16 [68]. Children/adolescents variables: age, HoNOSCA assessment [69], CGAS assessment [70], and FIHS assessment [71].	Grouping
	Australia, integrated mental health care sites	8,067 adult episodes (outliers trimmed: 7,244)	Experimental Bundled Episodes: 12 terminal groups. Variables: legal status, HoNOS assessment [55], diagnosis, suicidal risk, psychotic symptoms, and age.	Grouping
Leslie et al., 2000 [48]	USA, Veterans Affairs mental health outpatient clinics	53,700 adult patients	Global Assessment of Functioning [72]	Index

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
			Service-connected status: assessment of disability linked to military service	Index
			Service-connected status, but if patients were not service-connected, use GAF	Index
			Diagnosis: 12 groups (alcoholism, bipolar, dysthymia, generalized anxiety, major depressive, organic brain syndrome, other substance abuse disorder, panic disorder, personality disorder, post-traumatic stress disorder, schizophrenia, and other)	Grouping
DeLiberty et al., 2001 [73]	USA, Indiana Division of Mental Health	60,000 adults and children/adolescents	Serious Mental Illness (SMI): 9 groups. Level 1: by diagnoses. Level 2: by levels of difficulties	Grouping

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
Rosen et al., 2002 [49]	USA, Veteran Affairs inpatients and outpatients	1,039,712 adult patients (66.6% development, 33.3% validation)	Diagnostic Cost Group/Hierarchical Condition Category (DCG/HCC): ICD-9CM maps to 37 diagnostic groups, then aggregate into conditions categories (which a person can have multiple). Five hierarchies of conditions were then imposed so that minor diagnoses do not add to cost prediction.	Grouping
Gaines et al., 2003 [45]	New Zealand, 8 district health boards	Adults: 9,199, children/youths: 2,868	New Zealand Mental Health Classification and Outcomes Study (NZ-CAOS): 22 community terminal groups, out of 42 groups for all care settings. Adults (13 groups): assessment only, legal status, ethnicity, focus of care, and age. Children/Youths (9 groups): assessment only, ethnicity, age, HoNOSCA assessment [69].	Grouping
			MH-CASC [35]	Grouping

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
Selim et al., 2004 [74]	USA, Veterans Affairs ambulatory care at 4 sites in Boston.	2,425 adults	Physical and Mental Comorbidity Indices (PCI/MCI): Count of 36 physical diagnoses and 6 mental diagnoses	Index
			Conditional and Mental Comorbidity Indices (CCI/MCI): Count of 36 physical diagnoses (with symptoms) and 6 mental diagnoses	Index
Sloan et al., 2006 [50]	USA, Veterans Affairs inpatients and outpatients	914,225 adult patients (60% development, 40% validation)	PsyCMS: 46 categories based on ICD-9CM codes, with 4 hierarchies (alcohol use, drug use, anxiety disorder, and mood/psychotic disorder) imposed to assign patients into the highest expected cost category in a given hierarchy	Grouping
			Age (9 groups) and gender	Grouping
			VA-MH12: 12 categories of mental health diagnosis based on ICD-9CM codes	Grouping
			Adjusted Clinical Group/Aggregate Diagnostic Group (ACG/ADG)	Grouping

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
			DCG/HCC: 2 hierarchies (substance abuse and psychiatric disorders)	Grouping
			Chronic Illness and Disability Payment System (CDPS) [75]: 2 hierarchies (substance abuse and psychiatric) that grouped patients' ICD-9CM codes based on diagnosis and expected cost	Grouping
Independent Hospital Pricing Authority, 2015 [46]	Australia, ambulatory episodes from 3 states	9,976 community episodes (adults and children)	Australia Mental Health Care Classification (AMHCC): 46 community terminal groups, out of 91 groups for all care settings. Variables: 5 phases of care, age, HoNOS [55], Life Skills Profile (LSP-16) [68].	Grouping
			MH-CASC [35]	Grouping

Author, Year	Context	Sample Size	Case-Mix System(s)	Type
Martin et al., 2017 [76]	UK, 11 child and adolescent mental health service sites	4,573 completed outpatient periods (50% development, 50% validation)	Child and Adolescent Mental Health Services Need-Based (CAMHS): 19 terminal groups. Variables: getting advice/help/more help, diagnosis, and NICE guidance for mental health and substance use disorders [77]	Grouping

Table 2.2 Empirical results of case-mix systems predicting proxy measures of resource use, ordered by name of the case-mix system and year

Case-Mix System	Resource Measure	Performance Measure
CAMHS, complexity factors, contextual problems, education, employment, training [76]	Number of appointments for closed-cases (without activities for $\geq$ 6 months)	R^2 = 5.0%, R^2 (with provider effect) = 12.1%
CCI/MCI [74]	Number of total visits (6 months)	R^2 = 5.7%
	Number of medical visits (6 months)	R^2 = 3.4%
	Number of mental health visits (6 months)	R^2 = 14.3%
CCI/MCI and demographics	Number of total visits (6 months)	R^2 = 6.7%
variables [74]	Number of medical visits (6 months)	R^2 = 4.6%
	Number of mental health visits (6 months)	R^2 = 14.6%
CCI/MCI, demographics	Number of total visits (6 months)	R^2 = 7.5%
variables, and patient self-reported health status	Number of medical visits (6 months)	R^2 = 5.3%
[74]	Number of mental health visits (6 months)	R^2 = 15.9%
Chronicity [60]	Number of prior hospitalizations (categorical)	χ^2 = 419.5 (p = 0.000)
	Prescription of major psychotropic drugs (binary)	ANOVA F = 4.64 (p = 0.01)

Case-Mix System	Resource Measure	Performance Measure
DRG [58]	Number of outpatient sessions	Hartley's F max p-value < 0.01, Cochran's C p-value < 0.01, Barlett-Box F p-value < 0.01 (groups variances were not homogeneous)
LONCA [62]	Number of hospitalization, past 12 months (categorical)	Cramer's V = 0.17
MCAS [61]	Hospitalizations admission (next 2 years) or involuntary admission (next 18 months) to state hospital	$\chi^2 \geq 6.05$ ($p < 0.05$)
PCI/MCI [74]	Number of total visits (6 months)	$R^2 = 5.4\%$
	Number of medical visits (6 months)	$R^2 = 3.3\%$
	Number of mental health visits (6 months)	$R^2 = 14.4\%$
PCI/MCI and demographics variables [74]	Number of total visits (6 months)	$R^2 = 6.6\%$
	Number of medical visits (6 months)	$R^2 = 4.6\%$
	Number of mental health visits (6 months)	$R^2 = 14.6\%$
PCI/MCI, demographics variables, and patient self-reported health status [74]	Number of total visits (6 months)	$R^2 = 7.7\%$
	Number of medical visits (6 months)	$R^2 = 5.5\%$
	Number of mental health visits (6 months)	$R^2 = 15.8\%$
PsyCMS [50]	Annualized mental health and substance abuse outpatient visits	R^2 (retrospective) = 24.4%, R^2 (prospective) = 6.5%

Case-Mix System	Resource Measure	Performance Measure
VA-MH12 [50]	Annualized mental health and substance abuse outpatient visits	R^2 (retrospective) = 17.0, R^2 (prospective) = 4.6%

Table 2.3 Empirical results of case-mix systems predicting direct measures of resource use, ordered by name of the case-mix system and year

Case-Mix System	Resource Measure	Performance Measure
ACG [63].	Total annual Medicaid claims (in- and out-patient), except nursing homes, drug claims, and intermediate care facility for the mentally retarded	R^2 (adults) = 2.0%, R^2 (children) = 4.1%
ACG [63].	Total annual Medicaid mental health and substance abuse claims	R^2 (adults) = 2.1%, R^2 (children) = 1.7%
ACG [65]	Total annual mental health and substance abuse related insurance claims, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 2.5%, R^2 (children) = 1.3%, R^2 (combined) = 2.3%
ACG [65]	Total annual mental health and substance abuse related insurance claims and out-of-pocket payments, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 3.0%, R^2 (children) = 1.4%, R^2 (combined) = 2.7%
ACG/ADG [50]	Total annualized inpatient and outpatient cost of mental health and substance abuse care Annualized mental health and substance abuse outpatient visits	R^2 (retrospective) = 4.8%, R^2 (prospective) = 2.6% R^2 (retrospective) = 11.1%, R^2 (prospective) = 2.8%

Case-Mix System	Resource Measure	Performance Measure
ADG [65]	Total annual mental health and substance abuse related insurance claims, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 7.6%, R^2 (children) = 3.9%, R^2 (combined) = 6.8%
ADG [65]	Total annual mental health and substance abuse related insurance claims and out-of-pocket payments, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 9.0%, R^2 (children) = 4.1%, R^2 (combined) = 7.9%
AMHCC [46]	Direct cost: wage-weighted staff time, indirect cost: allocated equally among all contacts at a unit for an episode of care (various lengths)	R^2 = 26.6%
CDPS [50]	Total annualized inpatient and outpatient cost of mental health and substance abuse care	R^2 (retrospective) = 8.3%, R^2 (prospective) = 5.4%
	Annualized mental health and substance abuse outpatient visits	R^2 (retrospective) = 14.7%, R^2 (prospective) = 4.0%

Case-Mix System	Resource Measure	Performance Measure
Demographics [65]	Total annual mental health and substance abuse related claims, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 0.3%, R^2 (children) = 0.3%, R^2 (combined) = 0.3%
Demographics [65]	Total annual mental health and substance abuse related insurance claims and out-of-pocket payments, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 0.3%, R^2 (children) = 0.4%, R^2 (combined) = 0.3%
Demographics, diagnosis, and comorbidity [65]	Total annual mental health and substance abuse related insurance claims, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 8.6%, R^2 (children) = 4.2%, R^2 (combined) = 7.6%
Demographics, diagnosis, and comorbidity [65]	Total annual mental health and substance abuse related insurance claims and out-of-pocket payments, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 9.9%, R^2 (children) = 4.7%, R^2 (combined) = 8.7%

Case-Mix System	Resource Measure	Performance Measure
Demographics (age groups and gender) [50]	Total annualized inpatient and outpatient cost of mental health and substance abuse care	R^2 (retrospective) = 0.4%, R^2 (prospective) = 0.4%
	Annualized mental health and substance abuse outpatient visits	R^2 (retrospective) = 2.1%, R^2 (prospective) = 0.8%
Diagnosis (schizophrenia, personality disorder, and social withdrawal) [53]	Log of community care cost - which equals to total annual clinic cost allocated to patients based on their contact duration for the year.	R^2 = 13.9%, R^2 (schizophrenia) = 2.6%, R^2 (personality disorder) = 6.2%, R^2 (social withdrawal) = 5.8%
Diagnosis (12 groups) [48]	Annual direct and indirect costs of outpatient care	R^2 = 7.0%
DCG/HCC [49]	Annualized contacts with providers	R^2 = 27.9%
DCG/HCC [50]	Total annualized inpatient and outpatient cost of mental health and substance abuse care	R^2 (retrospective) = 9.5%, R^2 (prospective) = 5.7%
	Annualized mental health and substance abuse outpatient visits	R^2 (retrospective) = 15.7%, R^2 (prospective) = 4.0%
Experimental Bundled Episodes [35]	Wage-weighted staff time over 8-week long bundled episodes (across all care settings)	R^2 = 12.6%, R^2 (outliers trimmed) = 27.9%
GAF [48]	Annual direct and indirect costs of outpatient care	R^2 = 3.1%

Case-Mix System	Resource Measure	Performance Measure
HCC [65]	Total annual mental health and substance abuse related insurance claims, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 5.5%, R^2 (children) = 2.8%, R^2 (combined) = 4.9%
HCC [65]	Total annual mental health and substance abuse related insurance claims and out-of-pocket payments, for both inpatient and outpatient settings, excluding nursing home or intermediate care facility for the mentally retarded	R^2 (adults) = 6.2%, R^2 (children) = 3.1%, R^2 (combined) = 5.5%
High/Medium/Low user groups diagnosis and funding source [54]	Log of 3-year utilization of all mental health services, including inpatient settings	Misclassification = 35.6%, R^2 = 18.3%
High/Medium/Low user groups diagnosis and funding source[54]	Log of 1-year utilization of all mental health services, including inpatient and outpatient settings	Misclassification = 39.2%, R^2 = 4.3%
High/Medium/Low user groups by funding source, diagnosis, and age [54]	Log of 3-year utilization of all mental health services, including inpatient and outpatient settings	Misclassification = 56.0%, R^2 = 4.3%

Case-Mix System	Resource Measure	Performance Measure
High/Medium/Low user groups by funding source, diagnosis, and age [54]	Log of 1-year utilization of all mental health services, including inpatient and outpatient settings	Misclassification = 49.7%, R^2 = 1.0%
High/Medium/Low user groups by funding source [54]	Log of 3-year utilization of all mental health services, including inpatient and outpatient settings	Misclassification = 40.5%, R^2 = 4.0%
High/Medium/Low user groups by funding source [54]	Log of 1-year utilization of all mental health services, including inpatient and outpatient settings	Misclassification = 39.5%, R^2 = 2.9%
LSP sub-scales (antisocial and bizarre behavior) [53]	Log of community care cost - which equals to total annual clinic cost allocated to patients based on their contact duration for the year.	R^2 = 14.9, R^2 (antisocial) = 12.9%, R^2 (bizarre behavior) = 2.0%
LSP sub-scales (antisocial) [53]	Log of community care cost - which equals to total annual clinic cost allocated to patients based on their contact duration for the year.	R^2 = 12.9
MH-CASC [35]	Wage-weighted staff time over 8-week long episode	Adult: R^2 = 5.7%, R^2 (outliers trimmed) = 12.7% Children/Adolescents: R^2 = 12.4%, R^2 (outliers trimmed) = 4.1% Combined: R^2 = 4.1%, R^2 (outliers trimmed) = 14.8%

Case-Mix System	Resource Measure	Performance Measure
MH-CASC [45]	Cost based on staff activity data attributable to clients for an episode of care (various lengths)	Adults: $R^2 = 3.5\%$, Child/Youth: $R^2 = 5.3\%$, Combined: $R^2 = 4.1\%$
MH-CASC [46]	Direct cost: wage-weighted staff time, indirect cost: allocated equally among all contacts at a unit for an episode of care (various lengths)	$R^2 = 5.9\%$
NZ-CAOS [45]	Cost based on staff activity data attributable to clients for an episode of care (various lengths)	Adults: $R^2 = 13.2\%$, R^2 (outliers trimmed) = 14.5% Child/Youth: $R^2 = 12.9\%$, R^2 (outliers trimmed) = 14.2% Combined: $R^2 = 13.5\%$, R^2 (outliers trimmed) = 15.1%
PsyCMS [50]	Total annualized inpatient and outpatient cost of mental health and substance abuse care	R^2 (retrospective) = 11.2%, R^2 (prospective) = 6.4%
Service-Connected Disability [48]	Annual direct and indirect costs of outpatient care	$R^2 = 1.6\%$
Service-Connected Disability and GAF [48]	Annual direct and indirect costs of outpatient care	$R^2 = 2.5\%$
SMI [73]	Difference between reimbursement based on average cost vs. case-mix adjusted rates	Difference range = -40.0% (approx. -$700,000) to 30% (approx. $1,000,000)

Case-Mix System	Resource Measure	Performance Measure
VA-MH12 [50]	Total annualized inpatient and outpatient cost of mental health and substance abuse care	R^2 (retrospective) = 9.6%, R^2 (prospective) = 5.9%

2.4 Discussion

2.4.1 Principal Results

A modest number of studies examined case-mix classification systems to predict mental health care resource use in the community settings. A direct comparison in terms of predictive performance was not possible due to the variation in the measures of resource use, the follow-up duration, and performance metrics. In general, it can be said that the large majority of the variation in community mental health resource use was still not accounted for by these case-mix classification systems.

Although, the majority of the research on this topic came from the USA, the Australian system (AMHCC) was most comprehensive, covering all ages and care settings (inpatient and community settings) [46]. The most recent innovation was the five phases of care (assessment only, acute, functional gain, intensive extended, and consolidating gain) which reflects the goal of care [46]. These phases of care can also be viewed as a proxy for a person's health care needs and, by extension, a person's expected resource use driven by health care needs.

2.4.2 Input Variables

It is worth acknowledging that when a case-mix classification system is used in a funding formula, it must ensure that resources are allocated equitably. Therefore, whether a variable should be a case-mix variable is an important consideration. In the literature, the variables used for classification were often grouped into only a few categories such as: demographics,

Table 2.4 Input Variables and Their Alternative Case-Mix Classifications

Variable	Number of Models	Needs	Individual	Provider	Process	Historical
Diagnosis	23/33	x	x			
Age	12/33		x			
Health Conditions	10/33	x	x			
Social Relations	10/33	x	x			
Mental Status	8/33	x	x			
Gender	7/33		x			
Functional Status	7/33	x	x			
Harm to Self or Others	7/33	x	x			
Behavior	7/33	x	x			
Substance Use	6/33	x	x			
Medication Usage	5/33	x	x			
Service History	4/33		x			x
Legal Status	4/33		x			
Insurance Benefits	3/33					
Care Settings	3/33			x		
Roles Functioning and Finances	3/33	x	x			
Living Conditions	3/33	x	x			
Treatments	3/33				x	
Cognition	3/33	x	x			
Communication and Vision	2/33	x	x			
Veterans Status	2/33		x			x
Ethnicity	1/33		x			
Stress and Trauma	1/33	x	x			x

diagnosis, clinical status, or treatment variables. Discussions regarding their appropriateness as case-mix variables were also rare. Using an alternative classification of these variables, this study summarized the scope of case-mix variables used in the literature and discussed how case-mix variables can influence funding allocation (Table 2.4).

Needs Variables Variables that indicate the level of health care needs are those that not only have high explained variance of the resource use, but should also be variables that directly drive the resource use. For example, ethnicity in [45] and gender in [50, 63, 65] may have high correlation with resource use, but such correlation may be confounded by other factors such as systematic marginalization in the society that can make someone more vulnerable to mental health disorders and, by extension, to have higher expected level of resource use. Therefore, future research should consider needs variables that directly drive resource use, such as: diagnosis, functional status or severity of illness, instead of those that simply correlate with resource use for reasons other than clinical needs.

Individual vs. Provider Variables Provider variables, in essence, describe why it costs more in one facility compared to another, regardless of the person's health care needs. For example, these can be care setting, facility type, regional characteristics, staff qualifications, or teaching status. Using these variables as case-mix variables essentially reinforces the systematic inequalities that exist among the providers. Therefore, using variables related to the individuals, whenever possible, may help avoid this reinforcement. However, in some cases, reinforcing systematic inequalities may be desirable, such as: adjusting for facilities located in rural areas where resources and supplies may cost more to be delivered. Only the case-mix classification systems from Australia and New Zealand used care setting as a case-mix variable, but they were used as the first split to essentially join 2 separate case-mix systems for inpatient settings and community settings together [35, 45, 46].

Process Variables Process variables are those that describe treatments or services given to a person, found in [35, 45, 46, 76]. When using treatments or services as case-mix variables, they may encourage providers to do more of them for financial gain, if they are under the control of the providers. Similar to provider variables, consideration should be given to whether variables that describe the needs of the individual should be used as much as possible, or if there is a valid rationale for reinforcing differences in such variables.

Historical Variables Variables that describe historical use of services or treatments provided can be viewed as proxies for historical needs, such as prior hospitalization in [60], or usage in a prior year in [54]. The shortcoming of these variables is that they have limited ability to be modifiable and change with current needs. On the other hand, there are historical variables that are continued to be relevant to current needs, for example: past history abuse or violence in [35]. Historical variables therefore should not be entirely discounted, but the important consideration is whether historical variables have long-term relevance in describing a person's current health care needs, or whether another variable that is more dynamic and could change with a person's health care needs may be more appropriate.

Ambiguity of Variables Ambiguity may arise if the variables chosen to describe the patient type result in more than one way to classify an individual. This ambiguity may give providers an incentive to choose the classification that maximizes the reimbursement, especially if the differences in the expected resource use or reimbursement of the possible classifications are significant. Given the same input, a good case-mix system should be able to consistently output only one classification.

2.4.3 Output Variables

The use of proxy measures of resource use was common in this review, such as the number of visits or appointments (Table 2.2). In fact, the very first case-mix classification system (DRG) used length of stay as a proxy for an inpatient episode's cost [59]. This approach assumed that costs of care do not vary day-to-day during the hospitalization [27].

Similarly for direct measures of resource use, when assuming that the costs of care do not vary day-to-day or visit-to-visit, it is possible to calculate the costs of care for a particular case on a per-diem basis by multiplying the number of days/visits with expected cost per day/visit. The studies using direct measures of resource use found in this review; however, all calculated costs of care on an episodic basis with a pre-defined follow-up length (Table 2.3). An analysis from Australia showed that the preferred method of predicting resource use in community settings was a pre-defined episode with fixed length, due to the chronic nature of mental health care and community-based services are provided intermittently, instead of continuously as the inpatient settings [35].

The class of direct measures of resource use can be further divided into billed costs (i.e. claims data) or observed costs (i.e. staff time study). Billed costs have three main limitations [27]. First, they often include non-clinical administrative costs (such as management, and claims department), which could reduce the variance in the resource use measure if the administrative costs are high relative to costs of clinical care [27]. Second, billed costs were most likely derived by averaging over a large number of patients rather than the actual amount an individual patient consumed [27], which could also reduce the variance in the resource use measure. Third, additional variance can be added if there is a lot of variation in accounting practice across different facilities [27]. On the other hand, observed costs like staff time activities are more likely to closely match the actual resource consumption by individual patients and potentially more responsive to patients' characteristics [27], but may be harder to obtain than available administrative data [31].

2.4.4 Gaps in Current Research

There are many available case-mix systems that were developed for inpatient settings but were not tested for community settings. Creating a case-mix system is not a trivial process; however, considerable progress can be made by experimenting with existing case-mix systems developed for use in another setting. For example, the SCIPP is a good candidate for testing in community settings it has reported 26.3% explained variance of inpatient psychiatry cost using clinical characteristics, and higher than most of the identified case-mix systems [27].

It has been shown that children and adolescents also have unmet mental health care needs [78]. Most of the studies only focused on adult populations (Table 2.1). Therefore, future case-mix classification systems should also consider children and adolescent populations in the development of new case-mix systems.

Only three of the studies cross-validated the predictive performance of their systems on a different data set than the one used for model derivation [49, 50, 76]. Cross-validation can serve two purposes: to evaluate the generalizability of the model on unseen observations or future users of the health care system, and to compare competing models [79]. Future research should consider using cross-validation when evaluating the predictive performance because the uncross-validated performance metric may give an overestimation.

Lastly, it was not always clear if there exists a process or mechanism for updating the case-mix systems and exchanging knowledge. Therefore, it is important to have a robust feedback loop by conducting more replication studies to validate case-mix systems under different conditions, as new data become available if using administrative data, or with more participating sites and over different time periods if using staff time activity data. For example, Australia has an organization dedicated to continuous improvement of case-mix classification systems with more replication studies planned [46].

2.4.5 Limitations

This study was not without limitations. First, this study only examined articles written in English, which also limited our review to only English-speaking jurisdictions. Second, this study did not consider the implementation outcomes and policy impacts of the identified case-mix systems, which deserve a separate review in the future.

2.4.6 Conclusion

This study provided a summary of the scope of research in community mental health care case-mix classification. The research activity was modest, while the transition from institutionalization to community care continues to evolve. Consideration should be given to appropriateness and assumptions of the case-mix variables, resource use measure, and evaluation of predictive performance. More research, especially of replication type, is needed in community mental health to ensure resources are meeting the needs of the population as new data become available and as the health care system evolves over time.

Chapter 3
Study One: Mental Health Care Transition from Inpatient Psychiatry to Community Settings: Patterns from Waterloo-Wellington, Ontario, Canada

Abstract

Although mental health care has been gaining recognition as a priority in Canada and more funding was recommended by the Mental Health Commission of Canada, integration of services across the continuum of care and lifespan remained an elusive goal. One point of potential vulnerability was examined in this study, the transition between inpatient psychiatry and community mental health services in the province of Ontario, Canada. Individuals discharged from inpatient psychiatry were followed to observe their subsequent use of community mental health services. Using the Resident Assessment Instrument - Mental Health (RAI-MH) assessment at discharge from inpatient psychiatry as the baseline clinical profile, factors that are associated with readmission, usage and high usage of community mental health services post-discharge were examined. This study found that only 55% of the discharges would subsequently use publicly funded community mental health services. The clinical profile given by the RAI-MH assessment was shown to be associated with higher usage of community mental health services.

This study also showed that receiving community mental health services post-discharge may be beneficial in reducing readmission and the demand of intensive inpatient services for the system. The findings suggested that sharing and meaningful use of the clinical assessments, such as the RAI-MH, can play a larger role in achieving an integrated mental health care system.

3.1 Introduction

Mental illness is one of the leading causes of disability in Canada [4, 80, 81]. The Mental Health Strategy for Canada has called for improvements in access to mental health services across the continuum of care and lifespan [82–86]. One commonly suggested solution is to increase funding for mental health care proportionally to match the disease burden, often measured by health-adjusted life years or costs of health services and lost productivity [82, 87, 88]. For example, in the province of Ontario, the burden of mental illness accounts for about 10% of the total disease burden, but only accounts for about 7% of the health care spending [81, 89].

While funding is critically important, improvements to service access across the continuum of care can also be achieved by making mental health care services more integrated throughout the health care system. Over the years, the health care system in Ontario has gone through many reforms. An elusive goal has been to better integrate health care services across different care settings. For example, the creation of the Local Health Integration Networks (LIHNs) and the newly proposed Ontario Health Teams (OHTs) are intended to encourage integration across complex organizational boundaries and at the point of care [90]. Transition from one care setting to another has been identified as a point of vulnerability for patients, especially for those with chronic and complex health conditions [91–94]. This study aims to enhance the understanding of the transition from inpatient psychiatry to community mental health services by examining the patterns of use at the transition and factors associated with the usage of community mental health services.

Patients discharged from a psychiatric hospital in Waterloo-Wellington region, Ontario, Grand River Hospital (GRH), were followed to examine their usage of community mental health services at the Canadian Mental Health Association - Waterloo Wellington (CMHA-WW) subsequently. The GRH is an acute hospital with inpatient psychiatric beds and the CMHA-WW provides mental health services in the community settings. The CMHA-WW is the largest chapter of the Canadian Mental Health Association (CMHA) by staff count, and serves all age groups. CMHA-WW is a good example to study because it is a stand-alone entity that offers a comprehensive range of community mental health services.

3.1.1 Community Mental Health Services in Waterloo-Wellington region, Ontario

The Waterloo-Wellington region has 12 publicly funded agencies providing mental health and addiction services: three acute hospitals providing primarily inpatient psychiatry (includes GRH), three residential treatments and supportive housing agencies, three addiction services agencies, one agency providing primarily children/youth counseling, one agency primarily providing family counseling, and CMHA-WW. The bulk of community mental health services, except for addiction services, are provided by CMHA-WW due to their capacity and complete presence across the region.

There are 3 primary pathways of initiating adult community mental health services (Figure 3.1): (1) self-referral or referred by a third party to the centralized intake (Here 24/7 operated by CMHA-WW); (2) contact with the police or justice system; and (3) referral by family physician or care team to specialized geriatric services. Under normal circumstances, an adult in the region seeking community mental health services through self-referral would first come into contact with CMHA-WW via the centralized intake for an assessment and/or scheduling of services (Figure 3.1).

Figure 3.1 and Table 3.1 represents major programs or packages of services that have been designed to meet a certain objective or target populations (such as: adults vs. seniors, or by diagnosis). The source of the funding could also determine the scope of the program. Some programs were designed to connect the clients with appropriate services that they require, while others can be a specialized treatment program. A client may be enrolled in more than one program at a time depending on their needs.

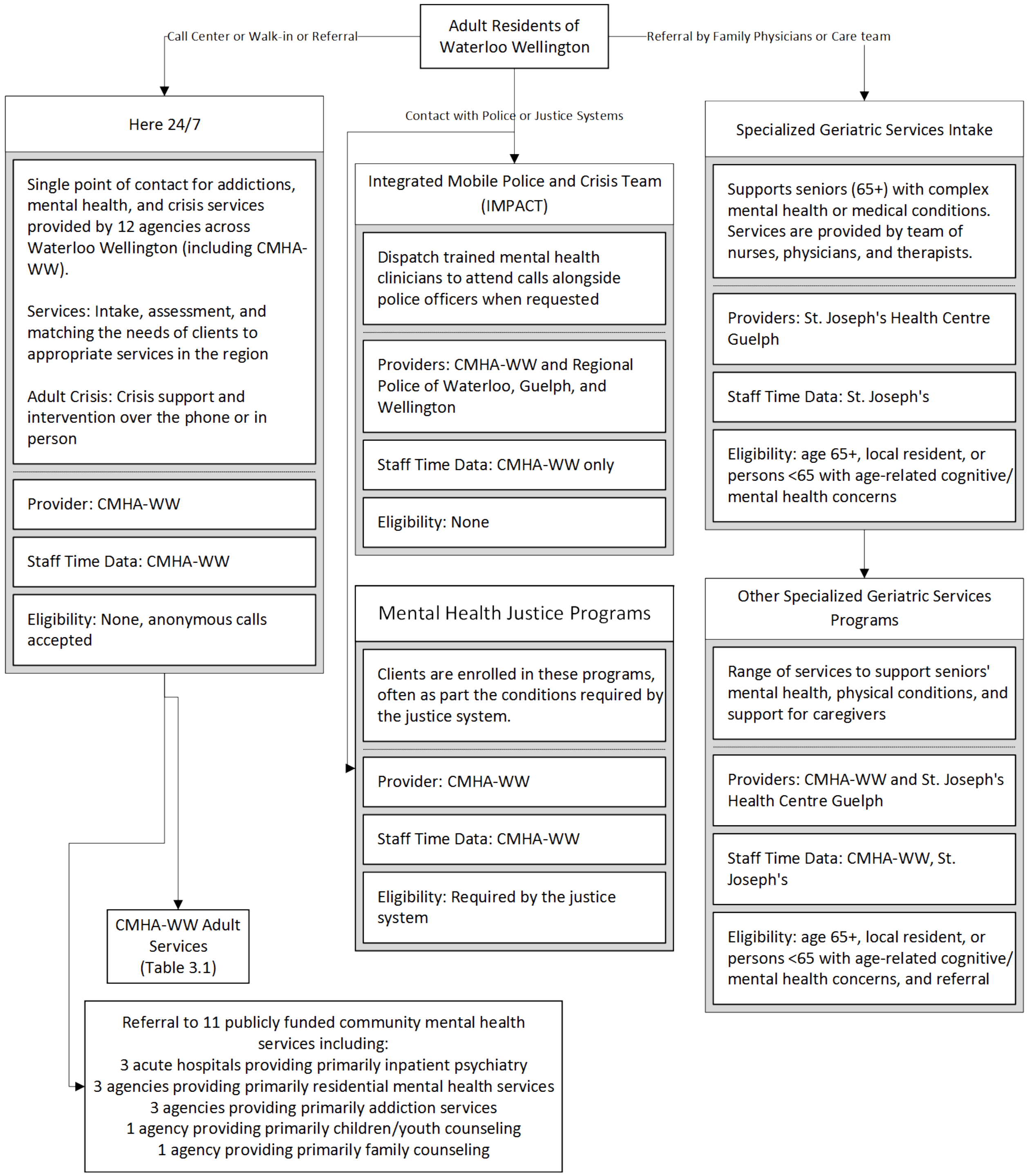

Figure 3.1 Pathways of initiating services at the community mental health agency, compiled based on the candidate's experience working at CMHA-WW and conversations with CMHA-WW staff.

Table 3.1 Portfolio of adult programs at the community mental health agency

Programs	Description	Eligibility
Adult Counseling & Treatment	Short term individual or group therapy sessions, including assessment diagnosis, and consultation.	18 or older, moderate to severe diagnosis, complex needs, concurrent disorders, have identified treatment goals, willing to work on goals outside counseling sessions. Exclude: age related cognitive decline, eligible for family or private therapist
Adult Psychiatry	Consultation with a psychiatrist in person or over tele-psychiatry	Referred by a family physician, local resident 18 or older, moderate/severe diagnosis, likely to become unstable without intervention
Dialectical Behavior Therapy	Cognitive behavioral treatment for suicide behavior and emotional dysregulation	18 or older. Excluded: diagnosed with schizophrenia, psychotic episode within 6 months, significant development disability, or learning disability
Eating Disorders	An initial assessment followed by primarily group-based therapy	Local residents, medically fit for services, not psychotic or suicidal, and to be monitored by physicians or nurse practitioner.
Early Psychosis - First Step	Early psychosis intervention program that assesses and treats young people experiencing first episode of psychosis.	Ages 14-35, local residents, first episode of psychosis within 1 year and has not been treated with medication for >6 months.
Flexible Assertive Community Treatment Team	Specialized team that provides flexible treatment from intensive care during crisis to less intensive care.	Local residents 18+ experiencing severe and persistent mental health issues
Intensive Support Coordination	Short and long-term support is available, assisting clients with personal planning, crisis planning, referral, and connection to other community resources, education, and employment goals.	Local residents 18+ with mental health issues that limit their ability to function on a daily basis
Mental Health Promotion - Family Initiatives	Group education with family member or care givers of adults with mental health or addiction issues.	family members or care givers of adult local residents with mental health or addiction issues, and also 18 or older
Self-help Services	Skills building, peer support groups. No intake, registration, or waitlist.	Local adult residents
Specialized Medical Addictions & Mental Health Outreach Services	Team of staff reach out to adults who are homeless or at risk of homelessness and disconnected from services for their mental health/addictions/concurrent issues.	18+, homeless or at risk of homelessness, living with mental health/addictions/concurrent issues

3.2 Methods

3.2.1 Data Sources

The discharge RAI-MH assessment, which is mandated for all adult inpatients psychiatric admission in Ontario, were obtained from GRH [95]. This data source provided the administrative information of the inpatient episode and many measures of clinical characteristics of a patient at the time of discharge. The discharge assessments are typically done within 3 days prior to discharge from inpatient psychiatry. All adult inpatient psychiatric discharges from the GRH between 2014 - 2018 were obtained. This time frame was chosen because the data represent the assessments done using the latest version of the RAI-MH assessment used in Ontario, and they represented the current organization structure of CMHA-WW since the last major organizational change in 2014. Discharges due to hospital-to-hospital transfers, discharges that were followed by a same-day readmission, and discharges due to in-hospital mortality were excluded because these were not likely to initiate services at the community mental health agency post-discharge.

From the community mental health agency, service records between 2012 – April of 2019 were obtained for all clients matching those discharged from GRH. Each service encounter was recorded as an event that contained the date of the service, the job title of the staff performing the service, duration of the staff time, and name of the program which was the basis for the service. Only the direct staff time was considered in the resource use measure, which is the resource that is driven directly by the health care needs of the clients.

Direct staff time only included direct contacts with clients such as: in-person services, over-the-phone services, or teleconference calls with clients. Indirect staff time includes all other client-specific activities, but are not driven by clients' clinical characteristics and typically not provided direct to a client or provided without client being present (such as: documentation, travel to/from client's meeting, and case review travel to/from client's meeting). Possible data entry errors were also manually checked and corrected if necessary, such as correcting for AM/PM in time entries for service events longer than a typical work shift of eight hours.

For group services, which were provided to more than one client at the same time, the staff time was divided equally among the number of clients registered for the group. If the number of clients attended the group session was lower than the number registered, the number of registrants were used instead because resources have already been assigned for the service from the organizational perspective.

The community mental health agency also partnered with other external organizations to deliver services by pooling human or capital resources together, such as the senior mental health and specialized geriatric programs, which is a joint venture between CMHA-WW and St. Joseph's Health Centre in Guelph, ON (Figure 3.1). Only the services that were provided solely by CMHA-WW salaried staff were included in this study for two reasons. First, the staff time activity data was available for some joint programs but not all. Second, there are services of the joint programs that may not be purely mental health services, such as general geriatric services that is more related to primary care. Additionally, services provided on a fee-for-service basis by the provincial health plan, such as psychiatrist or physician services, were not included. Overall, the usage of services were analyzed from the point of view of the community mental health agency and included only costs incurred by this agency. These data were expected to be of good quality because they were used for scheduling of appointments, monitored by management, and used to determine extra pay (such as: overtime or pandemic pay in 2020) for eligible direct face-to-face time.

Data were primarily linked using health card number and date of birth (Figure 3.2). In absence of health card number, possibly due to lack of health coverage at one point in a person's lifetime or changes in health card number, the secondary linkage method used the date of birth and full name instead.

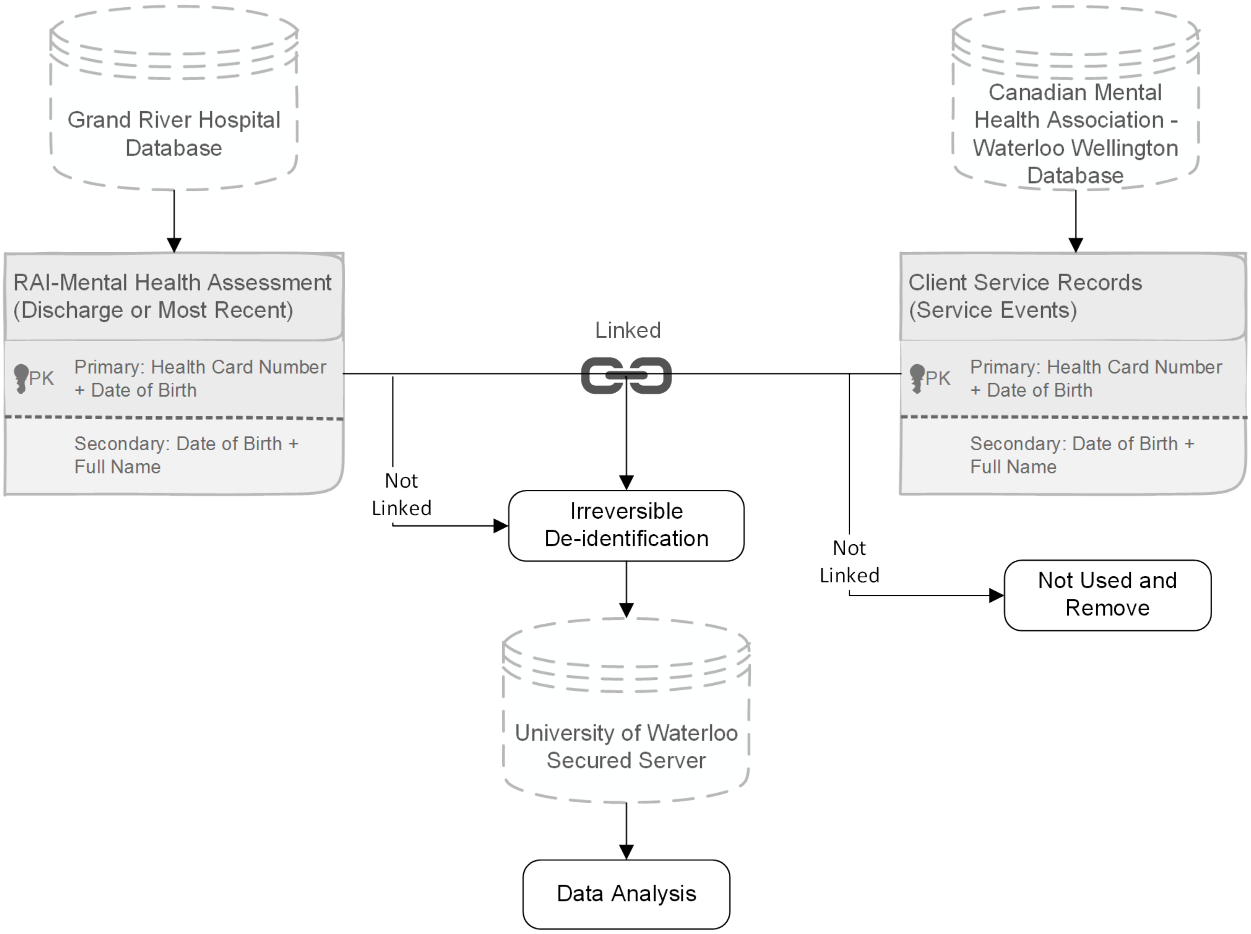

Figure 3.2 Linkage of Data Sources

Prior to data linkage, some pre-processing steps were applied to the identifiers. For health card number, the last two characters of the health card number were removed because these characters are changed every 5 years when a health card is renewed in Ontario, while the first 10 characters are fixed throughout a person's lifetime. All dashes and blank spaces were removed to resolve discrepancies of data entry. All dates were converted to YYYY-MM-DD format for consistency. For names, all the characters were converted to lower cases to eliminate case-sensitivity during linkage, and all punctuation or special characters were removed.

After linkage, personal identifiers were removed. All matched records were assigned a randomly generated ID to use for the study. Data were then stored at interRAI Canada, University of Waterloo secure server. This study received ethics clearance from the University of Waterloo (file number 40147) and the Grand River Hospital (file number 2018-0669). The ethics committee of the CMHA-WW approved this study on October 29th, 2018.

In addition, this study also compared the GRH data sample against the provincial population by obtaining additional de-identified RAI-MH assessment data from the Ontario Mental Health Reporting System (OMHRS) provided by the Canadian Institute for Health Information. Inpatient episodes from the same health region (the lowest level of identifier) as the GRH, which is Waterloo-Wellington, were excluded to remove the effect of autocorrelation. The use of OMHRS also received ethics clearance from the University of Waterloo (file number 19917).

3.2.2 Descriptive Analysis

The pattern of usage of community mental health services provided by CMHA-WW was described by examining usage pre- and post-inpatient psychiatry episode, time between discharge and service initiation. To quantitatively describe the sample, several clinical scales embedded in the assessment that measure both behavioral and physical characteristics were used [96].

The Activities of Daily Living Hierarchy Scale (ADL, range 0-6) measured a person's ability to perform personal hygiene, locomotion, toilet use, and eating [97]. The Instrumental Activities of Daily Living Scale (IADL, range 0-42) measured a person's higher level function required for daily living [97]. The Cognitive Performance Scale (CPS, range 0-6) measured the cognitive status [98]. The Depressive Severity Index (DSI, range 0-15) measured severity of the depressive symptoms [96]. The Positive Symptoms Scale short-form (PSSS, range 0-12) measured the frequency of positive symptoms [96]. The Aggressive Behavior Scale (ABS, range 0-12) measured the frequency and variety of aggressive behaviors [99]. The Risk of Harm to Others scale (RHO, range 0-6) estimates the risk of violent behavior that could harm others [100]. The Severity of Self-harm scales (SOS, range 0-6) estimates the risk of self-harm [96]. All of these scales indicate higher severity or lack of capacity with higher scores.

3.2.3 Modeling of Readmission

This analysis examined whether the rehospitalization clinical assessment protocol measured at discharge is associated with 30-day same hospital readmission using a multiple logistic regression [101]. This clinical assessment protocol is composed of multiple elements from the RAI-MH assessment that were shown to predict psychiatric readmission: number of prior hospitalizations, unemployment, substance use, positive symptoms, and risk of self-harm [101].

An admission or readmission provided inpatients with intensive services, such as intensive observation, diagnosis, and treatment [8]. Some of these services can also be provided by the community mental health agency via the Assertive Community Treatment (ACT) services. Unlike typical community services, ACT services does not require clients to initiate services or cancel services if appointments are missed, but aims to be assertive and persistent in engaging clients, especially hesitant ones [102, 103]. ACT teams are multidisciplinary that have shared caseloads among their members and make the care plan together to maintain continuity of care over time [102]. Evidence from randomized trials showed that ACT services have the effect of reducing hospital use by shifting intensive services into the community [104]. Therefore, this analysis also included usage of ACT (binary) within 30 days of discharge or being readmitted to inpatient psychiatry, whichever comes first, as a covariate. The Kaplan-Meier curves, a log-rank test, and Cox proportional hazards regression were also used to examine a possible relationship between usage of ACT services (binary) time until readmission to the same hospital, with a follow-up period of 30 days post-discharge.

3.2.4 Modeling of Community Mental Health Service Use Post-Discharge

Multiple logistic regression models were used to examine the relationships between clinical characteristics measured by the discharge assessment and usage (binary), high usage (top 10 and 20 percentiles of direct staff time) of community mental health services within 180 days of discharge, and usage (binary) of some specialized programs targeting persons at high risk of harming others (Mental Health Justice) and self-harm (Self-Help Skills for Safer Living).

Several reviews had summarized the known predictors of mental health resource use [36, 37, 105], which included age, sex, historical service usage, diagnosis, severity of illness, and behavior problems. Multiple logistic regression models were fitted using: age, sex, diagnosis, length of inpatient psychiatric episode as the historical service usage indicator, clinical scales (ADL, IADL, CPS, DSI, PSSS) as severity of illness indicators, and two safety clinical scales (SOS, RHO) as behavioral problems indicators [95]. Additionally, readmission could reduce the likelihood that a person would use the community mental health agency services, therefore readmission was also included in the model. Recognizing that these predictors could be correlated with each other, a model with all predictors was tried first, and non-significant predictors were removed one at a time to arrive at a final model that contained only statistically significant predictors. Further examinations were also done when unexpected associations were observed, such as for risks of self-harm and harming others.

Some observations were right-censored because the data were extracted earlier than 180 days post-discharge. Additionally, it could be argued that frequent users of the health systems may be treated differently if they were known to the system, such as people with multiple inpatient episodes, which meant that each assessment might not be an completely independent observation. Therefore, the sensitivity of the models were also examined on several subsamples of the dataset: full sample, sample with no censored observations, sample with first and single admissions only, and sample with last and single admissions only.

3.3 Results

There were a total of 4,688 discharge assessments (2,874 unique persons) obtained from GRH (Figure 3.3). Only about half of these discharges (2,312 discharges or 1,571 unique persons) subsequently received services from the community mental health agency post-discharge.

Due to the limited scope of the research ethics application, the number of adult clients who received services but were not previously admitted to inpatient psychiatry was not available. However, it was estimated that about 10,000 unique individuals came into contact with the community mental health agency per year, including children and contacts made through the centralized intake but subsequently referred to external agencies for services.

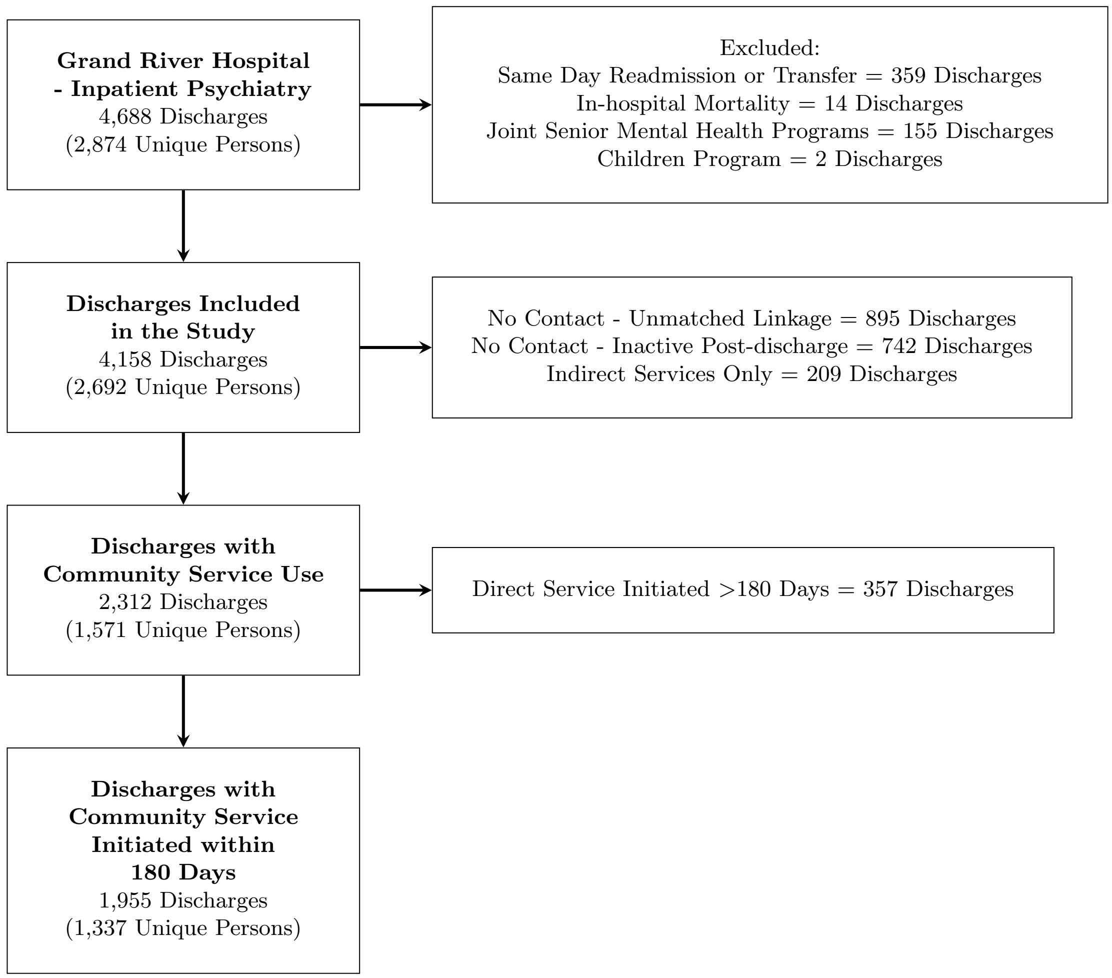

Figure 3.3 Record linkage between Grand River Hospital and Canadian Mental Health Association - Waterloo Wellington

3.3.1 Descriptive Analysis

3.3.1.1 Sample Characteristics

Table 3.2 showed that the distribution of sex and cognitive disorders were similar between the study sample and the rest of the province. However, age, clinical scales, diagnoses of schizophrenia, mood, and substance use disorders appeared to deviate from the rest of the province (Table 3.2).

Table 3.2 Demographics and clinical characteristics of inpatients discharged from Grand River Hospital versus the rest of Ontario

	GRH % (n = 4,158)	ON % (n = 174,338)	Chi-square p-value
Female	48.9% (2,034)	48.3% (84,268)	0.47
Age 18-25	23.5% (977)	18.1% (31,605)	<0.001
25-44	39.7% (1,649)	39.4% (68,762)	-
45-64	28.8% (1,199)	30.8% (53,713)	-
65+	8.0% (333)	11.6% (20,258)	-
ABS 0	88.7% (3,689)	82.4% (143,733)	<0.001
1-3	7.0% (293)	11.4% (19,833)	-
4-6	3.0% (124)	4.5% (7,831)	-
7-12	1.3% (52)	1.7% (2,941)	-
ADL 0	95.2% (3,958)	92.7% (161,536)	<0.001
1-3	3.6% (148)	6.5% (10,363)	-
4-6	1.3% (52)	1.4% (2,439)	-
CPS 0	90.1% (3,746)	76.5% (133,285)	<0.001
1-2	7.4% (307)	19.6% (34,184)	-
3-6	2.5% (105)	3.9% (6,869)	-
DSI 0	59.5% (2,472)	56.1% (97,791)	<0.001
1-3	29.1% (1,209)	28.4% (49,459)	-
4-7	9.3% (388)	11.4% (19,885)	-
8-15	2.1% (89)	4.1% (7,203)	-
IADL 0	85.2% (3,541)	67.6% (117,823)	<0.001
1-3	4.5% (186)	12.8% (22,299)	-
4-9	3.8% (160)	9.6% (16,725)	-
10-18	2.7% (111)	5.2% (9,056)	-
19-30	3.8% (160)	4.8% (8,435)	-
PSSS 0	82.8% (3,443)	68.4% (119,274)	<0.001
1-6	15.9% (660)	28.5% (49,620)	-
7-12	1.3% (55)	3.1% (5,444)	-
RHO 0-2	91.6% (3,807)	82.9% (144,549)	-
3-4	5.7% (239)	11.8% (20,642)	-
5-6	2.7% (112)	5.2% (9,147)	-
SOS 0-3	75% (3,118)	65.8% (114,760)	-
4	6.2% (257)	9.8% (17,086)	-
5-6	18.8% (783)	24.4% (42,492)	-
Schizophrenia	40.7% (1,691)	37.3% (65,035)	<0.001
Mood Disorders	39.7% (1,651)	44.0% (76,701)	<0.001
Cognitive Disorders	4.3% (179)	4.4% (7,660)	0.81
Substance Use Disorders	25.1% (1,042)	28.2% (49,085)	<0.001

Abbreviations: ABS, Aggressive Behavior Scale; ADL, Activities of Daily Living Hierarchy Scale; CPS, Cognitive Performance Scale; DSI, Depressive Severity Index; IADL, Instrumental Activities of Daily Living Scale; PSSS, Positive Symptom Scale Short-Form; RHO, Risk of Harm to Others; SOS, Severity of Self-Harm.

3.3.1.2 Contact with the community mental health agency prior to Inpatient Psychiatric Admission

Prior contact with the community mental health agency was defined as a direct service event which took place prior to admission to inpatient psychiatry. From all the psychiatric episodes included in this study (n = 4,158), about 60% were preceded by at least one contact with the community mental health agency prior to admission (Figure 3.4). However, the data of this study only tracked prior contacts with the community mental health agency up to 2012. The mean time since the last contact with the community mental health agency was 172.3 days (min = 1 day, median = 31.5 days, max = 6.4 years).

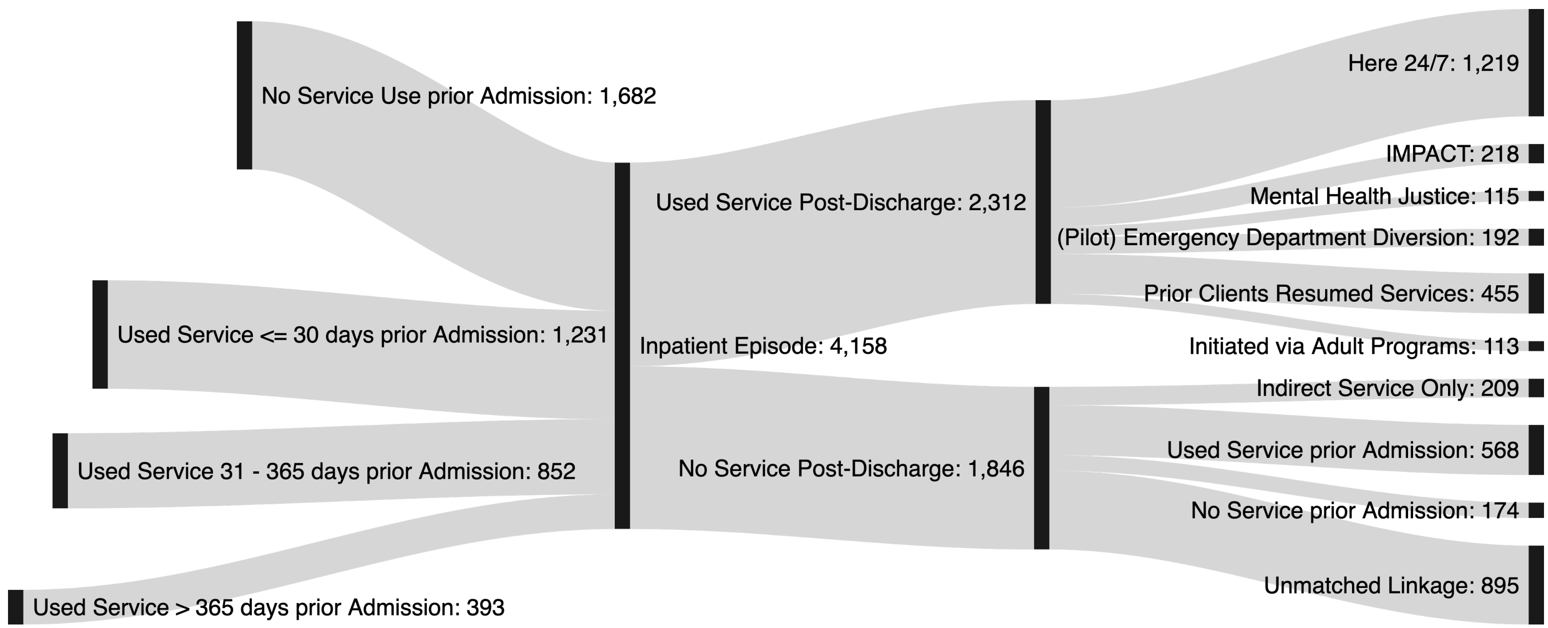

Figure 3.4 Flows of contacts with the community mental health agency prior to inpatient psychiatric admission, first direct service event with the community mental health agency post-discharge, and subgroups episodes without usage post-discharge (n=4,158 inpatient episodes) between 2012 – 4/2019

3.3.1.3 Pathways of Initiating Services at the community mental health agency Post-Discharge

After the inpatient psychiatric episode at GRH, about 55% of discharges used community mental health services at the community mental health agency post-discharge (Figure 3.4). The majority initiated their services through the centralized intake (Here 24/7), followed by prior clients resuming services they may have started prior to the inpatient episode. About 14% initiated services through the two criminal justice system programs (IMPACT and Mental Health Justice programs).

In the past, a pilot program called Emergency Diversion by the community mental health agency, which aimed to identify people who required immediate community support services but did not necessarily require inpatient services (Figure 3.4). Since it was only a pilot program, it was not included as a major pathways of initiating services in Figure 3.1.

For the rest, it was not immediately clear through which pathways services were initiated with the community mental health agency post-discharge (Figure 3.4). Since the data for this study only tracked back to 2012, the remainder of the sample were likely previous clients who had initiated services in some adult programs prior to the inpatient episode.

3.3.1.4 Discharges without Usage of Community Mental Health Services Post-Discharge

From the discharges that did not use the community mental health agency services, about 11% received indirect services (Figure 3.4). Indirect services implied that a staff from the community mental health agency had performed some activities on the client's file, such as an attempted contact, case review, referral review, or client refused services. About 31% of these episodes were preceded by a prior contact with the community mental health agency. It was not immediately clear from these data why these clients were inactive in using community services post-discharge.

A small portion of these discharges were successfully linked across two data sources but had neither contact prior-admission or post-discharge (Figure 3.4). This was likely due to the fact that the service use during this period was part of the joint partnership programs that were not included in this study.

3.3.1.5 Time to Service Initiation Post-Discharge

Of those that used the community mental health agency post-discharge within 180 days, the majority initiated a direct service event shortly after discharge (min = 0 days, median = 6 days, mean = 24 days, max = 180 days).

3.3.1.6 Compare First Episodes' Assessments vs. Last Episodes' Assessments

The demographics and clinical characteristics of the two subgroups of assessments done at different times were mostly similar (Table 3.3). Variables that showed statistical differences between the two subgroups were CPS, SOS, and mood disorder diagnosis.

Table 3.3 Summary proportions and chi-square test of independence of the characteristics of discharge assessments from the first inpatient psychiatric episode versus the last episode for persons with multiple inpatient admissions

	First Episodes (n = 772)	Last Episodes (n = 772)	p-value
Age 18-25	23.6% (182)	19.8% (153)	0.20
25-44	42.0% (324)	42.5% (328)	-
45-64	29.4% (227)	31.0% (239)	-
65+	5.1% (39)	6.7% (52)	-
ABS 0	87.7% (677)	88.5% (683)	0.89
1-3	8.3% (64)	7.4% (57)	-
4-6	2.7% (21)	2.6% (20)	-
7-12	1.3% (10)	1.6% (12)	-
ADL 0	96.2% (743)	95.6% (740)	0.91
1-3	3.2% (25)	3.5% (27)	-
4-6	0.5% (4)	0.6% (5)	-
CPS 0	88.6% (684)	92.2% (712)	0.01
1-2	10% (77)	6% (46)	-
3-6	1.4% (11)	1.8% (14)	-
DSI 0	58.8% (454)	62.6% (483)	0.16
1-3	31% (239)	26.6% (205)	-
4-7	8.2% (63)	9.5% (73)	-
8-15	2.1% (16)	1.4% (11)	-
IADL 0	84.7% (654)	83.9% (648)	0.38
1-3	5.6% (43)	4.7% (36)	-
4-9	3.9% (30)	4.1% (32)	-
10-18	3.5% (27)	3.2% (25)	-
19-30	2.3% (18)	4.0% (31)	-
PSSS 0	81.0% (625)	82.6% (638)	0.37
1-6	18.1% (140)	15.6% (123)	-
7-12	0.9% (7)	1.4% (11)	-
RHO 0-2	89.4% (690)	91.7% (708)	0.28
3-4	7.4% (57)	6.0% (46)	-
5-6	3.2% (25)	2.3% (18)	-
SOS 0-3	76.9% (594)	78.0% (602)	0.07
4	6.9% (53)	4.3% (33)	-
5-6	16.2% (125)	17.7& (137)	-
Schizophrenia	47.8% (369)	44.6% (344)	0.22
Mood Disorders	43.8% (388)	35.2% (272)	<0.001
Cognitive Disorders	2.3% (18)	3.4% (26)	0.28
Substance Use Disorders	24.7% (191)	26.2% (202)	0.56

Abbreviations: ABS, Aggressive Behavior Scale; ADL, Activities of Daily Living Hierarchy Scale; CPS, Cognitive Performance Scale; DSI, Depressive Severity Index; IADL, Instrumental Activities of Daily Living Scale; PSSS, Positive Symptom Scale Short-Form; RHO, Risk of Harm to Others; SOS, Severity of Self-Harm

3.3.2 Modeling of Readmission

Discharges that subsequently received ACT services within 30 days of discharge or until readmission had slower time until 30-day same hospital readmission than discharges that did not, 5.0% and 8.9% respectively (Figure 3.5). The estimated Cox hazard ratio was 0.54 [95% CI: 0.30 - 0.99] and the log-rank test p-value = 0.04. For context, the rate of 30-day psychiatric readmission to the same hospital of the rest of Ontario during the same period was 8.5%. Other research found that the rate of psychiatric 30-day readmission in Ontario to any hospital was between 7-9% [106, 107].

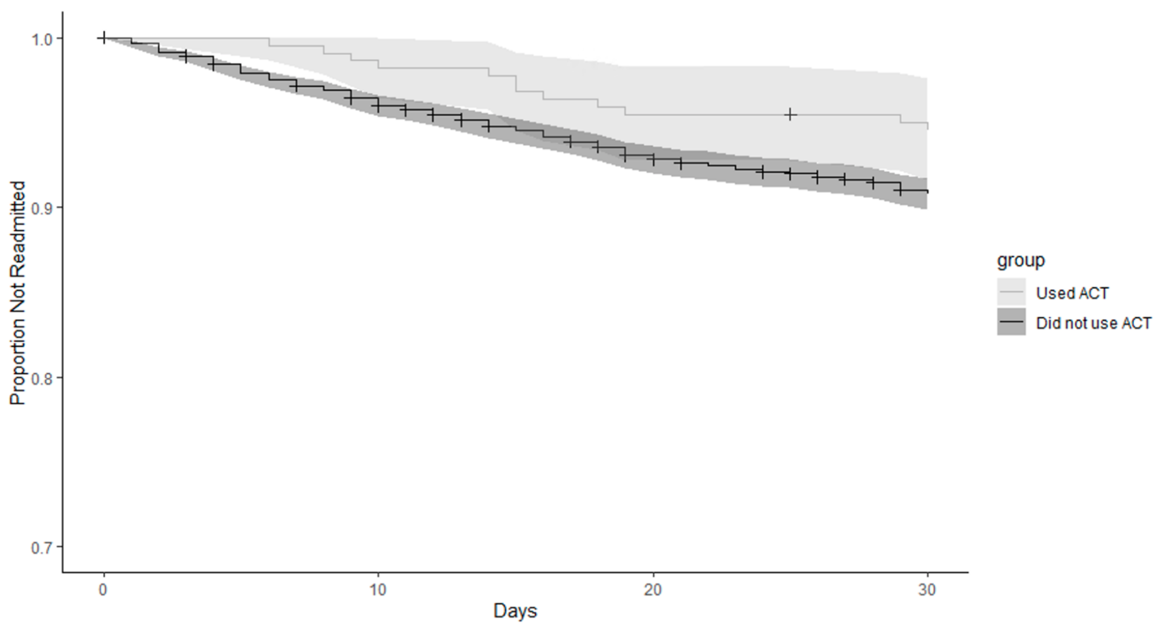

Figure 3.5 Kaplan-Meier curves for time to 30-day same hospital readmission of discharges that subsequently used Assertive Community Treatment (ACT) services between discharge and readmission. The shaded areas show the 95% confidence intervals, and the black crosses show observations that were right-censored.

The multiple logistic regression model confirmed the positive association between high risk of rehospitalization (indicated by the rehospitalization clinical assessment protocol = 2, embedded in the RAI-MH) and 30-day same hospital readmission (Table 3.4). The use of ACT services was negatively associated with readmission on the full sample and non-censored sub-sample. The sub-samples with only one assessment per person had the estimated odd ratios for use of ACT services < 1, but the estimated 95% CI were not statistically significant.

Table 3.4 Adjusted odds ratios in a multiple logistic regression model of 30-day same hospital readmission, their corresponding 95% confidence intervals, and p-values < 0.05 indicated by *

	Full Sample (True = 372, False = 3786)	Not Censored (True = 370, False = 3723)	First and Single Admissions (True = 369, False = 3017)	Last and Single Admissions (True = 190, False = 3196)
Rehospitalization Clinical Assessment Protocol (Ref=0)	-	-	-	-
1	0.91 [0.70-1.17]	0.90 [0.69-1.16]	0.96 [0.74-1.24]	0.83 [0.56-1.20]
2	1.49 [1.11-1.97]*	1.47 [1.09-1.95]*	1.71 [1.27-2.28]*	2.14 [1.49-3.04]*
ACT Services within 30 Days of Discharge or until readmission (True/False)	0.53 [0.28-0.93]*	0.53 [0.25-0.89]*	0.61 [0.31-1.09]	0.70 [0.32-1.31]
AUROC	0.55	0.55	0.54	0.58

3.3.3 Modeling of Community Mental Health Service Use Post-Discharge

3.3.3.1 Modeling Usage of Community Mental Health Services Post-Discharge

Overall, the models suggested that age, mood disorder, severe functional impairment indicated by IADL ≥ 10, and severe impairment in cognitive ability indicated by CPS ≥ 3 were negatively associated with usage of community services within 180 days post-discharge (Table 3.5). On the other hand, being female, schizophrenia, substance use disorders, severe depression indicated by DSI ≥ 8, and high risk of self-harm indicated by SOS ≥ 5 showed positive associations with usage of community mental health services. For all of the scales, the associations to the outcome was statistically significant at the high-range of the scales, and not observed at lower ranges of these scales.

Table 3.5 Adjusted odds ratios in a multiple logistic regression models of community mental health service usage provided by the community mental health agency within 180 days post-discharge, and their corresponding 95% confidence intervals and p-values < 0.05 indicated by *

	Full Sample (True = 1955, False = 2203)	Not Censored Admissions (True = 1872, False = 2118)	First and Single Admissions (True = 1556, False = 1830)	Last and Single Admissions (True = 1556, False = 1830)
Age (ref 18-24)	-	-	-	-
25-44	0.80 [0.68-0.94] *	0.80 [0.68-0.94] *	0.79 [0.66-0.95] *	0.82 [0.68-0.98] *
45-64	0.59 [0.50-0.71] *	0.60 [0.50-0.71] *	0.61 [0.50-0.74] *	0.61 [0.50-0.74] *
≥65	0.25 [0.18-0.35] *	0.26 [0.18-0.36] *	0.28 [0.19-0.39] *	0.24 [0.17-0.35] *
Female	1.17 [1.02-1.33] *	1.16 [1.02-1.33] *	1.17 [1.02-1.36] *	1.14 [0.99-1.32]
Schizophrenia	1.25 [1.07-1.46] *	1.24 [1.06-1.45] *	1.31 [1.10-1.56] *	1.22 [1.03-1.45] *
Mood Disorders	0.80 [0.69-0.94] *	0.77 [0.66-0.90] *	0.84 [0.71-0.99] *	0.80 [0.68-0.95] *
Substance Use Disorders	1.43 [1.23-1.66] *	1.42 [1.21-1.65] *	1.50 [1.28-1.78] *	1.47 [1.24-1.74] *
IADL (ref 1-3)	-	-	-	-
1-3	0.88 [0.64-1.20]	0.89 [0.65-1.22]	0.97 [0.69-1.37]	0.94 [0.66-1.35]
4-9	1.01 [0.72-1.42]	0.99 [0.70-1.40]	1.03 [0.71-1.50]	1.24 [0.85-1.80]
10-18	0.63 [0.41-0.96] *	0.64 [0.41-0.98] *	0.60 [0.36-0.98] *	0.79 [0.49-1.29]
19-30	0.35 [0.20-0.61] *	0.37 [0.21-0.64] *	0.35 [0.17-0.67] *	0.34 [0.18-0.62] *
CPS (ref 0)	-	-	-	-
1-2	0.96 [0.73-1.25]	1.02 [0.78-1.33]	1.00 [0.75-1.34]	0.91 [0.66-1.24]
3-6	0.46 [0.20-0.98] *	0.38 [0.16-0.86] *	0.35 [0.12-0.87] *	0.52 [0.21-1.15]
DSI (ref 0)	-	-	-	-
1-3	1.04 [0.90-1.20]	1.03 [0.89-1.19]	1.03 [0.88-1.21]	0.96 [0.81-1.12]
4-7	1.14 [0.91-1.42]	1.11 [0.88-1.40]	1.03 [0.80-1.32]	1.09 [0.85-1.40]
8-15	1.85 [1.18-2.94] *	1.75 [1.11-2.80] *	1.63 [1.01-2.65] *	2.14 [1.29-3.62] *
SOS (ref 0)	-	-	-	-
4	1.16 [0.89-1.52]	1.20 [0.91-1.59]	1.09 [0.82-1.46]	1.28 [0.95-1.73]
5-6	1.19 [1.01-1.41] *	1.20 [1.01-1.42] *	1.14 [0.95-1.37]	1.22 [1.01-1.46] *
AUROC	0.64	0.64	0.64	0.64

Abbreviations: CPS, Cognitive Performance Scale; DSI, Depressive Severity Index; IADL, Instrumental Activities of Daily Living Scale; SOS, Severity of Self-Harm

3.3.3.2 Modeling High Usage of Community Mental Health Services Post-Discharge

The mean staff time usage during the 180 days post-discharge was 7.2 hours (min <1 hour, median = 1.9 hour, max = 95.7 hours). The top 20 percentile consumed > 11.4 hours, and top 10 percentile consumed > 22.9 hours.

Age, risk of harm to others indicated by the RHO scale, and 30-day readmission were negatively associated with high usage of community services (Table 3.6). However, the RHO's association was only statistically significant for 3 out of 4 subsamples for the top 20 percentile, and not for the top 10 percentile of high usage. The readmission only showed a statistically significant association for the top 20 percentile because there was no readmission event observed for the top 10 percentile subsamples. Schizophrenia and length of the inpatient episode stay were positively associated with high usage of community mental health services.

Table 3.6 Adjusted odds ratios of the variables in a multiple logistic regression model of high usage of community mental health services provided by the community mental health agency within 180 days post-discharge, and their corresponding 95% confidence intervals and p-values < 0.05 indicated by *

	Top 20 Percentile Full Sample (True = 396, False = 1916)	Top 20 Percentile Last and Single Admissions (True = 377, False = 1852)	Top 20 Percentile First and Single Admissions (True = 279, False = 1559)	Top 20 Percentile Last and Single Admissions (True = 342, False = 1531)
Age (ref 18-24)	-	-	-	-
25-44	0.74 [0.57-0.96] *	0.75 [0.58-0.97] *	0.73 [0.54-0.98] *	0.67 [0.51-0.89] *
45-64	0.47 [0.34-0.63] *	0.50 [0.35-0.69] *	0.46 [0.31-0.66] *	0.46 [0.33-0.64] *
$\geq$ 65	0.19 [0.06-0.49] *	0.21 [0.06-0.53] *	0.20 [0.05-0.58] *	0.09 [0.02-0.32] *
Schizophrenia	1.55 [1.24-1.95] *	1.58 [1.26-2.00] *	1.56 [1.20-2.04] *	1.70 [1.33-2.18] *
RHO (ref 0-2)	-	-	-	-
3-4	0.70 [0.41-1.13]	0.66 [0.38-1.09]	0.72 [0.38-1.26]	0.78 [0.43-1.32]
5-6	0.29 [0.09-0.73] *	0.30 [0.09-0.75] *	0.41 [0.12-1.03]	0.27 [0.06-0.75] *
Inpatient Length of Stay (months)	1.11 [1.05-1.18] *	1.11 [1.04-1.18] *	1.11 [1.04-1.19] *	1.12 [1.05-1.20] *
30-day Readmission	0.11 [0.03-0.26] *	0.11 [0.03-0.27] *	0.12 [0.04-0.30] *	0.16 [0.04-0.43] *
AUROC	0.66	0.65	0.67	0.66
	Top 10 Percentile Full Sample (True = 199, False = 2113)	Top 10 Percentile Not Censored Admissions (True = 195, False = 2034)	Top 10 Percentile First and Single Admissions (True = 135, False = 1703)	Top 10 Percentile Last and Single Admissions (True = 168, False = 1705)
Age (ref 18-24)	-	-	-	-
25-44	0.72 [0.51-1.01] *	0.72 [0.52-1.02]	0.81 [0.54-1.21]	0.63 [0.43-0.91] *
45-64	0.43 [0.28-0.66] *	0.43 [0.27-0.65] *	0.49 [0.29-0.81] *	0.44 [0.28-0.69] *
$\geq$65	0.09 [0.01-0.45] *	0.10 [0.01-0.47] *	0.00 [0.00-0.00]	0.10 [0.01-0.49] *
Schizophrenia	1.74 [1.28-2.36] *	1.74 [1.28-2.37] *	1.69 [1.18-2.45] *	1.86 [1.34-2.60] *
RHO (ref 0-2)	-	-	-	-
3-4	0.68 [0.32-1.27]	0.62 [0.28-1.19]	0.82 [0.35-1.67]	0.83 [0.38-1.63]
5-6	0.31 [0.05-1.02]	0.31 [0.05-1.01]	0.47 [0.08-1.56]	0.20 [0.01-0.93]
Inpatient Length of Stay (months)	1.15 [1.07-1.23] *	1.14 [1.07-1.23] *	1.17 [1.08-1.26] *	1.14 [1.06-1.22] *
30-day Readmission	0.00 [0.00-0.00]	0.00 [0.00-0.00]	0.00 [0.00-0.00]	0.00 [0.00-0.66]
AUROC	0.68	0.68	0.69	0.67

Abbreviations: RHO, Risk of Harm to Others

3.3.3.3 Modeling of Outcomes Related to Risk of Harm to Others and Risk of Self-harm

One surprising result was that risk of harm to others (indicated by the RHO scale) did not show a positive association with usage or high usage of community mental health services as normally expected (Table 3.5 and 3.6). In fact, risk of harm to others appeared to be negatively associated with the likelihood of being high resource users (Table 3.6). Additionally, the SOS scale was not observed to be a statistically significant covariate in modeling of high usage. However, when changing the outcome to be enrollment to specialized services within the community mental health agency for those with risk of harm to others (Justice System programs) or risk of self-harm (Skills for Safer Living Support group therapy) among those that received services, the RHO and SOS showed the expected positive association with enrollment in those programs (Tables 3.7, 3.8).

For enrollment in the Justice System programs, the RHO scale at elevated level showed the expected positive association with enrollment in these specialized programs (Table 3.7), along with schizophrenia and substance use disorders. On the other hand, mood disorders showed negative associations with these programs.

For enrollment in the self-harm reduction group therapy (Skills for Safer Living program), the SOS scale also showed an expected positive association with enrollment in this program at both mid and high-level of the scale, along with being female, elevated level of depression indicated by the DSI (Table 3.8). Age, schizophrenia, substance use disorders, and 30-day readmission showed negative association with this program.

Table 3.7 Adjusted odds ratios of the variables in a multiple logistic regression model of enrollment in specialized programs for risk of harming others, and their corresponding 95% confidence intervals and p-values < 0.05 indicated by *

	Justice System Programs Full Sample (True = 358, False = 3800)	Justice System Programs Not Censored (True = 351, False = 3639)	Justice System Programs First and Single Admissions (True = 299, False = 3087)	Justice System Programs Last and Single Admissions (True = 273, False = 3113)
Schizophrenia	1.34 [1.05-1.71] *	1.33 [1.04-1.71] *	1.39 [1.07-1.82] *	1.37 [1.04-1.80] *
Mood Disorders	0.65 [0.48-0.86] *	0.64 [0.47-0.85] *	0.69 [0.50-0.93] *	0.60 [0.43-0.84] *
Substance Use Disorders	1.63 [1.28-2.05] *	1.59 [1.25-2.01] *	1.67 [1.28-2.15] *	1.53 [1.17-2.00] *
RHO (ref 0-2)	-	-	-	-
3-4	1.16 [0.74-1.75]	1.09 [0.68-1.67]	1.16 [0.70-1.83]	1.26 [0.75-2.01]
5-6	2.14 [1.26-3.46] *	2.06 [1.20-3.38] *	2.42 [1.39-4.01] *	2.46 [1.35-4.21] *
AUROC	0.62	0.61	0.62	0.62

Abbreviations: RHO, Risk of Harm to Others

Table 3.8 Adjusted odds ratios of the variables in a multiple logistic regression model of enrollment in specialized programs for risk of self-harm, and their corresponding 95% confidence intervals and p-values < 0.05 indicated by *

	Self-harm Reduction Group Therapy Full Sample (True = 139, False = 4019)	Self-harm Reduction Group Therapy Not Censored (True = 136, False = 3854)	Self-harm Reduction Group Therapy First and Single Admissions (True = 101, False = 3285)	Self-harm Reduction Group Therapy Last and Single Admissions (True = 128, False = 3258)
Age (ref 18-24)	-	-	-	-
25-44	0.81 [0.52-1.24]	0.85 [0.55-1.32]	0.81 [0.49-1.34]	0.85 [0.54-1.34]
45-64	0.84 [0.54-1.3]	0.86 [0.55-1.35]	0.78 [0.46-1.31]	0.94 [0.59-1.49]
≥65	0.10 [0.02-0.34] *	0.11 [0.02-0.36] *	0.14 [0.02-0.47] *	0.12 [0.02-0.38] *
Female	2.61 [1.74-4] *	2.78 [1.84-4.31] *	2.48 [1.55-4.11] *	2.41 [1.59-3.73] *
Schizophrenia	0.20 [0.1-0.35] *	0.17 [0.08-0.32] *	0.07 [0.02-0.2] *	0.23 [0.11-0.42] *
Substance Use Disorders	0.55 [0.33-0.87] *	0.53 [0.31-0.84] *	0.54 [0.3-0.92] *	0.53 [0.31-0.86] *
DSI (ref 0)	-	-	-	-
1-3	1.22 [0.8-1.83]	1.28 [0.84-1.92]	1.32 [0.81-2.1]	1.23 [0.8-1.88]
4-7	1.94 [1.18-3.12] *	2.11 [1.27-3.41] *	1.80 [0.99-3.15] *	1.82 [1.07-3] *
SOS (ref 0-3)	-	-	-	-
4	2.44 [1.37-4.19] *	2.33 [1.29-4.05] *	2.68 [1.4-4.91] *	2.55 [1.38-4.51] *
5-6	2.67 [1.82-3.92] *	2.53 [1.71-3.73] *	2.66 [1.69-4.19] *	2.90 [1.94-4.34] *
30-day Readmission	0.22 [0.05-0.58] *	0.21 [0.05-0.57] *	0.25 [0.06-0.68] *	0.39 [0.09-1.08]
AUROC	0.81	0.81	0.83	0.8

Abbreviations: DSI, Depressive Severity Index; SOS, Severity of Self-Harm

3.4 Discussion

By using routinely collected data, this study was able to provide insights on the transition from inpatient psychiatry to community mental health in Waterloo-Wellington region, Ontario. The results suggested that services appear to be appropriately targeted to those with greater needs. Specifically, usage and high usage were associated with complex conditions like schizophrenia (Table 3.6). Specialized programs were more likely provided to those at risk of harming others and self-harm (Tables 3.7, 3.8). The use of ACT services was also associated with lowered 30-day same hospital readmission rate, and the rate was also lower than that for the rest of the province during the same period (Tables 3.2, 3.4).

About half of discharges from inpatient psychiatry would subsequently use publicly funded mental health services in the community (Figure 3.4). The results suggested that typical usage level for those with mild or moderate conditions is about one hour of in-person service per week for about seven to eight weeks.

This study only focused on examining the expenditures of services provided by the community mental health agency, and not services provided by primary care services, such as family physicians. Additionally, CMHA-WW does not provide addiction services, therefore the actual usage was likely underestimated for this sample. While increased complexity appeared to be associated with subsequent usage of community mental health services post-discharge, factors associated with aging such as cognitive decline and independent living appeared to be associated with reduced usage (Table 3.5), possibly due to the exclusion of joint senior mental health programs of our study.

Most of the services were initiated with the community mental health agency shortly after discharge. However, the time to service initiation should not be interpreted as wait time or time spent on a waitlist. The data obtained for this study only contained when and what services were provided. Therefore, this study could not identify how long the wait time was, if any.

Since the readmission to inpatient psychiatry would reduce the time spent in the community, it was expected and observed in this study that readmission was associated with lowered usage of community mental health services (Table 3.4 and 3.8). By extension, the results also confirmed the validity of the RAI-MH rehospitalization clinical assessment protocol in predicting inpatient psychiatric readmission [101]. Although this study could only provide the temporal association and not causality between readmission and usage of ACT, the results are aligned with randomized trials of ACT in reducing hospital use by providing intensive care in the community [104]. There is also evidence to suggest that readmission does not only reflect the quality of inpatient care, but also the continuing care post-discharge in other parts of the health care system [101, 108–110]. Therefore, it was possible the non-usage of community mental health services post-discharge may also contribute to worsening of health conditions or unmet needs, which increased the likelihood of requiring more intensive care from the inpatient services.

Sensitivity of censoring and repeated assessments were examined. Overall, in all of the models, the directions of the associations indicated by the odd ratios were consistent across subsamples. The statistical significance indicated by the p-values showed minor sensitivity across some subsamples, which was likely due to power of different sample sizes when considering the consistency of the associations indicated by the confidence intervals.

This study highlighted the utility of the RAI-MH assessment beyond the point of care as it was able to explain the patterns of usage of community mental health services, despite the fact that the clinical data were collected by an external organization to the resource use data. Specifically, indicators of clinical complexities such as schizophrenia, substance use, depression, self-harm were associated with usage of community mental health services post-discharge (Table 3.5). Similarly, high and specialized service usage were closely linked with other clinical complexity, such as risk of harm to others, and length of inpatient psychiatric episode (Table 3.6).

Previous research suggested that information sharing across care settings could improve quality of care, and reduce confusion about care plan for patients and their care givers [111, 112]. If the assessment is made available across the health care continuum, the discharge assessment from one setting can aid the care planning in the next care setting. Information from the RAI-MH such as the diagnosis and clinical scales could be potentially be used as the basis for service assignment by the community mental health agency, because they closely mirrored how the community mental health agency allocates services. For example, the RAI-MH discharge assessment showed that persons at higher risk of harm to others were more likely to be allocated services specific to the justice system specialized programs, and persons at risk of self-harm were more likely to be allocated services specific to self-harm reduction from the community mental health agency (Tables 3.7, 3.8).

This study was not without limitations. First, the data from the community mental health agency capture the majority of adult community mental health services but it was not inclusive of all community mental health services and primary care, such as addiction services provided by other local agencies, children services, physicians fee-for-service services provided by the province, senior mental health services provided by joint programs with other organizations, and possible unobserved service use by non-local residents discharged from GRH. One of the possible directions for future research is to recruit additional agencies that provide community mental health services, such as addiction services, to get a fuller picture of health care service usage patterns in the community. Second, the outcomes of this study were within 180 days of discharge. While this is a long follow-up window, it was unable to take into account factors that could shift the window such as wait time, therefore it may not closely match an actual episode of care. Future research should explore other possible methods of constructing an episode of care in the community settings, and incorporate additional data sources to examine complexities related to wait time. Third, the outcome measures were based on direct staff time, which did not take into account the variation in individual staff's compensations. In follow-up studies, we aim to incorporate compensation data to enhance the outcome measures. Lastly, the sample of this study deviated in both demographic and clinical characteristics from the rest of the province (Table 3.2), which further supported the need for an expanded study in the future to obtain a fuller picture.

This study offered a glance at the transition between inpatient psychiatry and community mental health services in Waterloo-Wellington region, ON and showed that the clinical complexity, indicated by the RAI-MH assessment of patients at discharge, and service use patterns were closely linked. The use of community mental health services was associated with better outcome post-discharge (indicated by 30-day readmission) and could potentially reduce the demand on intensive services from the inpatient settings.

Chapter 4
Study Two: Evaluating Existing Case-Mix Classification Systems for Use in Community Mental Health

Abstract

Case-mix classification systems play an important role in funding formulas. As mental health care has been shifting from institutionalization to community-based care, there is an immediate need for a case-mix classification system specific to community mental health care in order to support its prospective funding. Since developing case-mix classification is not a trivial task, this study examined with two existing case-mix classification systems, the System for Classification of In-Patient Psychiatry (SCIPP) and Australian Mental Health Care Classification (AMHCC), for their potential utility in community mental health care. The two systems were adapted to predict community mental health resource use from the clinical characteristics measured at discharge from inpatient psychiatry. Different approaches in constructing a community episode of care that were examined. An episode of 90 days from the first contact with the community mental health agency appeared to be most responsive to indicators of clinical needs, and most practical for implementation.

The results showed that the SCIPP was able to explain 6% [95% CI: 3.7, 8.9] of variance in resource use of a community mental health agency in Ontario, Canada. However, whether an individual had visited the community mental health agency within 30 days prior to the inpatient episode was an important factor in predicting resource use. Incorporating prior contact improved the explained variance to 11.9% [95% CI: 8.7, 15.5] of resource use and 14.1% [95% CI: 10.6, 17.9%] of log-transformed resource use. The SCIPP showed to be fair and equitable across its sub-groups. The AMHCC was not immediately usable beyond the context of Australia due to inoperationalizable concept of phases of care, which requires a clinical decision that is not possible to derive from a clinical assessment. The simplified AMHCC using only age, and two clinical assessments achieved only 1.2% of explained variance in the context of this study. Compared to using diagnosis alone, SCIPP was observed to be more predictive of community mental health resource use.

4.1 Introduction

Since 2012, the province of Ontario has implemented a new health care funding formula, which a portion of the total funding to hospitals is driven by the type of patient [26]. The underlying assumption is that there are shared characteristics among individuals that drive the usage of health care services. A population can then be divided into groups of people with similar characteristics that are expected to consume similar levels of health care services. These "case-mix" groups represent the types of cases that are observed to utilize different levels of resources within a health care system [23]. There is a tentative plan to expand case-mix funding to community mental health services [26].

Currently, in Ontario, funding for community mental health services is still based on global budget. In order to support the expansion of the new funding approach outlined by the province, a case-mix classification system for community mental health is needed to support the transition. Most of the research on case-mix classification for mental health care has primarily focused on the inpatient settings [113]. In order to support continuity of care across settings and fast-tracking the development of a new case-mix classification system, it may be advantageous to determine if existing systems can be used in different settings than the ones in which they were developed [113]. There are two possible candidates systems that both achieved high explained variance of resource use in their development studies of about 26% for both, the System for Classification of Inpatient Psychiatry (SCIPP) [27, 52] and Australian Mental Health Care Classification (AMHCC) [46].

Using resource use data from a community mental health agency in Waterloo-Wellington region of Ontario, this study examined how well the SCIPP and AMHCC perform as a predictor of resource use in a less intensive care settings. The results of this study are expected to contribute to the gap in community mental health case-mix research, and the ongoing development of funding formulas for the Ontario mental health care system.

4.1.1 System for Classification of Inpatient Psychiatry

SCIPP was developed specifically for inpatient psychiatry settings. SCIPP differentiates 47 levels of resource use (Figure 4.1), measured in per-diem wage-weighted staff time, based on clinical profile, measured by the RAI-MH assessment [95]. The development study of SCIPP observed 2000 inpatients in 34 psychiatric units. The RAI-MH is a mandated comprehensive assessment for all inpatient psychiatric beds in Ontario, containing over 300 clinical measurements and risk scales derived from the clinical measurements [95]. Each group has a weight attached, also known as the case-mix index (CMI) (Figure 4.1). The value of the case-mix index represents the resource intensity of a group relative to a referenced inpatient, derived during the development of SCIPP.

4.1.2 Australian Mental Health Care Classification

The AMHCC was developed using a sample of 9,976 community episodes [46]. The AMHCC requires age, the Health of the Nation [55], and Life Skills Profile-16 assessments [46, 67], as well as a clinical decision of the phase of care. The phase of care for the AMHCC is defined as the primary goal of care identified prospectively by a clinician engaging with the client to create their care plan and prior to providing care, which can be one of: acute (short term reduction in symptom severity and/or personal distress), functional gain (improve personal, social, and occupational functioning, or promote psychosocial adaptation), intensive extended (prevention or minimization of further deterioration), consolidating gain (maintain or improve functioning and/or prevent relapse), and assessment only (information gathering only) [46]. More detailed guidance on classification of the phase of care are provided by the Independent Hospital Pricing Authority of Australia [114].

4.2 Methods

4.2.1 Overview

To operationalize the examination of the two case-mix classification systems, some modification were required to apply these systems in the context of this study's data. The SCIPP requires the Resident Assessment Instrument - Mental Health (RAI-MH) assessment as input for its algorithm. Since the RAI-MH or a compatible assessment from the interRAI suite of mental health instruments [52] was not available from the community mental health settings in Ontario, this study instead used the RAI-MH of patients discharged from inpatient psychiatry and linked that to their usage data of community mental health services post-discharge.

Although the two assessments used by the AMHCC could be crosswalked from the RAI-MH using closely related assessment items, the classification of phase of care appeared to be a decision made by a clinician [114], which cannot be retrospectively derived from the RAI-MH data alone. In absence of the phases of care information, this study could only replicate a simplified version of the AMHCC using age, HoNOS assessment, and LSP-16 assessment.

4.2.2 Clinical Characteristics Data

Discharge RAI-MH assessments were obtained from GRH, a hospital with inpatient psychiatric beds located in the Waterloo Region, Ontario. Their records were linked with service use records of the Canadian Mental Health Association - Waterloo Wellington (CMHA-WW), a local community mental health agency. According to CMHA-WW, their chapter is the largest of the Canadian Mental Health Association in Canada by staff volume and portfolio of services offer covering all ages, which made CMHA-WW a good example to study.

The RAI-MH discharge assessments from the GRH were linked with the community mental health service records using the Ontario health card number and date of birth. If the health card number linkage was not possible, then linkage using full name and date of birth was used. This study included all adults who were assessed with the RAI-MH, discharged from the GRH inpatient psychiatry. Children or people who received children mental health services post-discharge were excluded. Forensic inpatient admissions were excluded because their inpatient stay were often required by law and beyond the scope of our research ethics application. In-hospital mortality, transfer, or same day readmission were also excluded because these discharges would not subsequently use community services. Those who were admitted to inpatient psychiatry for ≤ 72 hours were excluded because they were assessed using a short version of the RAI-MH which only captured the intake administrative items.

The data covered the calendar years between 2014 and 2018. This time frame was chosen because CMHA-WW was established through merging of smaller chapters in 2014 and matched the current organizational structure. The use of record linkage data was approved by the Grand River Hospital (# 2018-0669), the CMHA-WW ethics board in October 2018, and the University of Waterloo (# 40147).

This study also compared the demographics and clinical characteristics of our study sample to the rest of the province of Ontario by obtaining RAI-MH discharge assessments from the Ontario Mental Health Reporting System (OMHRS) for the same time period. Since OMHRS also contained the GRH data and masked the individual hospitals in the dataset, this study excluded all three hospitals in the same region (Waterloo-Wellington) as the GRH to remove the effect of autocorrelation. The use of OMHRS also received ethics clearance from the University of Waterloo (file number 19917).

4.2.3 Resource Use Measure

To capture the actual resource use, this study relied on the service events data available from the community mental health agency. Each service event received by the community mental health agency clients was recorded with a staff's job title and the duration of time spent with client. This study only included the direct services, which were services that were provided directly to the client or by telecommunication. Indirect services that were not driven by the clinical characteristics included all other client specific activities, such as documentation or case review. Other non-client specific activities such as staff training or meetings were not available in the data.

Using the median wage rates provided by the community mental health agency for each job title, this study calculated a wage-weighted staff time cost measure by multiplying the hourly wage of staff by the duration spent. For group services, the staff time was divided equally among the number of clients registered for the group, even if someone did not actually attend because resources were already allocated from the organization's perspective.

These data were generally expected to be of good quality because they were used for scheduling of appointments, monitored by management, and used to determine extra pay for eligible direct face-to-face time. Service events longer than eight hours, length of a typical work shift, were reviewed with an analyst from the community mental health agency to determined whether an error had occurred and corrected if necessary. In many cases, the error was due to data entry of the time stamp by incorrectly selecting AM versus PM. It was also unlikely for a service event to span multiple days, so incorrect date entries were also identified and corrected.

4.2.4 Episode of Care

The first challenge is examining the two case-mix classification systems was to define a community episode of care in order to construct a resource use measure. The concept of an episode of care is commonly used throughout the health care system to quantify the resource use as a result of the care activity provided to an user of health care services [115]. The AMHCC defines an episode of care as a period between initiation of services or transfer from the inpatient setting to the close of the case by the care team or transfer into another setting [116]. Beyond the Australian context, it was not immediately possible to replicate this definition of episode of care using this study's data.

Most of the research on this topic has focused on the inpatient settings, where an episode of care has definitive start and end points that are easily observed. Therefore an episode of care defined by the SCIPP was also not applicable. An episode of care in the community settings is much more complex because the services can be provided for both longer period of time and intermittently instead of continuously like the inpatient settings. As a result, the care received long after the discharge from inpatient psychiatry may not be related to the health care needs at discharge. Therefore, it was important to determine an appropriate observational window for the period of ongoing care post-discharge that reflected the care received in the community driven by the needs at discharge.

Attempts to define an episode of care for community mental health as either an episode of illness or an episode of disease were also problematic. An episode of illness is defined as the period of suffering due to symptoms [115, 117, 118]. This definition introduced additional complexities in defining the symptoms and observing the symptoms in order to construct an episode. An episode of disease is defined as the period from onset of disease to death or a completion, such as a cure [117]. Since the nature of mental health care can be chronic, factors like remission or relapse also add to the complexity of defining the completion of the disease.

A community episode of care should reflect the ongoing nature of care. In other words, the ongoing care does not have to be continuous on a daily basis, but there should not be a significant interruption, which can be defined as a change in care setting [45] or an extended period of inactivity [118]. This study considered readmission to inpatient psychiatry as an indicator of change in care setting. Empirical data were used to determine the length of inactivity that indicates an extended period of inactivity that was representative of the observed data.

Nine different ways of defining an episode of care were examined by considering combinations of three elements of an episode of care: the unit of counting of resource use as the basis for reimbursement, the start, and the end. These approaches can be categorized in two categories: interval basis and episodic basis (Table 4.1).

Table 4.1 Approaches to construct an episode of care

Type	Start	End	Unit of Counting
Interval basis	First contact post-discharge AND within 30 days of discharge	Last service event prior to activity interruption OR change in care setting. Experiments: intervals of $\{1, 7, 14\}$ days	Interval
Episodic basis	Discharge Date	Discharge date + pre-defined fixed follow-up length of $\{30, 60, 90\}$ days	Whole episode
Episodic basis	First contact post-discharge AND within 30 days of discharge	First service event date + pre-defined fixed follow-up length of $\{30, 60, 90\}$ days	Whole episode

The SCIPP was developed to predict resource use on a interval basis [27], which calculated resource use for each day of an inpatient episode. This approach is also referred to as per-diem basis because each unit of counting equals to one day [23]. This study experimented with other interval lengths such as 7 or 14 days (Table 4.1). The start was defined as the first contact post-discharge indicated by the first service event at the community mental health agency, which is a commonly used starting point [115]. An interruption indicated by a change in care setting or extended period of inactivity ended the episode.

Another approach is the episodic basis, which has a fixed start and a fixed follow-up length for everyone. This study experimented with two options for the start: discharge from inpatient psychiatry or the first contact with the community mental health agency post-discharge. Different follow-up lengths that have typically been used for an episodic basis were experimented, such as 30, 60, or 90-day. Although the length and the end of the episode were pre-defined, an episode would still end pre-maturely if there was a change in care setting. However, the activity interruption criterion was not taken into account in this approach because one of the rationales behind this approach was that it removed the complexity of having to define an interruption in terms of time. Additionally, the follow-up length can theoretically be longer, such as 180 or 365 days; however, it would not be very practical for the cash flow if the provider was reimbursed only once or twice a year.

To facilitate comparison across experiments, this analysis restricted the first contact to: within 30 days of discharge, full 90 days observation window, and no readmission within 90 days of discharge, so that all experimental episode constructions contained the same data sample. An appropriate episode of care definition should enable the cost measure to be responsive to the clinical needs and intensity of service usage. Several proxies of clinical needs and service use were used to compare the possible episode of care definitions: primary diagnosis, number of service events, length of previous inpatient stay, number of service days, enrollment (binary) in specialized programs that aim to help individuals with severe and persistent mental health issues like Assertive Community Treatment (ACT).

4.2.5 SCIPP as Predictor of Community Mental Health Services Usage Post-Discharge

In a previous study (Chapter 3), clinical severity were shown to be associated to usage (binary) and high usage (top 10^{th} and 20^{th} percentile) within 180 days of discharge. SCIPP has a set of weights, the case-mix index, attached to each of its terminal nodes that represents the resource intensity of a group relative to other groups. In other words, the case-mix index can be viewed as an indicator of clinical or functional severity. The association between the case-mix index of SCIPP and usage (binary) within 30, 60, 90, and 180 days post-discharge was examined using logistic regression. Additionally, association between SCIPP's CMI and high usage (top 10^{th} and 20^{th} percentiles) within 180 days of discharged was examined using logistic regression. Since some SCIPP groups had small sample sizes, this analysis examined the sensitivity of these associations for groups with at least 10 observations of usage post-discharge.

4.2.6 Explained Variance in Community Mental Health Resource Use of SCIPP

Given a classification system like SCIPP, there are two ways to evaluate the predictive performance. First, the SCIPP can be used as a regression tree, in which the predicted value for each leave is the group mean. This approach is equivalent to a multiple regression model whose predictors are dummy-coded SCIPP group classification. Second, the SCIPP's case-mix index (CMI) can be considered as a risk score, which can then be regressed against the observed costs. The resource use data are often positively skewed, therefore log-transforming the resource use to approximate a normal distribution was also considered.

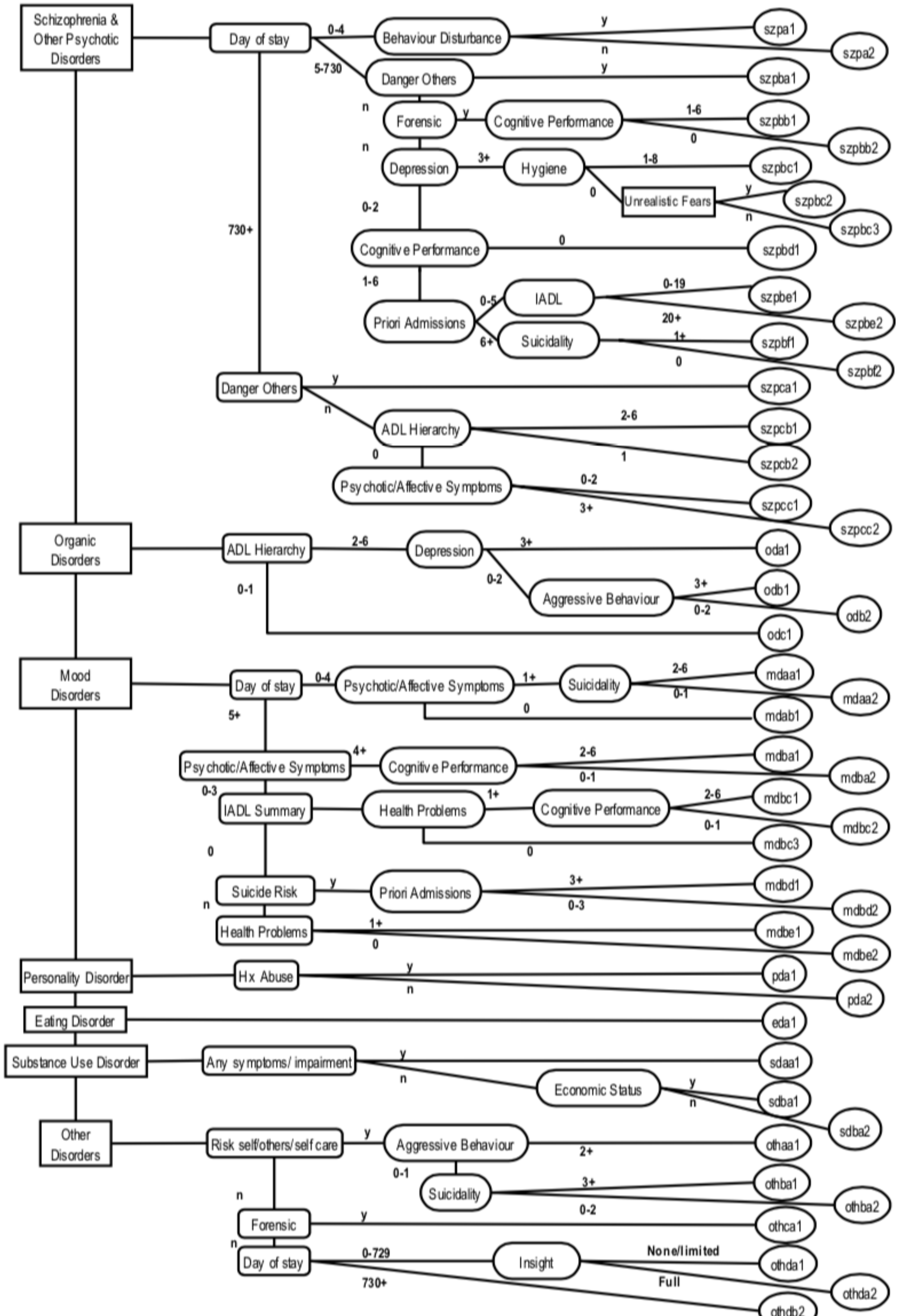

Figure 4.1 System for Classification of Inpatient Psychiatry (SCIPP) [27]

To assess the overall fit of SCIPP, the observed and predicted resource use were used to calculate the *coefficient of determination* (R^2). However, there are increasing concerns that decisions made by automated algorithms can perpetuate existing biases, which results in unfair allocation of health care across ethnicities or sexes [119]. Therefore, this analysis was also stratified by each SCIPP group and sex. The RAI-MH does not contain a variable on race, but contains a variable indicating indigenous status. However, indigenous status is a protected data element that was not available for study under our current ethics application.

In addition to using a global metric like the R^2, several metrics were used at the sub-group level to performed more granular assessment for the most appropriate episode construction method. The *coefficient of variation* is a ratio of the standard deviation to the group mean. It is an indicator of how homogeneous a group is in terms of resource use. Since a classification system assumes that every case in a group should have very similar clinical needs and resource use, the coefficient of variance is expected to be close to zero. In addition, the covariance, $covariance(G, \hat{Y} - Y)$ where $G\{0,1\}$ indicates the membership of group g, between group classification and the residuals was shown to be a proxy for fairness [120]. A negative value indicates that being classified into a particular group resulted in larger under-reimbursement relative to the rest of the population, or unfair in resource allocation relative to other groups.

4.2.7 Effects of Time to Service Initiation and Prior Contact

Currently, the community mental health agency does not use the discharge assessments as the basis for their resource allocation, and instead they may use information from both intake and previous contacts to allocate resources. Therefore, prior contact could potentially influence on the performance of SCIPP because it could contain information not available from the discharge assessment, and by extension not available to SCIPP. This analysis defined prior contact as a contact with the community mental health agency within 30 days prior to the admission to inpatient psychiatry.

In addition, the discharge assessment may be a current snapshot of someone's clinical profile at discharge, but not everyone initiated services with the community mental health agency immediately following discharge. Therefore, time to service initiation was also tested for its potential influence on the performance of SCIPP.

4.2.8 Explained Variance of Community Mental Health Resource Use of the Simplified AMHCC

The main limitation of the data of this study is that it was not possible to derive the phases of care retrospectively based on the RAI-MH assessment. Therefore, a simplified version of the AMHCC was created without the phases of care. This analysis should reflect the potential explained variance of the remaining components of the AMHCC, which are age, HoNOS, and LSP-16 assessments.

For each RAI-MH assessment in the data, the 12 items in the HoNOS, and 16 items in the LSP-16 were derived according to Appendicies A and B. The AMHCC used three versions of the HoNOS for younger than 18, 18-64, and 65 and older groups [114]. The data in this study contained mostly individuals between 18-64 and a small number 65 and older, therefore only the HoNOS and HoNOS 65+ were used.

The sum total of all items in the HoNOS is referred to as the HoNOS complexity and split into high or moderate groups [114]. For the individuals at moderate HoNOS, their LSP-16 items were also summed and split into high and moderate levels [114]. In the AMHCC, the thresholds between high and moderate of the HoNOS and LSP-16 were different for each of the phases of care [114]. In absence of the phases of care, these thresholds were determined empirically by fitting a regression tree model to the data of each age group, and using HoNOS and LSP-16 complexities as predictors (Figure 4.2).

The R^2 was calculated for the HoNOS complexity, LSP-16 complexity, and simplified AMHCC (Figure 4.2) against the observed resource use. The simplified version only contained 6 terminal groups, compared to the possible 46 terminal groups if the five phases of care were available. Since the simplified AMHCC does not entirely reflect the original AMHCC, the analysis of terminal groups was not performed.

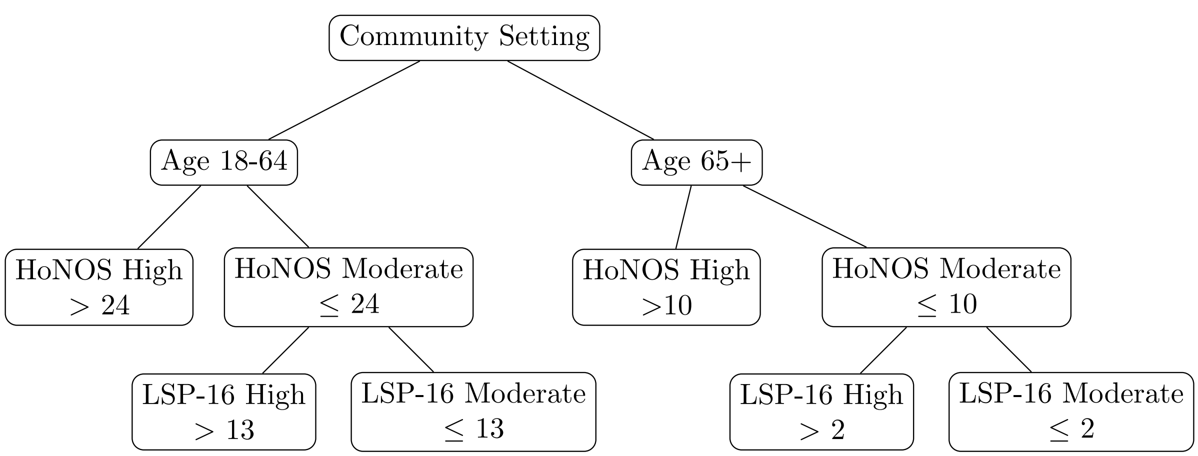

Figure 4.2 Simplified version of the Australian Mental Health Care Classification without phases of care

4.3 Results

This study identified 1,207 discharges that allowed all nine approaches of episode constructions to have the same number of observations and non-censored episodes either due to readmission or date of data extract (Figure 4.3). This sample represents only a small portion of all clients that came into contact with the community mental health agency, which is estimated to be about 10,000 per year. Table 4.2 showed the counts of observations for each SCIPP groups from the sample included in this study, as well as from the original SCIPP development study, the OMHRS, and all discharges from GRH.

The study sample observed 31 out of 47 SCIPP groups. Not all SCIPP groups were observed in this study sample, such as groups of long inpatient episode of $\geq$ 730 days and forensic inpatients (Figure 4.2). Five discharges were not classifiable into a SCIPP group due to missing data and invalid data (Figure 4.3). In addition, groups with organic disorders or neurocognitive disorders were also largely absent from the sample, as they would qualify for senior mental health services offered by inter-agency programs between the community mental health agency and other agencies, which this study did not have complete resource use data.

The distribution of observations per group for the study sample appeared relatively similar to the provincial distribution (Figure 4.2). In contrast, the distribution of observations per group in the original SCIPP study appeared uniform, due to the study's objective to obtain wide range of clinical needs [27].

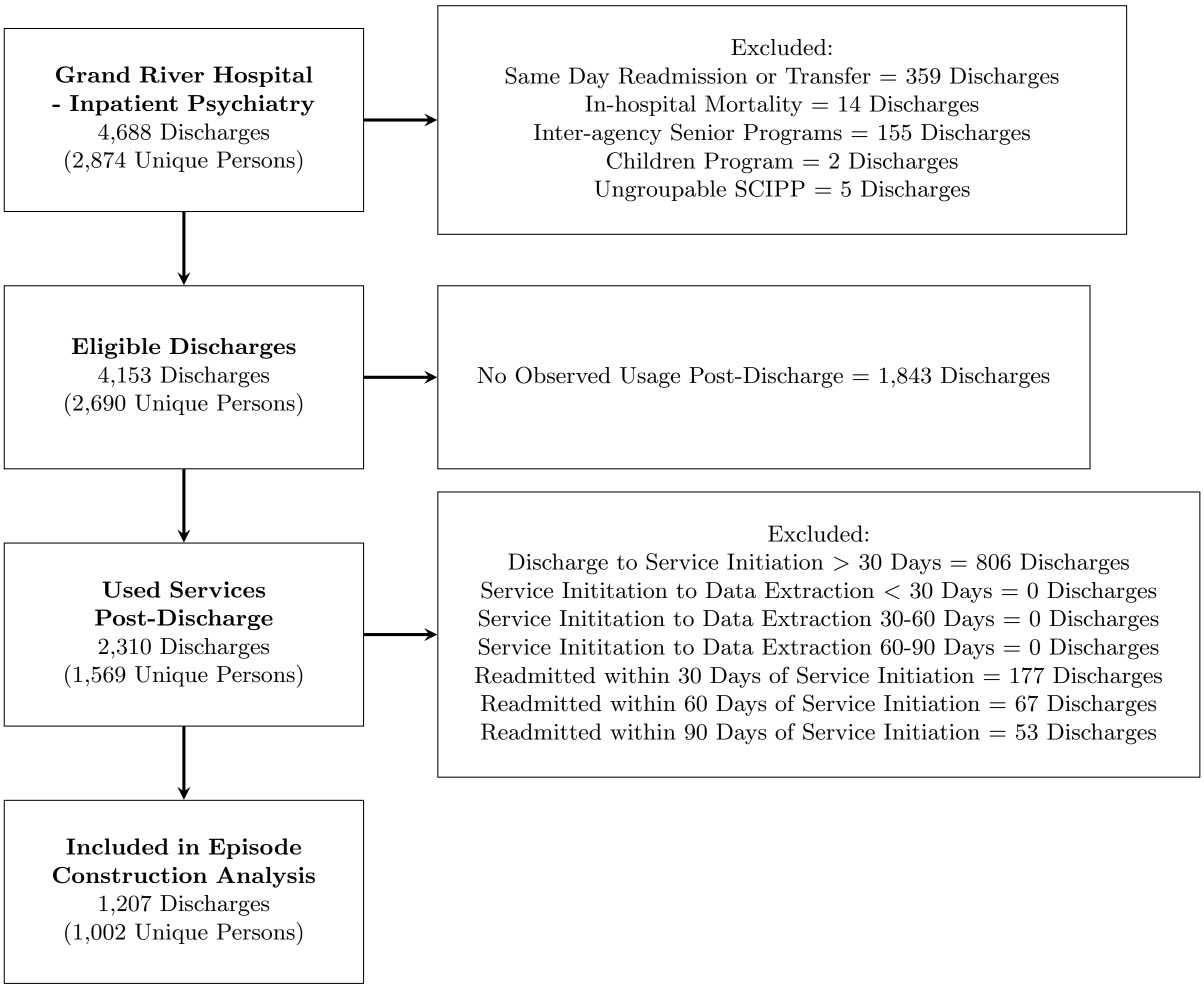

Figure 4.3 Record linkage and sample selection

Table 4.2 Count of observations in each SCIPP groups from the original development study, discharges in OMHRS between 2014-2018, all of GRH discharges, and included in this study

SCIPP Groups	Observations per Group in Original Study	Observations per Group in OMHRS, n=174,338	Observations per Group in all GRH Discharges, n=4,688	Observations per Group included in this Study, n=1,207
SZPA1	34 (1.7%)	1,187 (0.7%)	15 (0.3%)	2 (0.2%)
SZPA2	41 (2.1%)	2,347 (1.3%)	43 (0.9%)	11 (0.9%)
SZPBA1	39 (2.0%)	4,551 (2.6%)	83 (1.8%)	13 (1.1%)
SZPBB1	65 (3.3%)	675 (0.4%)	0 (0%)	0 (0%)
SZPBB2	30 (1.5%)	1,796 (1%)	0 (0%)	0 (0%)
SZPBC1	40 (2.0%)	1,171 (0.7%)	30 (0.6%)	3 (0.2%)
SZPBC2	45 (2.3%)	2,479 (1.4%)	62 (1.3%)	22 (1.8%)
SZPBC3	52 (2.6%)	5,992 (3.4%)	218 (4.7%)	50 (4.1%)
SZPBD1	49 (2.5%)	34,492 (19.8%)	1,315 (28.1%)	388 (32.1%)
SZPBE1	28 (1.4%)	7,968 (4.6%)	103 (2.2%)	28 (2.3%)
SZPBE2	39 (2.0%)	845 (0.5%)	25 (0.5%)	3 (0.2%)
SZPBF1	32 (1.6%)	2,292 (1.3%)	13 (0.3%)	4 (0.3%)
SZPBF2	33 (1.7%)	1,797 (1.0%)	14 (0.3%)	3 (0.2%)
SZPCA1	31 (1.6%)	8 (0%)	2 (0%)	0 (0%)
SZPCB1	46 (2.3%)	22 (0%)	1 (0%)	0 (0%)
SZPCB2	57 (2.9%)	9 (0%)	0 (0%)	0 (0%)
SZPCC1	55 (2.8%)	107 (0.1%)	2 (0%)	0 (0%)
SZPCC2	38 (1.9%)	15 (0%)	0 (0%)	0 (0%)
ODA1	59 (3%)	823 (0.5%)	56 (1.2%)	0 (0%)
ODB1	30 (1.5%)	594 (0.3%)	19 (0.4%)	0 (0%)
ODB2	75 (3.8%)	1,736 (1%)	68 (1.5%)	0 (0%)
ODC1	64 (3.2%)	4,038 (2.3%)	106 (2.3%)	9 (0.7%)
MDAA1	33 (1.7%)	1,701 (1%)	15 (0.3%)	6 (0.5%)
MDAA2	30 (1.5%)	1,273 (0.7%)	16 (0.3%)	3 (0.2%)
MDAB1	44 (2.2%)	2,923 (1.7%)	54 (1.2%)	17 (1.4%)
MDBA1	37 (1.9%)	733 (0.4%)	5 (0.1%)	0 (0%)
MDBA2	37 (1.9%)	3,742 (2.1%)	31 (0.7%)	9 (0.7%)
MDBC1	34 (1.7%)	653 (0.4%)	1 (0%)	0 (0%)
MDBC2	63 (3.2%)	3,387 (1.9%)	20 (0.4%)	1 (0.1%)
MDBC3	59 (3%)	8,541 (4.9%)	96 (2%)	14 (1.2%)
MDBD1	33 (1.7%)	0 (0%)	0 (0%)	0 (0%)
MDBD2	45 (2.3%)	28,253 (16.2%)	614 (13.1%)	156 (12.9%)
MDBE1	76 (3.8%)	4,989 (2.9%)	56 (1.2%)	22 (1.8%)
MDBE2	40 (2%)	17,520 (10%)	777 (16.6%)	191 (15.8%)
PDA1	41 (2.1%)	3,617 (2.1%)	73 (1.6%)	39 (3.2%)
PDA2	40 (2%)	3,729 (2.1%)	193 (4.1%)	69 (5.7%)
EDA1	30 (1.5%)	592 (0.3%)	3 (0.1%)	1 (0.1%)
SDAA1	34 (1.7%)	6,987 (4%)	144 (3.1%)	48 (4%)
SDBA1	30 (1.5%)	1,561 (0.9%)	56 (1.2%)	16 (1.3%)
SDBA2	33 (1.7%)	2,090 (1.2%)	68 (1.5%)	20 (1.7%)
OTHAA1	31 (1.6%)	667 (0.4%)	18 (0.4%)	3 (0.2%)
OTHBA1	32 (1.6%)	2,982 (1.7%)	121 (2.6%)	28 (2.3%)
OTHBA2	39 (2%)	469 (0.3%)	6 (0.1%)	0 (0%)
OTHCA1	57 (2.9%)	42 (0%)	0 (0%)	0 (0%)
OTHDA1	41 (2.1%)	1,357 (0.8%)	108 (2.3%)	24 (2%)
OTHDA2	39 (2%)	1,001 (0.6%)	22 (0.5%)	4 (0.3%)
OTHDB2	38 (1.9%)	2 (0%)	0 (0%)	0 (0%)

4.3.1 SCIPP as Predictor of Community Mental Health Services Usage and High Usage Post-Discharge

The sample was likely under-powered to allow the statistical significant detection of association between the relative resource intensity of the inpatient setting (indicated by the SCIPP CMI) and usage of community mental health services post-discharge. This may be because some SCIPP groups had small observed sample sizes (Tables 4.2, 4.3, 4.4).

When focusing the analysis on groups that had at least 10 observations of usage post-discharge, a positive association between the SCIPP CMI and usage of community mental health services within 180 days was observed (Table 4.4). For shorter follow-up periods, from 30-90 days, the odd ratios of SCIPP CMI overlapped 1.00.

Table 4.3 Odds ratios in a logistic regression model of usage post-discharge within different time windows and sample selections, their corresponding 95% confidence intervals, and p-values < 0.05 indicated by *

	Usage within 30 days (True = 1504, False = 2649)	Usage within 60 days (True = 1672, False = 2481)	Usage within 90 days (True = 1787, False = 2366)	Usage within 180 days (True = 1953, False = 2200)
SCIPP CMI	0.9 [0.71-1.14]	0.86 [0.68-1.08]	0.87 [0.69-1.10]	0.9 [0.72-1.13]
AUROC	0.51	0.51	0.5	0.5

Table 4.4 Odds ratios in a logistic regression model of usage post-discharge within different time windows with ≥ 10 observations of usage per SCIPP group, their corresponding 95% confidence intervals, and p-values < 0.05 indicated by *

	Usage within 30 days, observed usage per group $\geq$ 10 (True = 1463, False = 2446)	Usage within 60 days, observed usage per group $\geq$ 10 (True = 1630, False = 2279)	Usage within 90 days, observed usage per group $\geq$ 10 (True = 1740, False = 2169)	Usage within 180 days, observed usage per group $\geq$ 10 (True = 1896, False = 2013)
SCIPP CMI	1.27 [0.96-1.67]	1.28 [0.97-1.68]	1.3 [0.99-1.72]	1.39 [1.06-1.83]*
AUROC	0.49	0.52	0.52	0.52

An additional logistic regression model also showed that there was no statistically significant association detected between the SCIPP's CMI and high usage within 180 days of discharge (Table 4.5). Models of SCIPP's CMI as predictors of usage and high use of community mental health service usage post-discharge both showed low discrimination, indicated by the low AUROC.

Table 4.5 Odds ratios in a logistic regression models of high usage within 180 days post-discharge among those with observed usage post-discharge with different sample selections, their corresponding 95% confidence intervals, and p-values < 0.05 indicated by *

	Top 20^{th} Percentile (True = 462, False = 1848)	Top 20^{th} Percentile, group size ≥ 10 (True = 453, False = 1785)	Top 10^{th} Percentile (True = 231, False = 2079)	Top 10^{th} Percentile, group size ≥ 10 (True = 225, False = 2013)
SCIPP CMI	0.78 [0.51-1.16]	0.80 [0.50-1.27]	0.86 [0.49-1.45]	0.86 [0.46-1.56]
AUROC	0.52	0.52	0.51	0.51

4.3.2 Episode of Care

For the interval basis approach in constructing an episode, this study first determined a sensible definition of activity interruption by calculating the gaps between all two consecutive service events. The majority of the gaps (95%) were < 28 days (Q1 = 0 day, Q2 = 2 days, mean = 9.9 days, Q3 = 7 days). Based on this result, an interruption of at least 4 weeks (28 days) triggered the end of an interval-based episode. The ongoing community care post-discharge had a mean of 16.5 service days [95% CI: 14.6-18.4] stretching over mean of 82.9 days [95% CI: 74.3-91.5].

Restricting the length of the observation window using the episode-based approach expectedly captured fewer services days, ranging from means of 3.7 - 8.0 services days (Table 4.6). Additionally, it appeared that about half of the service days within the 90-day periods were received in the first 30 days, indicated by the mean number of service days of episodes of 30-day versus episodes of 90-day (Table 4.6). About half of the resource use was also consumed within the first 30 days as well. This is consistent with receiving an initial intake assessment, which consumed a large portion of staff time up-front; however, service provision tapered off over time.

Table 4.6 Summary statistics of alternative episode construction approaches

	Days with Service mean [95% CI]	Mean Resource Use ($) [95% CI]	Median Resource Use ($) [Q1,Q3]	Coefficient of Variation in Resource Use ($)
Interval 1-day	16.5 [14.6, 18.4]	28.9 [28.3, 29.4]	19.2 [7.8, 37.9]	1.38
Interval 7-day	16.5 [14.7, 18.4]	38.4 [37.5, 39.3]	25.1 [5.0, 52.5]	1.50
Interval 14-day	16.5 [14.7, 18.4]	73.1 [71.1, 75.0]	49.8 [19.4, 97.0]	1.20
Discharge + 30 days	3.7 [3.5, 3.9]	93.9 [85.9, 101.9]	40.9 [14.0, 124.6]	1.50
Discharge + 60 days	5.9 [5.6, 6.2]	153.5 [141.1, 165.9]	59.7 [17.4, 221.7]	1.43
Discharge + 90 days	7.8 [7.4, 8.3]	211.8 [194.0, 229.5]	72.5 [19.6, 294.2]	1.49
1st Contact + 30 days	4.1 [3.9, 4.2]	102.5 [94.1, 110.9]	46.6 [15.5, 141.5]	1.45
1st Contact + 60 days	6.2 [5.9, 6.5]	161.1 [148.2, 174.0]	62.8 [17.9, 232.2]	1.41
1st Contact + 90 days	8.0 [7.6, 8.5]	217.2 [199.2, 235.2]	74.4 [19.9, 309.4]	1.47

The examination of experimental episode constructions and their correlation with indicators of clinical needs and service use showed that overall the correlations were weaker for the log-transformed resource use (Table 4.7).

The interval basis approach (episodes of 1-, 7-, 14-day intervals) appeared to be less correlated with indicators of clinical needs, such as: diagnosis of schizophrenia, major diagnoses, and usage of ACT program, than the episodic basis approach (Table 4.7). This approach also did not show correlation with indicators of service usage, such as: service events count, and service days.

Table 4.7 Correlations with resource use indicators of alternative episode construction approaches

Episode Type	Resource Use ($) vs. Inpatient Days of Stay R^2 [95% CI]	Log-Resource Use ($) vs. Inpatient Days of Stay R^2 [95% CI]	Resource Use ($) vs. Service Events Count R^2 [95% CI]	Log-Resource Use ($) vs. Service Events Count R^2 [95% CI]
Interval 1-day	0.001 [0, 0.002]	0.004 [0.002, 0.006]	0.005 [0.004, 0.008]	0.010 [0.007, 0.012]
Interval 7-day	0.002 [0.001, 0.003]	0.004 [0.002, 0.006]	0.040 [0.034, 0.046]	0.044 [0.037, 0.050]
Interval 14-day	0.004 [0.002, 0.007]	0.007 [0.004, 0.011]	0.080 [0.068, 0.091]	0.097 [0.085, 0.109]
Discharge + 30 days	0.015 [0.004, 0.031]	0.010 [0.002, 0.024]	0.595 [0.559, 0.629]	0.483 [0.442, 0.523]
Discharge + 60 days	0.022 [0.009, 0.042]	0.012 [0.003, 0.027]	0.690 [0.660, 0.718]	0.552 [0.514, 0.589]
Discharge + 90 days	0.027 [0.012, 0.047]	0.014 [0.004, 0.031]	0.703 [0.673, 0.73]	0.571 [0.533, 0.607]
1st Contact + 30 days	0.015 [0.005, 0.032]	0.011 [0.003, 0.026]	0.597 [0.560, 0.631]	0.498 [0.457, 0.537]
1st Contact + 60 days	0.022 [0.009, 0.042]	0.013 [0.003, 0.028]	0.685 [0.654, 0.713]	0.554 [0.516, 0.591]
1st Contact + 90 days	0.027 [0.012, 0.048]	0.014 [0.004, 0.031]	0.702 [0.673, 0.730]	0.573 [0.535, 0.608]
Episode Type	Resource Use ($) vs. Service Days Count R^2 [95% CI]	Log-Resource Use ($) vs. Service Days Count R^2 [95% CI]	Resource Use ($) vs. ACT Usage Spearman's Rho [95% CI]	Log-Resource Use ($) vs. ACT Usage Spearman's Rho [95% CI]
Interval 1-day	0.004 [0.003, 0.007]	0.010 [0.007, 0.013]	0.100 [0.08, 0.110]	0.1 [0.080, 0.11]
Interval 7-day	0.034 [0.029, 0.040]	0.041 [0.035, 0.048]	0.110 [0.090, 0.130]	0.110 [0.090, 0.130]
Interval 14-day	0.070 [0.060, 0.081]	0.094 [0.082, 0.106]	0.180 [0.150, 0.200]	0.180 [0.150, 0.200]
Discharge + 30 days	0.545 [0.506, 0.582]	0.579 [0.542, 0.614]	0.300 [0.240, 0.350]	0.300 [0.240, 0.350]
Discharge + 60 days	0.664 [0.632, 0.694]	0.619 [0.584, 0.652]	0.330 [0.280, 0.380]	0.330 [0.280, 0.380]
Discharge + 90 days	0.693 [0.663, 0.72]	0.633 [0.599, 0.665]	0.340 [0.290, 0.390]	0.340 [0.290, 0.390]
1st Contact + 30 days	0.540 [0.501, 0.577]	0.585 [0.549, 0.620]	0.300 [0.250, 0.350]	0.300 [0.250, 0.350]
1st Contact + 60 days	0.662 [0.630, 0.692]	0.624 [0.590, 0.657]	0.330 [0.280, 0.380]	0.330 [0.280, 0.380]
1st Contact + 90 days	0.694 [0.664, 0.721]	0.635 [0.601, 0.666]	0.350 [0.300, 0.400]	0.350 [0.300, 0.400]
Episode Type	Resource Use vs. Schizophrenia Spearman's Rho [95% CI]	Log-Resource Use vs. Schizophrenia Spearman's Rho [95% CI]	Resource Use vs. Major Diagnoses R^2 [95% CI]	Log-Resource Use vs. Major Diagnoses R^2 [95% CI]
Interval 1-day	0.040 [0.020, 0.050]	0.040 [0.020, 0.050]	0.003 [0.002, 0.005]	0.008 [0.005, 0.010]
Interval 7-day	0.010 [0.000, 0.030]	0.010 [0.000, 0.030]	0.002 [0.001, 0.004]	0.002 [0.001, 0.003]
Interval 14-day	0.050 [0.030, 0.070]	0.050 [0.030, 0.070]	0.004 [0.002, 0.007]	0.007 [0.004, 0.011]
Discharge + 30 days	0.140 [0.090, 0.200]	0.140 [0.090, 0.200]	0.025 [0.011, 0.046]	0.023 [0.009, 0.043]
Discharge + 60 days	0.150 [0.090, 0.200]	0.150 [0.090, 0.200]	0.026 [0.011, 0.047]	0.023 [0.009, 0.042]
Discharge + 90 days	0.150 [0.090, 0.200]	0.150 [0.090, 0.200]	0.026 [0.011, 0.046]	0.023 [0.009, 0.043]
1st Contact + 30 days	0.150 [0.090, 0.200]	0.150 [0.090, 0.200]	0.025 [0.011, 0.045]	0.024 [0.010, 0.043]
1st Contact + 60 days	0.150 [0.090, 0.200]	0.150 [0.090, 0.200]	0.025 [0.011, 0.046]	0.023 [0.009, 0.043]
1st Contact + 90 days	0.150 [0.090, 0.200]	0.150 [0.090, 0.200]	0.026 [0.011, 0.046]	0.024 [0.010, 0.043]

Abbreviation: ACT, Assertive Community Treatment; Major Diagnoses (categorical variable): schizophrenia, cognitive disorders, mood disorders, personality disorders, eating disorders, and others

For the episodic basis approach (episodes of 30, 60, 90 days from either discharge or first contact with the community mental health agency), the longer episode length of 90-day appeared to be more correlated with indicators of service use than the 30 or 60-day lengths (Table 4.7). The difference in correlation was not as noticeable for indicators of clinical needs for longer follow-up periods.

The comparison of using different episode starting points, either discharge date or first contact with the community mental health agency, showed that there was no major difference in responsiveness to indicators of service use or clinical needs (Table 4.7). In practice, it may be easier to use the first service contact as the start of an episode because it does not assume that the discharge date is readily available or known by the community mental health agency. Additionally, the discharge date is only applicable to individuals who previously had an inpatient episode, therefore using the first service contact has a broadly applicability to the whole population. Overall, a 90-day community episode of care starting from the first contact appeared to be the most appropriate given the data.

4.3.3 Explained Variance in Community Mental Health Resource Use of SCIPP

Using the SCIPP groups as categorical variables resulted in better explained variance than using the SCIPP's CMI (Table 4.8). The log-transformed resource use resulted in less explained variance than the raw resource use data. SCIPP at the time of inpatient psychiatry discharge was able to explain up to 6% of the variance of resource use for 90 days from the first contact post-discharge.

This result aligned with the analysis of different episode constructions in showing that an episode of 90 days from the first contact was the most appropriate for the data (Table 4.7). Additionally, the low explained variance of the discharge SCIPP's CMIs showed that the CMIs, which were meant to predict inpatient service use, was not good predictors of usage and high usage of community mental health services post-discharge (Tables 4.3, 4.4, 4.5).

Table 4.8 Explained Variance of SCIPP for Community Mental Health Resource Use

Episode Type	Resource Use ($) vs. SCIPP Groups R^2 [95% CI]	Log-Resource Use ($) vs. SCIPP Groups R^2 [95% CI]	Resource Use ($) vs. SCIPP CMI R^2 [95% CI]	Log-Resource Use ($) vs. SCIPP CMI R^2 [95% CI]
Interval 1-day	0.016 [0.013, 0.02]	0.036 [0.031, 0.041]	0.000 [0.000, 0.001]	0.002 [0.001, 0.003]
Interval 7-day	0.028 [0.023, 0.033]	0.021 [0.017, 0.026]	0.000 [0.000, 0.000]	0.000 [0.000, 0.001]
Interval 14-day	0.046 [0.037, 0.055]	0.043 [0.035, 0.052]	0.000 [0.000, 0.001]	0.001 [0.000, 0.003]
Discharge + 30 days	0.048 [0.027, 0.074]	0.055 [0.033, 0.082]	0.001 [0.001, 0.007]	0.004 [0.000, 0.014]
Discharge + 60 days	0.058 [0.035, 0.085]	0.054 [0.032, 0.081]	0.001 [0.001, 0.008]	0.003 [0.000, 0.012]
Discharge + 90 days	0.060 [0.037, 0.089]	0.053 [0.031, 0.08]	0.001 [0.001, 0.007]	0.002 [0.000, 0.011]
1st Contact + 30 days	0.051 [0.029, 0.078]	0.055 [0.033, 0.083]	0.001 [0.001, 0.008]	0.003 [0.000, 0.013]
1st Contact + 60 days	0.059 [0.035, 0.087]	0.052 [0.03, 0.079]	0.001 [0.001, 0.007]	0.002 [0.000, 0.011]
1st Contact + 90 days	0.060 [0.037, 0.089]	0.053 [0.031, 0.08]	0.001 [0.001, 0.007]	0.002 [0.000, 0.011]

Abbreviations: SCIPP, System for Classification of Inpatient Psychiatry; CMI, Case-mix Index

At the sub-group level, the coefficient of variation and confidence intervals of the means of the observed resource use for episodes of 90 days from the first contact suggested that there was still a lot of within group variation (Table 4.9).

The covariance and its confidence intervals between group classification and the residual errors indicated that the two variables were independent for all groups. When stratifying by sex, the covariance was -2.90 [-11.61, 5.39], indicating that sex and residual difference were also independent, with zero included in the confidence interval.

Table 4.9 Characteristics of sample stratified by SCIPP groups and by sex based on an episode of care of 90 days from first contact

Sub-groups	Observed n	Observed Resource Use ($) Mean [95% CI]	SCIPP CMI	Observed CMI	Coefficient of Variation	Covariance(Group, Residual) [95% CI]
SZPA1	2	436 [-2254, 3126]	2.17	2.01	0.69	0.00 [-0.71, 0.66]
SZPA2	11	137 [2, 272]	1.40	0.63	1.46	0.00 [-1.66, 1.53]
SZPBA1	13	70 [14, 125]	1.49	0.32	1.31	0.00 [-1.80, 1.67]
SZPBC1	3	230 [-487, 946]	1.37	1.06	1.26	0.00 [-0.87, 0.80]
SZPBC2	22	278 [116, 439]	1.08	1.28	1.31	0.00 [-2.34, 2.16]
SZPBC3	50	194 [115, 273]	0.89	0.89	1.43	0.00 [-3.48, 3.22]
SZPBD1	388	286 [249, 322]	0.76	1.31	1.29	0.00 [-8.16, 7.54]
SZPBE1	28	432 [229, 634]	1.37	1.99	1.21	0.00 [-2.63, 2.43]
SZPBE2	3	181 [-521, 884]	0.97	0.84	1.56	0.00 [-0.87, 0.80]
SZPBF1	4	106 [-36, 248]	0.95	0.49	0.84	0.00 [-1.00, 0.93]
SZPBF2	3	44 [-77, 164]	0.69	0.20	1.11	0.00 [-0.87, 0.80]
ODC1	9	191 [-45, 426]	1.04	0.88	1.60	0.00 [-1.50, 1.39]
MDAA1	6	99 [-67, 266]	2.12	0.46	1.60	0.00 [-1.23, 1.14]
MDAA2	3	297 [-189, 782]	1.68	1.37	0.66	0.00 [-0.87, 0.80]
MDAB1	17	120 [39, 202]	1.39	0.55	1.32	0.00 [-2.06, 1.90]
MDBA2	9	502 [-158, 1162]	1.16	2.31	1.71	0.00 [-1.50, 1.39]
MDBC2	1	90 [NA, NA]	1.08	0.42	NA	0.00 [-0.50, 0.46]
MDBC3	14	260 [111, 409]	0.83	1.20	0.99	0.00 [-1.87, 1.73]
MDBD2	156	196 [150, 242]	0.79	0.90	1.49	0.00 [-5.86, 5.42]
MDBE1	22	92 [39, 146]	0.78	0.43	1.31	0.00 [-2.34, 2.16]
MDBE2	191	169 [135, 205]	0.64	0.78	1.42	0.00 [-6.37, 5.89]
PDA1	39	206 [95, 317]	1.38	0.95	1.67	0.00 [-3.09, 2.86]
PDA2	69	181 [126, 237]	0.98	0.83	1.28	0.00 [-4.05, 3.75]
EDA1	1	133 [NA, NA]	0.72	0.61	NA	0.00 [-0.50, 0.46]
SDAA1	48	108 [57, 158]	1.15	0.50	1.62	0.00 [-3.41, 3.16]
SDBA1	16	275 [26, 524]	0.81	1.27	1.70	0.00 [-2.00, 1.85]
SDBA2	20	119 [40, 199]	0.58	0.55	1.43	0.00 [-2.23, 2.06]
OTHAA1	3	100 [-72, 271]	1.72	0.46	0.69	0.00 [-0.87, 0.80]
OTHBA1	28	103 [49, 157]	1.33	0.47	1.36	0.00 [-2.63, 2.43]
OTHDA1	24	113 [30, 196]	1.16	0.52	1.74	0.00 [-2.44, 2.25]
OTHDA2	4	389 [-207, 986]	0.86	1.79	0.96	0.00 [-1.00, 0.93]

Abbreviations: SCIPP, System for Classification of Inpatient Psychiatry; CMI, Case-mix Index

4.3.4 Effects of Time to Service Initiation and Prior Contact

The explained variance of SCIPP for community mental health resource use was examined at different points in time after discharge from inpatient psychiatry. The results showed that the explained variance decreases quickly within the first few days of discharge (Figure 4.4). The decrease was observed for both groups of those who had contact with the community mental health agency within 30 days prior to their inpatient episode and those who did not. The decrease lasted longer for the group without prior contact. The explained variance performance stabilized for both groups around the 30-day post-discharge mark. Although the confidence intervals overlapped, the explained variance of SCIPP appeared higher for those with prior contact than those without prior contact after the trends stabilized.

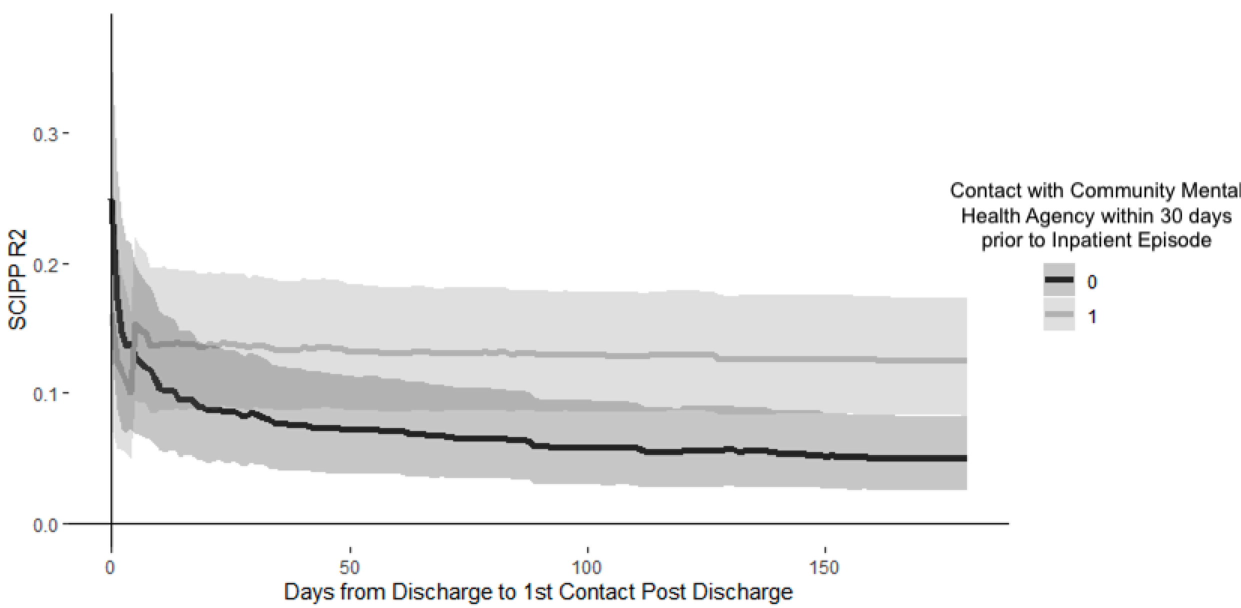

Figure 4.4 Explained variance and 95% confidence intervals of resource use for episodes of 90 days from first contact post-discharge using SCIPP, stratified by contact within 30 days prior to inpatient admission

When contact prior to the inpatient episode was included as another split in the regression tree, the explained variance of resource use 90 days from first contact post-discharge using SCIPP improved to 11.9% for the raw resource use and 14.1% for the log-transformed resource use (Table 4.10). However, the explained variance of the CMI did not improve with the addition of prior contact variable.

Table 4.10 Explained Variance of SCIPP and Prior Contact within 30 days of inpatient admission for community mental health resource use

Episode Type	Resource Use ($) vs. Prior Contact + SCIPP Groups R^2 [95% CI]	Log-Resource Use ($) vs. Prior Contact + SCIPP Groups R^2 [95% CI]	Resource Use ($) vs. Prior Contact + SCIPP CMI R^2 [95% CI]	Log-Resource Use ($) vs. Prior Contact + SCIPP CMI R^2 [95% CI]
Interval 1-day	0.017 [0.013, 0.02]	0.038 [0.033, 0.044]	0.000 [0.000, 0.001]	0.002 [0.001, 0.003]
Interval 7-day	0.028 [0.023, 0.034]	0.022 [0.018, 0.027]	0.000 [0.000, 0.000]	0.000 [0.000, 0.001]
Interval 14-day	0.049 [0.04, 0.058]	0.053 [0.044, 0.063]	0.000 [0.000, 0.001]	0.001 [0.000, 0.003]
Discharge + 30 days	0.089 [0.061, 0.122]	0.130 [0.096, 0.167]	0.001 [0.001, 0.007]	0.004 [0.000, 0.014]
Discharge + 60 days	0.114 [0.083, 0.15]	0.134 [0.100, 0.172]	0.001 [0.001, 0.008]	0.003 [0.000, 0.012]
Discharge + 90 days	0.121 [0.088, 0.157]	0.143 [0.108, 0.181]	0.001 [0.001, 0.007]	0.002 [0.000, 0.011]
1st Contact + 30 days	0.094 [0.065, 0.127]	0.133 [0.099, 0.17]	0.001 [0.001, 0.008]	0.003 [0.000, 0.013]
1st Contact + 60 days	0.117 [0.085, 0.152]	0.136 [0.101, 0.173]	0.001 [0.001, 0.007]	0.002 [0.000, 0.011]
1st Contact + 90 days	0.119 [0.087, 0.155]	0.141 [0.106, 0.179]	0.001 [0.001, 0.007]	0.002 [0.000, 0.011]

Abbreviations: SCIPP, System for Classification of Inpatient Psychiatry; CMI, Case-mix Index

4.3.5 Explained Variance in Community Mental Health Resource Use of the Simplified AMHCC

The results showed that the HoNOS and LSP-16 complexities were very limited in their ability to explain the variance of observed community mental health resource use in the context of this study. Both achieved $R^2 < 1\%$ (Table 4.11).

Table 4.11 Explained Variance of the HoNOS complexity, LSP-16 Complexity, and simplified AMHCC of community mental health resource use

Episode Type	Resource Use ($) vs. HoNOS Complexity R^2 [95% CI]	Log-Resource Use ($) vs. HoNOS Complexity R^2 [95% CI]	Resource Use ($) vs. LSP-16 Complexity R^2 [95% CI]	Log-Resource Use ($) vs. LSP-16 Complexity R^2 [95% CI]	Resource Use ($) vs. Age + HoNOS + LSP-16 R^2 [95% CI]	Log-Resource Use ($) vs. Age + HoNOS + LSP-16 R^2 [95% CI]
Interval 1-day	0.000 [0.000, 0.001]	0.000 [0.000, 0.001]	0.000 [0.000, 0.001]	0.001 [0.000, 0.001]	0.001 [0.000, 0.002]	0.001 [0.000, 0.002]
Interval 7-day	0.002 [0.001, 0.003]	0.000 [0.000, 0.001]	0.001 [0.000, 0.002]	0.001 [0.000, 0.002]	0.003 [0.002, 0.005]	0.003 [0.001, 0.005]
Interval 14-day	0.003 [0.001, 0.005]	0.000 [0.000, 0.002]	0.001 [0.000, 0.003]	0.002 [0.000, 0.004]	0.006 [0.003, 0.010]	0.005 [0.003, 0.009]
Discharge + 30 days	0.002 [0.000, 0.010]	0.001 [0.001, 0.007]	0.000 [-0.003, 0.004]	0.000 [-0.003, 0.003]	0.007 [0.001, 0.020]	0.005 [0.000, 0.016]
Discharge + 60 days	0.001 [0.001, 0.006]	0.000 [-0.001, 0.006]	0.000 [-0.004, 0.003]	0.000 [-0.002, 0.004]	0.009 [0.001, 0.023]	0.007 [0.001, 0.020]
Discharge + 90 days	0.000 [-0.002, 0.004]	0.000 [-0.001, 0.006]	0.000 [-0.004, 0.003]	0.000 [-0.002, 0.005]	0.011 [0.002, 0.027]	0.008 [0.001, 0.022]
1st Contact + 30 days	0.002 [0.000, 0.010]	0.001 [0.000, 0.009]	0.000 [-0.003, 0.004]	0.000 [-0.002, 0.005]	0.008 [0.001, 0.022]	0.006 [0.000. 0.019]
1st Contact + 60 days	0.001 [0.001, 0.007]	0.000 [-0.001, 0.006]	0.000 [-0.004, 0.003]	0.000 [-0.002, 0.004]	0.009 [0.001, 0.023]	0.007 [0.001. 0.020]
1st Contact + 90 days	0.000 [-0.002, 0.004]	0.000 [-0.001, 0.006]	0.000 [-0.004, 0.003]	0.000 [-0.002, 0.005]	0.012 [0.003, 0.027]	0.009 [0.001, 0.023]

Abbreviations: HoNOS, Health of the Nation Outcome Scales; LSP-16, Life Skills Profile - 16

In order to operationalize the simplified AMHCC (Figure 4.2), two regression tree models were fitted to stratified sub-samples of two age groups (18-64 and 65+) to determine empirical high complexities thresholds for the HoNOS and LSP-16. The results showed that thresholds for high complexities that maximized the R^2 on empirical data are: HoNOS > 24 and LSP-16 > 13 for the 18-64 age group, and 10 and 2 respectively for the 65+ group (Figure 4.2). The simplified AMHCC (Figure 4.2) using the combination of age, HoNOS and LSP-16 achieved a maximum of 1.2% of explained variance in observed resource use for an episode of 90 days from first contact (Table 4.11).

4.4 Discussion

Overall, the results suggested that SCIPP partitioned clients of community mental health services into relatively homogeneous groups in term of resource use better than using diagnosis alone (Tables 4.7, 4.8). SCIPP at inpatient psychiatry discharge, by itself, explained a modest amount (6%) of variance of community mental health resource use, while diagnosis only explained up to 2.6% of variance. Age, HoNOS, and LSP-16 components of the AMHCC explained 1.2% of variance.

In general, case-mix classification systems for community mental health generally have lower explained variance than other care settings, such as nursing home care of more than 50% [32]. A previous review found the range of explained variance for community mental health was between <1% to 26% (Chapter 2) [113]. The majority of systems reviewed achieved R^2 between 5-10%, and newer systems from Australia and New Zealand were the best performers [113]. First, it is worth noting that many systems found in the literature review predicted resource use on an annual basis, which could make magnitude of the variance relative to the total resource use less prominent and resulted in better performance. Second, the reported performance from most studies were the performance based on the training data and not on a prospectively validation dataset like this study, which could have resulted in an overestimation [113]. The highest prospectively validated was by the PsyCMS ($R^2 = 6.4\%$) developed using a sample from the US Veterans Affairs close to 1 million but it aims to explain both inpatient and outpatient services on an annual basis [50]. The PsyCMS classification system uses diagnoses as the main differentiator of levels of resource use.

The coefficient of variance of SCIPP showed that there was still large within group variance of resource use (Table 4.9). The covariance of group classification and residual errors showed that equitable allocation of resource was observed across SCIPP groups or by sex (Table 4.9). Time to service initiation and prior contact with the community mental health agency could have an influence on the explained variance of SCIPP. Longer delay in service initiation post-discharge appeared to have a negative effect on explained variance of SCIPP for those who did not have prior contact with the community mental health agency (Figure 4.4). One possible explanation is that the lag between the inpatient discharge and community service initiation could make the information from the discharge assessment less accurate as a measure of a person's actual clinical characteristics at service initiation. Therefore, the predictive power of the discharge assessment could be reduced as the gap increases. These results pointed towards the need to have contemporaneous clinical data at the time of community service initiation.

When taking into account prior contact with the community mental health agency, the performance of the model that included SCIPP improved up to 14.1% (Table 4.10). One possible explanation for the difference in explained variance for the sub-group who had prior contact and the ones who did not was that the community mental health agency may have had a better sense of the person's needs because they were familiar to the staff. Those who had a prior contact also might already have a care plan ready to start immediately after discharge from inpatient psychiatry. Therefore, despite the effect of delay between discharge and service initiation, the resource use post-discharge was likely structured, and by extension, more predictable. Whereas those who did not have a prior contact may first need to go through an in-depth assessment in order to create a care plan. Compounded with the effect of delay between discharge and service initiation, the resource use for this sub-group was likely more variable than a pre-determined care plan.

More research is needed to refine and adapt SCIPP for use in the community setting. Alternatively, there may be a need for a new case-mix classification system for community settings. The use of inpatient days of stay may be relevant for inpatient psychiatry but was not observed to be informative to resource use post-discharge. However, a closely related length of service indicator such as days since community episode initiation may also be relevant in predicting community resource use. Specifically, it was observed that the first 30 of 90 days of a community episode could consume up to half of the total resource use of the episode (Table 4.6). Prior contact with the community mental health agency could potentially be another differentiator with a similar effect to inpatient length of stay, in which having a prior contact with the community mental health agency resulted in higher explained variance of community resource use.

Future research should also consider a few fundamental differences in the two settings. SCIPP was developed for inpatient care settings, which have a different pattern of resource consumption. Such differences in usage pattern were significant enough to warrant a separate classification systems tailored for each care settings in previous research from Australia and New Zealand [35, 45]. Additionally, differences across care settings could have contributed to why the SCIPP's CMI did not perform well as a predictor of resource use in the community settings. The results showed that the relative resource intensity of the inpatient psychiatry (indicated by the SCIPP CMI) may not be well calibrated for the community resource use data (Tables 4.3, 4.4, 4.5, 4.8). Some factors that are associated with high resource intensity of SCIPP were risk of self-harm and harming others (Figure 4.1, 4.2). In a previous study (Chapter 3), it was shown that these factors are not always associated with high community service resource use. However, these factors were associated with receiving specialized services designed for clients at risk of self-harm and harming others, which may have a different pattern of usage compared to the inpatient settings. For example, the specialized programs serving clients at risk of self-harm provided services in a group setting that had the effect of lowering the intensity of resource consumption per client. More research is needed to fully understand the pattern of usage of clients at high risk of harming others because it involves the interaction between the services provided by justice system and services provided by the community mental health agency that the data of this study do not fully cover.

For the AMHCC, due to the inability to derive the phases of care from the RAI-MH data, the analysis was only partial, focusing only on the performance of age, HoNOS, and LSP-16. These components were observed to be limited in explaining only 1.2% of variance in the context of this study. This analysis used a rudimentary crosswalk from the RAI-MH to the HoNOS and LSP-16 using closely related items but not entirely exact, which could also contributed to the low observed performance. Although it is typical for the explained variance to be lower in replication studies compared to original development study, the results suggested that the bulk of the explanatory power of the AMHCC was likely came from the phases of care component. In its development study, the phases of care explained up to 20.6% out of 26.6% of the total variance observed [46]. Therefore, the poor observed explained variance of the AMHCC without phases of care was not entirely surprising.

The use of phases of care appeared to be a barrier in applying the AMHCC beyond the Australian context. The phases of care attempted to bring together two concepts of needs of care and goals of care [114], which are determined at the start of a community episode by a clinician. The definitions of the phases of care were not immediately operationalizable or derivable from other well-known mental health assessments, including the HoNOS, LSP-16, and RAI-MH.

Since the phases of care are determined by a clinician, it is possible that some degree of subjectivity may be part of the determination. In terms of resource allocation, room for subjectivity may give rise to potential of up-coding or gaming for financial incentive. In terms of risk adjustments or population comparison applications, more research is needed to establish the inter-rater reliability of the phases of care and the extent of their validity in other countries given that they are not entirely driven by the measurable clinical characteristics.

The main limitation of this study was selection bias of the sample containing only people who were recently discharged from inpatient psychiatry. This sample may represent a narrow set of people at higher clinical needs than the average clients served by the community mental health agency. With an implementation of an assessment tool that is compatible with the SCIPP at the community mental health agency, a replication study that uses the full population of community clients can offer a better estimation of performance. Most importantly, a more contemporaneous community mental health clinical assessments measured at the time of community service that are paired with community resource use are needed to counter the effect of time to service initiation delay observed in this study.

Additionally, the study sample also excluded clients with cognitive disorders who are served by the inter-agency programs, which would have been classified to be at high resource intensity according to SCIPP. The community mental health agency of this study does not provide addiction services, therefore resource use related to addiction services were also absent from the study data. The resource use data for these subgroups is also needed in future replication study.

This study also made contribution to the topic of episode of care by testing several methods of episode of care construction using empirical community mental health resource use data. The results suggested that a simple 90-day period from the 1st contact is an appropriate method to construct a community episode of care. It appeared to be clinically relevant and responded well to other indicators of service use intensity (Table 4.7). Although an interval basis approach in episode construction was not shown to be optimal, they offered additional insights into the usage pattern of community mental health services. The data confirmed the intermittent nature of community service use, which could be ongoing for a long time period (mean of 82.9 days) (Table 4.6). The mean of the activity interruption gaps (9.9 days) showed that the frequency of service use was likely to be between a weekly or bi-weekly basis.

This study contributed to existing case-mix classification research through experimentation with different episode construction methods for non-continuous care in community settings, testing an inpatient psychiatric and a community mental health case-mix classification systems in a new context. Despite its limitations, the use of discharge SCIPP grouping classification can offer an improved prediction of resource use for community mental health compared to using diagnosis alone for the subset of people previously discharged from inpatient psychiatry. SCIPP has the potential to allocate resources in an equitable manner for community mental health services. The AMHCC was not immediately usable beyond the Australian context. More research is needed to refine and tailor SCIPP for community settings to maximize its potential as a viable classification system. Alternatively, further development of a new community mental health case-mix classification can also be done by using contemporaneous assessment data at the time of community service initiation and the community resource use data.

Chapter 5
Study Three: Classifying Resource Use for Community Mental Health Services at the Transition from Inpatient Psychiatry Using Machine Learning

Abstract

Case-mix classification systems connects the measures of clinical needs and health care resource use at the individual level. Machine learning algorithms were used to classify levels of observed community mental health service resource use for a sample of adults discharged from inpatient psychiatry. Cross-validation results showed that the achievable explained variance of community mental health resource use by the clinical information measured at discharge from inpatient psychiatry is about 12%. Simple decision trees models showed comparable performance as complex models in cross-validation. Although machine learning can uncover patterns in the observed data, human expertise is still required in the development of case-mix classification system to ensure that the resource allocation does not only fit the observed data but also supports the objective of delivering better outcomes. This study showed that combining insights from machine learning and human expertise can aid in the development of case-mix classification systems.

5.1 Introduction

Individuals who use community mental health services vary widely in terms of their clinical needs. From the point of view of a health care system, it is possible to measure the clinical needs in a standardized manner. Examples of standardized measures of clinical needs may include psychiatric diagnosis (e.g., the Diagnostic and Statistical Manual of Mental Disorders) and clinical severity scores related to symptoms or levels of risk of adverse outcomes. Ideally, the delivery health care services should be related to clinical needs. The absence of such a connection for the mental health care system may lead to inequity gaps in health care funding, biases in monitoring performance based on outcomes of care, and hindrance in long-term capacity planning [34].

Case-mix classification systems aim to connect the measures of clinical needs and resource consumption. A case-mix classification system partitions the population into relatively homogeneous groups based on resource utilization and similar clinical needs [23, 34]. It enables a clinically meaningful way of relating the health care needs to the resources required for providing care, the expected health outcomes at the individual level, and quality measures and capacity planning at the organization or health system level [34].

The province of Ontario, Canada has planned to convert their current funding for mental health based on a fixed global budget to a funding formula in which a portion of the funding is based on case-mix classification [26]. Case-mix classification systems have the potential to contribute to equity in treatment allocation and more efficient use of health care funding [19].

However, most case-mix classification systems for mental health focused only on the inpatient psychiatry settings, and not the community mental health settings [113], despite the fact that mental health care has shifted from facility-based care to community care in the past few decades [8, 10–12].

Few case-mix systems have attempted to track patients across care settings, except for those from Australia and New Zealand that attempted to unify two separate classifications for inpatient and community settings together [113]. Since of the approaches to care between these settings share some similarities and overlaps, there should be some continuity of relations between need and resource use between inpatient and community-based care. In a previous study (Chapter 4), the System for Classification of In-Patient Psychiatry (SCIPP) at discharge from inpatient settings explained about 6% of variance in resource use of community mental health resource use in the 90 days after discharge from psychiatry, and up to 14% of the log-transformed resource use when taking into account contact with the community mental health agency prior to the inpatient episode. The SCIPP was developed specifically for the inpatient settings and some aspects were not compatible with the community settings. For example, the use of inpatient days of stay as an indicator of resource intensity is not likely to be related to resource intensity in the community over time.

Using data from one of the largest community mental health agencies in Ontario and Canada, this study experimented with several strategies to build models for predicting community mental health resource use post-discharge using the clinical assessment done at discharge from inpatient psychiatry. The results are expected to contribute to the ongoing effort of Ontario to develop a case-mix classification for community mental health settings.

5.2 Methods

5.2.1 Data Sources

People discharged from inpatient psychiatry at a hospital in the Waterloo-Wellington region of Ontario (Grand River Hospital) were followed to capture their usage of publicly-funded community mental health services post-discharge. Their provincially mandated Resident Instrument Assessment - Mental Health (RAI-MH) discharge assessments were obtained as the source of clinical input that is expected to predict subsequent usage of community mental health services. The RAI-MH assessment is required to be done at admission, discharge, and every 90 days for all inpatients admitted to psychiatric beds in Ontario. The assessments measures many clinical domains at the individual level, including embedded clinical severity scales and assessment protocols that can summarize the assessment in several key areas [52].

Records from two organizations were linked using primarily health card number and date of birth. If the linkage was not successful with the primary method, a secondary method of using full name and date of birth was subsequently attempted. The data spanned the 2014-2018 calendar years. The use of these data and record linkage was approved by the Grand River Hospital (# 2018-0669), the CMHA-WW ethics board in October 2018, and the University of Waterloo (# 40147).

The study sample included inpatients who initiated services from the community mental health agency within 30 days post-discharge. In a previous study (Chapter 4), delay in service initiation could reduce predictive utility of the clinical data, possibly due to the changes in the clinical profile over time compared to at discharge. Therefore, this study limited the time between discharge and service initiation. To ensure that episodes of care observations were comparable, this study only included episodes with no readmission within 90 days, and not right-censored.

Additionally, those who were younger than 18 at discharge, received children mental health services post-discharge, and forensic inpatient admissions were excluded because the ethics application only approved the study of adults. Discharges due to transfer, same day readmission, or in-hospital mortality were also excluded because these discharges would not subsequently use community mental health services. Admissions to inpatient psychiatry for ≤ 3 days were excluded because only the intake administrative information part of the RAI-MH was required to be filled out.

In a previous study (Chapter 4), an episode of 90 days starting from discharge appeared to be an appropriate episode of care used for follow-up to capture resource use post-discharge. The resource use data used in this study was the wage-weighted staff time of direct services provided by a community mental health agency in the Waterloo-Wellington region (Canadian Mental Health Agency - Waterloo Wellington).

Direct services that were directly provided to a client either in person or virtually, and not covered under fee-for-service arrangement (such as services provided by psychiatrist or physicians). Indirect services are activities such as: documentation, case review, and care team consultation. Non-client specific activities, such as staff training or meetings, were not available and also not included in the resource use data of this study. Services provided through an inter-agency programs were not included because of data gaps from partner agencies for some services.

Each service encounter at the community mental health agency was recorded with the staff time duration and job title. Using the median wage rates, this study calculated the wage-weighted staff time cost measure for each service encounter. For group services, the wage-weighted staff time was divided equally for all clients registered because resources had already been allocated from the agency's point of view.

These data were expected to be of good quality because they were used for scheduling of appointments, monitored by management, and used to determine extra pay for eligible direct face-to-face time. Additionally, manual verification of data was done for service events longer than a typical work shift of eight hours. Errors in time entry of AM versus PM, or date entries were corrected.

5.2.2 Exploring Limits of Explained Variance Using Machine Learning

Predictive models of resource use often required human expertise to select relevant predictors. For example, the SCIPP was constructed iteratively by a working group of experts and clinicians [27]. However, since the early 2000s when SCIPP was developed, there has been great advancements in machine learning techniques and software that automated the process of building predictive models. Several machine learning techniques were used to explore the limits of explained variance in resource use that can be predicted by the discharge assessment. The machine learning algorithms were used in this study can be broadly classified in two groups: generalized linear models and tree-based models. The machine learning algorithms used were included R package *caret* [121].

Generalized linear models are algorithms that model the output as a weighted linear combination of the input variables. The most basic form is the **linear regression**. While adding more variables to a model could improve the explained variance, it could also lead to having too many co-linear variables. To mitigate such issue, approaches in penalizing the magnitude of each coefficients were used in this study.

Specifically, two additional generalized linear models beside linear regression were considered, Least Absolute Shrinkage and Selection Operator regression or **L1** [122] and the Ridge regression or **L2** [123]. Similar to the typical linear regression, L1 and L2 methods also estimate the coefficients for each variables by minimizing the residual errors between the predicted and observed output. However, these methods add a penalty term proportional to the magnitude of the coefficients to the residual errors, which have the effect of inflating the residual errors and shrinking the magnitude of the coefficients to produce potentially simpler models. The main difference is that L1 adds a penalty term that is proportional to the absolute value of the coefficients, and L2 adds a penalty term that is proportional to the squared value of the coefficients.

Both methods require a constant value to scale the penalty parameters to be specified (also known as a hyperparameter) to be specified a priori instead of being estimated by learning from the data. In order to choose the most optimal value, for each candidate value of the hyperparameter, a 50-fold cross validation with 20 repeats was used in this study to find values of hyperparameter that maximized the mean R^2 on the testing subsets. The possible values of the hyperparameters were trialed randomly because empirical experiments showed that random trials of the hyperparameters are both computationally efficient and yielded acceptable performance compared to manual or pre-defined search [124]. Since the sample size was limited, the cross-validation was also configured to maintain the same proportion of major diagnoses that is representative of the dataset across all folds, according to Table 3.2.

Tree-based models makes prediction by recursively dividing a population using input variables into smaller subsets of observations that are relatively homogeneous. Tree-based models can be advantageous because the grouping of observations into homogeneous groups mirrors classifications in biology (such as phenotype) and medicine (such as disease classification) [125]. The first case-mix classification ever developed, Diagnosis Related Group, was a tree-based model [59]. Other popular tree-based case-mix classification systems include the Resource Utilization Groups [30]. Additionally, issues related to co-linearity and interaction among variables are handled in tree-based models because subsequent splits of tree are conditional upon the parent split.

The simplest form of tree-based model in this study was the Classification and Regression Tree algorithm (CART) [126]. **CART** searches through all input variables to find a cut off value that minimizes the sums of square errors. Similar to linear regression, CART can build a complex model with many variables but could lead to generalizability issues. To prevent CART from building overly complex model, a complexity hyperparameter is used. The complexity parameter sets a threshold that the overall R^2 must increase by in order to split further [121]. Cross-validation was also used to find the optimal value for this hyperparameter.

An alternative approach to using the sums of square errors as the criterion for choosing a split is to use conditional inference tests [127]. For each possible split, the conditional inference tree (**ctree**) algorithm first tests for independence between all input variables and the output, and selects the variable with strongest association with the output indicated by the p-value. The algorithm stops when the null hypothesis cannot be rejected, which was set for $p < 0.05$.

Tree-based models can also be more complex by combining more than one tree to make a ensemble model. Two approaches in ensembling were used, random forest [128] and extreme gradient boosting (xgboost) [129]. **Random forest** makes improvement over a simple regression tree by aggregating predictions of many trees using majority votes [128]. Each tree is built using a bootstrap sample of the data and restricting the number of candidate variables available for splitting. The hyperparameters to be optimized using cross validation is the number of trees and number of candidate variables available for trial at each split. This approach can be viewed as an attempt to simulate a broader range of potential observed cases and aggregate the results to minimize biases that may exist in the full data sample. **Xgboost** makes improvements on random forest by sequentially aiming to improve the prediction of the previously trained tree by minimizing the residual errors using a gradient descent algorithm (or partial derivative to find a minimum value) [129]. This approach can be viewed as making step-wise improvements instead of having each tree as an independent prediction attempt like random forest. Cross-validation was also used to optimize the following hyperparameters: the learning rate that weights the contribution of each tree toward the final prediction, the minimum error reduction required to make further split, the maximum depth of each tree, minimum number in each node, ratio of the sample used for training, and number of variables available to construct each tree.

5.2.3 Exploring the Minimum Functional Set of Input Variables

The role of the input variables is important in maximizing the clinical relevance of a case-mix system. In a previous study (Chapter 2), the role of input variables was discussed extensively. There are several categories of variables: health care needs, process, historical, and individual-level versus provider-level variables. The emphasis should be to given to variables that are indicators of health care needs at the individual-level, such as diagnosis, clinical severity scales, and clinical assessment protocols. Process variables should be avoided since they provide the incentive to do increase volume of high price procedures. Historical variables can be used if they have a long-term relevance in describing a person's health care needs; however, historical variables are insensitive to change in clinical conditions. Since this study examined community mental health resource use from a single organization, no facility-level variables were considered.

In this study, four subsets of input variables were examined for their utility in predicting resource use. The simplest set of input variables is age (groups of 18-25, 26-44, 45-64, 65+), sex, and diagnosis (coded as binary for schizophrenia, cognitive disorders, mood disorders, personality disorders, eating disorders, substance use disorders, and all others), which are usually available from most care settings. The most complex set of input variables is simply to use all clinical variables available from the RAI-MH assessment.

The RAI-MH assessment can be summarized using clinical scales and clinical assessment protocols (CAPs), which could be a more efficient way to include key clinical domains rather than using numerous related items individually. The third set of input variables considered was diagnosis, clinical scales, CAPs, and prior contact with community mental health agency within 30 days prior to inpatient psychiatry admission. The last subset added comorbidity count, Charlson Comorbidity Index (CCI) [130] to diagnosis, clinical scales, CAPs, and prior contact. The clinical scales and CAPs of the RAI-MH are listed in Tables 5.1, 5.2.

Table 5.1 List of clinical scales (continuous variables) embedded in the RAI-MH assessment

Clinical Scales	Purpose	Components
Aggressive Behavior Scale (ABS)	Measures frequency and diversity of aggressive behaviors	Verbal abuse, physical abuse, socially inappropriate/disruptive, resists care
Activities of Daily Living (ADL)	Measures ability to carry out daily living activities	Personal hygiene, locomotion, toilet use, eating
Anhedonia	Count of symptoms of anhedonia	Anhedonia, withdrawal from activities of interest, lack of motivation, reduced social interactions
Cut down, anger, guilt, eye-opener (CAGE)	Screens for substance use	Count of 4 indicators of substance use
Cognitive Performance Scale (CPS)	Measures cognitive status	Daily decision making, short-term memory, expression, self-performance in eating
Depression Rating Scale (DRS)	Count of indicators of negative mood	Negative statements, persistent anger, unrealistic fear, repetitive health complaints, repetitive anxious complaints, sad facial expressions, crying
Depressive Severity Index (DSI)	Count of depressive symptoms	Sad facial expression, negative statements, self-deprecation, guilt, hopelessness
Instrumental Activities of Daily Living Capacity (IADL)	Measures higher level function based on others' perceptions of a person's ability to perform IADLs	Meal preparation, ordinary housework, managing finances, managing medications, phone use, shopping transportation
Mania	Count of symptoms of mania	inflated self-worth, hyperarousal, irritability, increased sociability, pressured speech, labile effect, sleep problems due to hypomania
Pain	Measure of pain	Pain frequency, pain intensity
Positive Symptoms Scale Short(PSSS)	Count of positive symptoms	Hallucinations, command hallucinations, delusions, abnormal thought process
Positive Symptoms Scale Long (PSSL)	Count of positive symptoms	PSSS, inflated self-worth, hyperarousal, pressured speech, abnormal movement
Risk of Harm to Others (RHO)	Measures risk of harm to other people	History of violence or extreme behavior, ABS, positive symptoms, insights into mental health, delusions, sleeping problems
Self-Care Index (SCI)	Measures ability to care for self	Cognitive skills for decision making, insight into mental health, positive symptoms, making self understood, abnormal thought process, poor hygiene, decreased energy, mania, anhedonia
Severity of Self-Harm (SOS)	Measure risk of harm to self	Self-injury ideation, history of suicide attempt, positive symptoms, severity of depression, CPS, family concerned about self-injury, suicide plan
Social Withdrawal Scale (SWS)	Count of social withdrawal indicators	Lack of motivation, reduced interaction, decreased energy, flat/blunted affect, anhedonia, loss of interest

Table 5.2 List of clinical assessment protocols (categorical variables) embedded in the RAI-MH assessment

Clinical Assessment Protocols (CAPs)	Purpose
Harm to Others (mhcHARMOTH)	Indicates imminent risk of harm to others
Self-Harm (mhcSELFHARM)	Indicates high and imminent risk of self-harm
Self-Care (mhcSELFCR)	Indicates risk of inability to care for self
Social Relationships (mhcSOREL)	Indicates risk of social isolation friendship or family dysfunction
Informal Support (mhcINFSUPP)	Indicates need for support related to mental health symptoms, physical disability, or cognitive impairment
Support System for Discharge (mhcSSDIS)	Indicates difficulties post-discharge due to lack of resources
Interpersonal Conflict (mhcIPCON)	Indicates conflict within specific relationships and widespread conflict
Traumatic Life Events (mhcTRAUMA)	Indicates impact of prior traumatic life events or immediate safety concerns
Criminal Activity (mhcCRIM)	Indicates risk of criminal behavior
Personal Finance (mhcFINAN)	Indicates hardship due to loss of income or poverty, incapable of managing property
Education and Employment (mhcEDEMP)	Indicates risk of losing employment, dropping out of school, and need for employment and education support
Control Intervention (mhcCTRLINT)	Indicates the need for control interventions
Medication Management and Adherence (mhcMEDMGT)	Indicates problems with medication management due to cognitive deficits positive symptoms, or stop taking medication due to side effects
Rehospitalization (mhcREHOSP)	Indicates risk of being readmitted within 180 days or sooner
Smoking (mhcTOBUSE)	Indicates risk of withdrawal symptoms and recent tobacco use
Substance Use (mhcSUBUSE)	Indicates problematic or history of problematic substance use
Weight Management (mhcWTMGT)	Indicates problematic body composition and eating behaviors
Exercise (mhcEXER)	Indicates physical activity for persons capable of physical activity or persons less capable due to a medical condition
Sleep Disturbance (mhcSLEEP)	Indicates sleep problems due to cognitive impairment
Pain (mhcPAIN)	Indicates persons with severe and persistent pain
Falls (mhcFALLS)	Indicates risk of future falls

5.2.4 Models Evaluation

The primary metric used was the R^2 and its 95% confidence interval. A global metric, like the R^2, can only provide an overall evaluation of the models [120]. While it is useful to explore the achievable explained variance from the discharge assessment, other considerations regarding clinical validity, practicality, calibration, homogeneity of terminal groups, indicator of inequitable allocation, spread and predictive utility of relative resource intensity were also taken into account [30, 59].

5.2.4.1 Clinical Considerations

As observed in a previous study (Chapter 3), observed resource use does not always increase in the same direction as higher clinical severity. For example, persons at risk of self-harm are often allocated group therapy, which has the effect of lowering observed resource use due to dividing the staff time among group participants.

Therefore, in addition to the simplest functional tree-based models produced by machine learning, another tree was manually built to refine the machine learning-based models and put in additional restrictions. For each split, from the top 10 candidate variables, based on the sum of square errors [126], choose the top variable that split the higher clinical severity in the same direction as the higher observed resource use in daughter nodes. If not available, the top variable that minimized the sum of squares error was used. To simplify the tree, the minimum number of observations in each terminal group was set to be at least 20. The total number of splits was set to be seven, same as the SCIPP, to restrict the model complexity.

5.2.4.2 Practical Considerations

A good case-mix system should be feasible to use in daily practice. Although some machine learning algorithms, such as random forest and xgboost, could produce high predictive performance they are consider “black-box” algorithms, which the mapping of input variables to the final prediction is uninterpretable to humans. A complex multiple regression, such as linear regression, L1, and L2, can be interpreted by humans but they are also not user-friendly in a clinical setting. To use such models, a clinician must either use a calculator or computer to produce the output of the regression equation.

In additional to being compatible with the biological classification paradigm, single tree-base models, or decision trees, are much more practical in a clinical setting because they do not require any computation by a clinician. A tree-based model allows the users to visually follow a decision flow chart with a series of binary decisions to arrive at the output of the model. Therefore, further analysis of model emphasized the simplest functional tree-based models.

5.2.4.3 Calibration

Using the calibration curve could help examine the consistency of the models in predicting resource use across the distribution of observed resource use, and identify areas of weakness of the models. In this study, the calibration curve examined each predicted decile for the proportion of the observed data that is actually smaller or equal to the predicted decile.

5.2.4.4 Evaluation of Terminal Groups

The aim of a classification system is to divide the population into relatively homogeneous groups [125]. To evaluate the homogeneity of the models, the coefficient of variation was used to examine each group of the models for within-group variations. In addition, the spread of the relative resource intensity, or observed case-mix index (CMI), was used as an indicator how well a model separate different levels of resource intensity in the data.

For each terminal group in the tree-based model, the covariance between group classification and the residuals was shown to be a proxy for equitable allocation, $covariance(G, \hat{Y} - Y)$ where $G\{0, 1\}$ indicates the membership of group g [120]. A negative value indicates that being classified into a particular group resulted in larger under-reimbursement relative to the rest of the population, or unfair in resource allocation relative to other groups.

5.2.4.5 Relative Resource Intensity as an Indicator of Usage of Community Mental Health Services Post-Discharge

Another mark of validity is whether the observed relative resource intensity of the terminal groups compared to the average observed case (indicated by the observed CMI) is associated with usage of community mental health services post-discharge (binary). To examine this association, the simplest functional models were used to predict the CMI for all eligible discharges, with both observed usage and non-usage post-discharge. Then, logistic regression models were fitted to examine the association between predicted CMI and usage of community services post-discharge.

5.3 Results

This study identified 1,207 discharges from 1,002 unique persons (Figure 5.1). This sample represents only a small portion of all individuals who came into contact with the community mental health agency that is estimated to be about 10,000 annually, which included children/youth services or contacts through the centralized intake but subsequently referred to another agency.

The mean resource use during an episode of 90 days post-discharge was \$217 (min = \$3, Q1 = 20, Q2 = \$74, Q3 = 308, max = \$3,039). The distribution appeared to be positively skewed with a long right tail of a few but resource intensive clients.

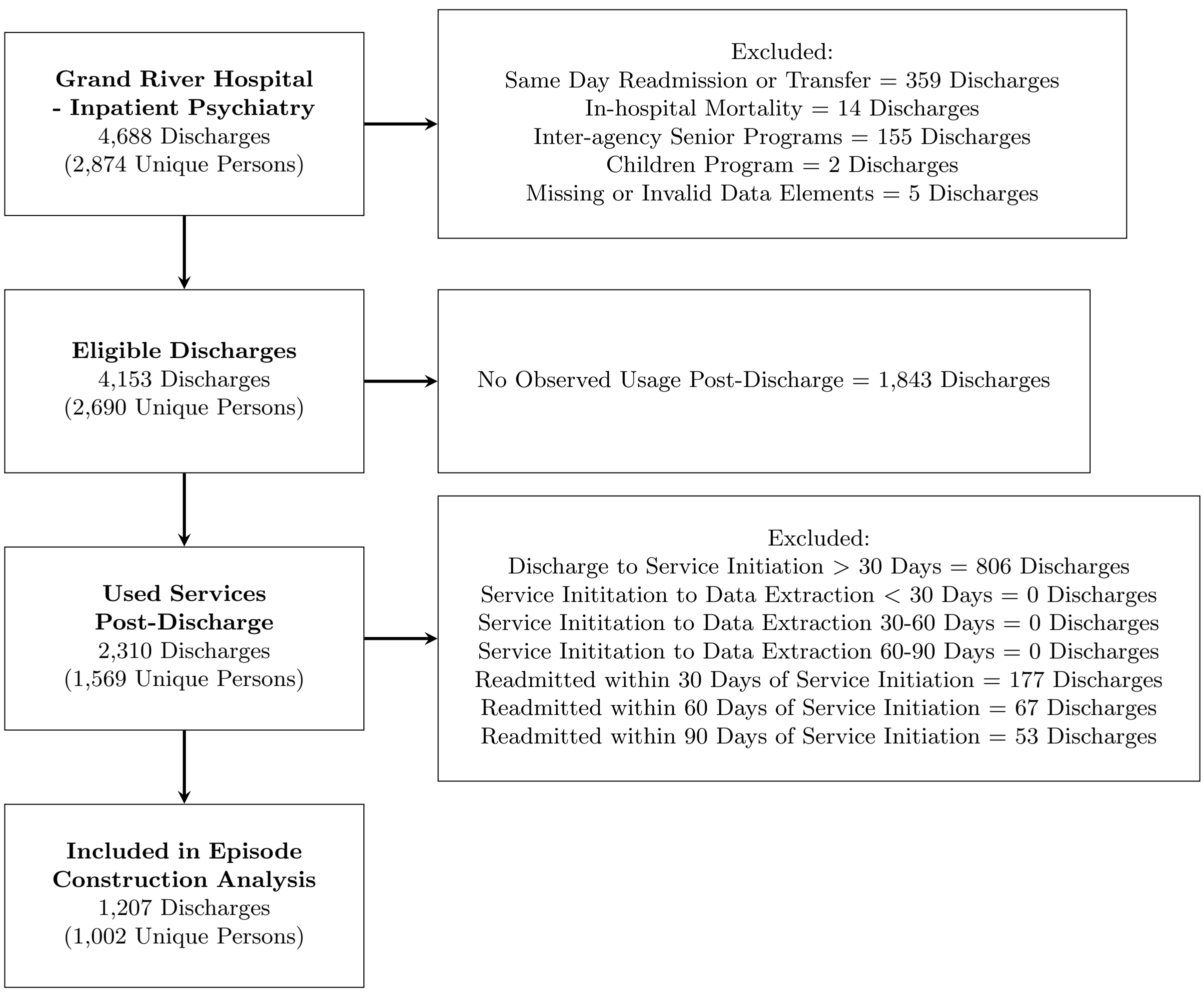

Figure 5.1 Record linkage and sample selection

The results showed that using machine learning could produce models that had high performance on the training data, up to about 93%, but the results of the cross-validation showed that the range of achievable explained variance of models trained using the current data set to predict future observations is about 12% (Table 5.3). Models trained to predict log-transformed resource use achieved higher explained variance than models trained to predict the raw resource use data.

The ensemble algorithms, random forest and xgboost, produced very high performing models on the training data, but the performance did not sustain in the cross-validation (Table 5.3). This behavior is also observed when using all available variables, which achieved high performance on the training data but did not maintain the performance gain in comparison to simpler sets input variables. As also observed in a previous study (Chapter 4), prior contact with the community mental health agency within 30 days before the inpatient episode was a strong predictor of subsequent usage of community mental health services.

The age, sex, and diagnosis subset did not show similar performance as the other sets (Table 5.3). Unlike other subsets that showed higher performance on the training data than cross validation data, this set of variables showed the opposite effect. This behavior is an indicator that the model was perhaps too simple. It did not capture the signal from the training data but was also able to be more generalizable by avoid learning the noise in the data.

Table 5.3 R^2 of models trained using different subsets of input variables from the discharge assessments to predict resource use and log-transformed resource use. Bolded models were analyzed further as simplest functional models.

		Resource use		Log-transformed resource use	
Vars	Algorithm	R^2 [95% CI] on training data	Mean cross-validated R^2 [95% CI]	R^2 [95% CI] on training data	Mean cross-validated R^2 [95% CI]
age, sex,	CART	7.04% [4.50, 10.05]	9.07% [8.43, 9.71]	6.18% [3.80, 9.05]	6.87% [6.35, 7.39]
diagnosis	ctree	5.75% [3.46, 8.55]	8.86% [8.18, 9.54]	5.23% [3.04, 7.92]	5.93% [5.48, 6.38]
	linear regression	4.45% [2.45, 6.99]	7.89% [7.38, 8.41]	4.32% [2.35, 6.82]	7.16% [6.62, 7.70]
	L1	4.45% [2.45, 6.99]	7.89% [7.38, 8.41]	4.32% [2.35, 6.82]	7.16% [6.62, 7.71]
	L2	4.43% [2.43, 6.96]	7.92% [7.40, 8.44]	4.32% [2.35, 6.82]	7.16% [6.62, 7.71]
	random forest	8.26% [5.51, 11.45]	9.96% [9.28, 10.63]	7.78% [5.11, 10.9]	7.73% [7.17, 8.28]
	xgboost	7.96% [5.26, 11.1]	9.73% [0.09, 10.38]	6.90% [4.38, 9.88]	7.67% [7.13, 8.21]
all variables	CART	8.33% [5.57, 11.52]	11.69% [10.93,12.44]	20.27% [16.33, 24.41]	13.83% [13.08, 14.59]
	ctree	14.84% [11.30, 18.69]	9.00% [8.42,9.59]	26.75% [22.52, 31.06]	12.17 [11.48, 12.87]
	linear regression	27.80% [23.55, 32.13]	9.10% [8.47, 9.72]	28.75% [24.48, 33.09]	8.99% [8.40, 9.58]
	L1	27.18% [22.95, 31.5]	10.00% [9.34, 10.66]	14.67% [11.14, 18.85]	14.83% [14.10, 15.56]
	L2	21.94% [17.91, 26.14]	10.37% [9.68, 11.05]	24.75% [20.6, 29.03]	10.51% [9.86, 11.15]
	random forest	89.65% [88.49, 90.71]	11.01% [10.34,11.67]	93.01% [92.21, 93.73]	14.95% [14.19, 15.71]
	xgboost	47.51% [43.36, 51.52]	11.97% [11.24, 12.71]	33.62% [29.28, 37.95]	15.75% [14.98, 16.52]
prior contact,	**CART**	**22.43% [18.37, 26.65]**	**11.19% [10.38, 12.00]**	**16.4% [12.72, 20.35]**	**11.80% [11.12, 12.48]**
diagnosis,	ctree	9.76% [6.79, 13.13]	9.23% [8.69, 9.78]	14.20% [10.71, 17.99]	11.39% [10.73, 12.06]
scales, CAPs	linear regression	10.07% [7.61, 14.19]	9.48% [8.89, 10.07]	13.12% [9.74, 16.83]	11.66% [10.98, 12.35]
	L1	10.07% [7.61, 14.19]	9.50% [8.91,10.09]	10.06% [7.05, 13.48]	12.59% [11.87, 13.31]
	L2	10.47% [7.41, 13.93]	9.82% [9.21, 10.43]	13.18% [9.80, 16.68]	11.79% [11.10, 12.48]
	random forest	74.74% [72.18, 77.11]	8.77% [8.16, 9.38]	69.68% [66.72, 72.44]	11.17% [10.5, 11.83]
	xgboost	28.27% [24.00, 32.60]	9.23 [8.63, 9.83]	19.91% [15.98, 24.03]	13.85% [13.11, 14.59]
prior contact,	CART	22.43% [18.37, 26.65]	11.19% [10.38, 12.00]	16.40% [12.72, 20.35]	11.77% [11.09, 12.45]
diagnosis,	ctree	9.76% [6.79, 13.13]	9.25% [8.7, 9.8]	13.33% [9.93, 17.05]	10.94% [10.29, 11.59]
scales, CAPs,	linear regression	11.17% [8.01, 14.7]	9.48% [8.89, 10.06]	13.53% [10.11, 17.28]	11.35% [10.69, 12.02]
comorbidities	L1	11.16% [8.00, 14.68]	9.53% [8.94, 10.12]	9.77% [6.80, 13.15]	12.62% [11.90, 13.34]
	L2	10.62% [7.53, 14.09]	9.91% [9.30, 10.52]	13.18% [9.80, 16.89]	11.79% [11.10, 12.48]
	random forest	69.89% [66.94, 72.63]	8.68% [8.09, 9.27]	75.55% [73.05, 77.86]	10.62% [9.99, 11.26]
	xgboost	27.9% [23.65, 32.23]	9.20% [8.60, 9.80]	21.03% [17.04, 25.20]	13.84% [13.09, 14.59]

Abbreviations: CAPs, Clinical Assessment Protocols; CART, Classification and Regression Trees; ctree, Conditional Inference Trees; xgboost, Extreme Gradient Boosted Trees

The cross-validated performance of prior contact, diagnosis, clinical scales, and CAPs subset was very close to the maximum achievable explained variance of using more variables (Table 5.3). The addition of comorbidities indicators did not appear improve predictive performance beyond prior contact, diagnosis, clinical scales, and CAPs.

Out of the two single tree algorithms, CART showed better performance than conditional inference tree. Therefore, CART models using prior contact, diagnosis, clinical scales, and CAPs to predict resource use and log-transformed resource use were analyzed more in-depth as simplest functional models with potential for implementation. These two models are subsequently referred to as CART and CART-Log, respectively. In addition, another tree model was manually built, subsequently referred to as CART-Manual, to refine the two machine learning-based simplest functional trees.

5.3.1 Simplest Functional Models

The CART model identified 30 terminal groups and the CART-Log model identified 14 terminal groups (Figures 5.2, 5.3). Not all predictors included in the subsets of prior diagnosis, clinical scales, and CAPs were used by these models. Both models identified prior contact as the most important predictor to be used as the very first split. Variables related to safety (SOS, RHO, ABS), depression (mood diagnosis, DRS, DSI), and daily functions (IADL) were common in both models, suggesting that they are important clinical indicators of health care needs. The CART model has a mean of 57.61 [95% CI: 16.72, 98.5] observations per group and the CART-Log model has a mean of 119.07 [95% CI: 14.03, 224.1] observations per group (Table 5.4, 5.5).

The last model (CART-Manual) was manually developed to predict resource use by taking insights from the CART and CART-Log model, as well as choosing the variable that split the higher clinical severity in the same direction as the higher observed resource use in daughter nodes whenever possible, restricting the minimum number of observations in each terminal group to 20 and maximum of seven splits. The CART-Manual explained 11.1% [95% CI: 8.0, 14.7] of variance in resource use (Figure 5.4), 11.3% and 5.3% less than the non-cross validated R^2 of the CART and CART-Log models, respectively. Since the model was built manually, the cross-validated R^2 was not available. The CART-Manual identified 12 terminal groups, with mean of 100.58 [95%CI: 46.49, 154.67] observations per group (Table 5.6).

The mean coefficients of variation of the CART model: 1.12 (min = 0.7, Q1 = 0.85, Q2 = 1.09, Q3 = 1.24, max = 1.98), CART-Log model: 1.39 (min = 0.58, Q1 = 1.12, Q2 = 1.39, Q3 = 1.60, max = 2.84), and CART-manual model: 1.30 (min = 0.83, Q1 = 1.12, Q2 = 1.28, Q3 = 1.41, max = 1.8)

The CMI range of the CART model was 0.35 to 3.67, a 10.5 times difference between the lowest and highest groups (Table 5.4). The CMI range of the CART-Log model was 0.27 to 5.72, a 21.2 times difference between the lowest and highest groups (Table 5.5). The CMI range of the CART-Manual mode was 0.48 to 2.41, a 5.0 times difference between the lowest and highest groups (Table 5.6).

The covariance and its confidence intervals of the group assignment and residual difference showed group assignment and residual are independent for the CART and CART-Manual models (Tables 5.4, 5.6). For the CART-Log model, the covariance of the group assignment and residual difference indicated that 2 of its terminal groups (groups 1 and 3) could have been under-reimbursed and 2 of its groups (groups 11 and 12) could have been over-reimbursed compared to the remaining groups (Table 5.5).

When stratified the sample by sex, the covariance and its confidence intervals of group assignment and residual difference indicated that sex and residual difference are independent in all models. Specifically, the covariance [95% CI] of female and residual difference for the CART model: -1.29 [-9.21, 6.14], CART-Log model: -5.44 [-14.01, 2.92], CART-Manual model: -3.25 [-11.72, 4.85].

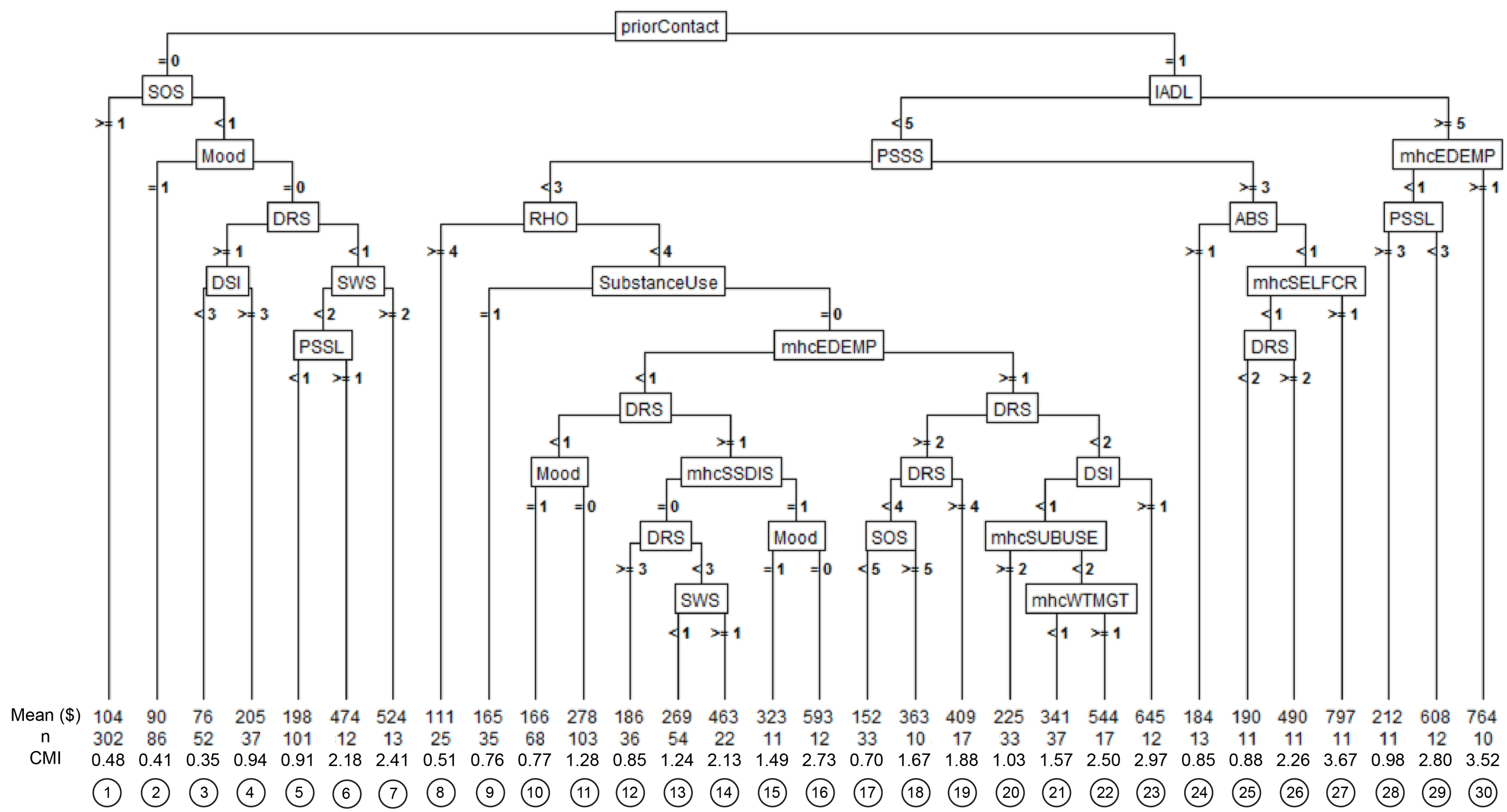

Figure 5.2 Decision tree (CART) trained using prior contact with community mental health agency, diagnosis, clinical scales, and assessment protocols to predict resource use. Each terminal node shows the predicted resource use, number of observations of the node, case-mix index, and node number

Abbreviations: ABS, Aggressive Behavior Scale; CART, Classification and Regression Tree; CMI, Case-mix Index; DRS, Depression Rating Scale; DSI, Depressive Severity Index; IADL, Instrumental Activities of Daily Living; mhcEDEMP, Education and Employment Clinical Assessment Protocol; mhcSELFCR, Self-care Clinical Assessment Protocol; mhc-SSDIS, Support Systems for Discharge Clinical Assessment Protocol; mhcSUBUSE, Substance Use Clinical Assessment Protocol; mhcWTMGT, Weight Management Clinical Assessment Protocol; PSSL, Positive Symptoms Scale Long-form; PSSS, Positive Symptoms Scale Short-form; RHO, Risk of Harm to Other Scale; SOS, Severity of Self-harm; SWS, Social Withdrawal Scale;

Table 5.4 Characteristics of terminal nodes of decision tree (CART) trained using prior contact with community mental health agency, diagnosis, clinical scales, and assessment protocols to predict resource use

Group	n	Observed Resource Use ($) Mean [95% CI]	Predicted Resource Use ($)	Coefficient of Variation	Observed CMI	Covariance(Group, Residual Difference)
Group 1	302	104 [83,125]	104	1.80	0.48	0.00 [-6.87, 6.35]
Group 2	86	90 [52,128]	90	1.98	0.41	0.00 [-4.08, 3.77]
Group 3	52	76 [46,106]	76	1.42	0.35	0.00 [-3.22, 2.98]
Group 4	37	205 [119,290]	205	1.25	0.94	0.00 [-2.73, 2.53]
Group 5	101	198 [132, 264]	198	1.68	0.91	0.00 [-4.39, 4.06]
Group 6	12	474 [195, 753]	474	0.93	2.18	0.00 [-1.57, 1.46]
Group 7	13	524 [245, 802]	524	0.88	2.41	0.00 [-1.64, 1.51]
Group 8	25	111 [66, 155]	111	0.97	0.51	0.00 [-2.26, 2.09]
Group 9	35	165 [89, 241]	165	1.34	0.76	0.00 [-2.66, 2.46]
Group 10	68	166 [118, 214]	166	1.20	0.77	0.00 [-3.66, 3.38]
Group 11	103	278 [220, 337]	278	1.08	1.28	0.00 [-4.43, 4.10]
Group 12	36	186 [119, 253]	186	1.06	0.85	0.00 [-2.70, 2.50]
Group 13	54	269 [205, 333]	269	0.88	1.24	0.00 [-3.28, 3.03]
Group 14	22	463 [319, 607]	463	0.70	2.13	0.00 [-2.12, 1.96]
Group 15	11	323 [53, 592]	323	1.24	1.49	0.00 [-1.51, 1.39]
Group 16	12	593 [134, 1053]	593	1.22	2.73	0.00 [-1.57, 1.46]
Group 17	33	152 [93, 210]	152	1.09	0.70	0.00 [-2.59, 2.39]
Group 18	10	363 [146, 579]	363	0.84	1.67	0.00 [-1.44, 1.33]
Group 19	17	409 [243, 575]	409	0.79	1.88	0.00 [-1.87, 1.73]
Group 20	33	225 [137, 312]	225	1.10	1.03	0.00 [-2.59, 2.39]
Group 21	37	341 [245, 436]	341	0.84	1.57	0.00 [-2.73, 2.53]
Group 22	17	544 [318, 769]	544	0.81	2.50	0.00 [-1.87, 1.73]
Group 23	12	645 [83, 1207]	645	1.37	2.97	0.00 [-1.57, 1.46]
Group 24	13	184 [92, 276]	184	0.83	0.85	0.00 [-1.64, 1.51]
Group 25	11	190 [37, 344]	190	1.20	0.88	0.00 [-1.51, 1.39]
Group 26	11	490 [88, 892]	490	1.22	2.26	0.00 [-1.51, 1.39]
Group 27	11	797 [325, 1270]	797	0.88	3.67	0.00 [-1.51, 1.39]
Group 28	11	212 [54, 370]	212	1.11	0.98	0.00 [-1.51, 1.39]
Group 29	12	608 [309, 906]	608	0.77	2.80	0.00 [-1.57, 1.46]
Group 30	10	764 [328, 1199]	764	0.80	3.52	0.00 [-1.44, 1.33]

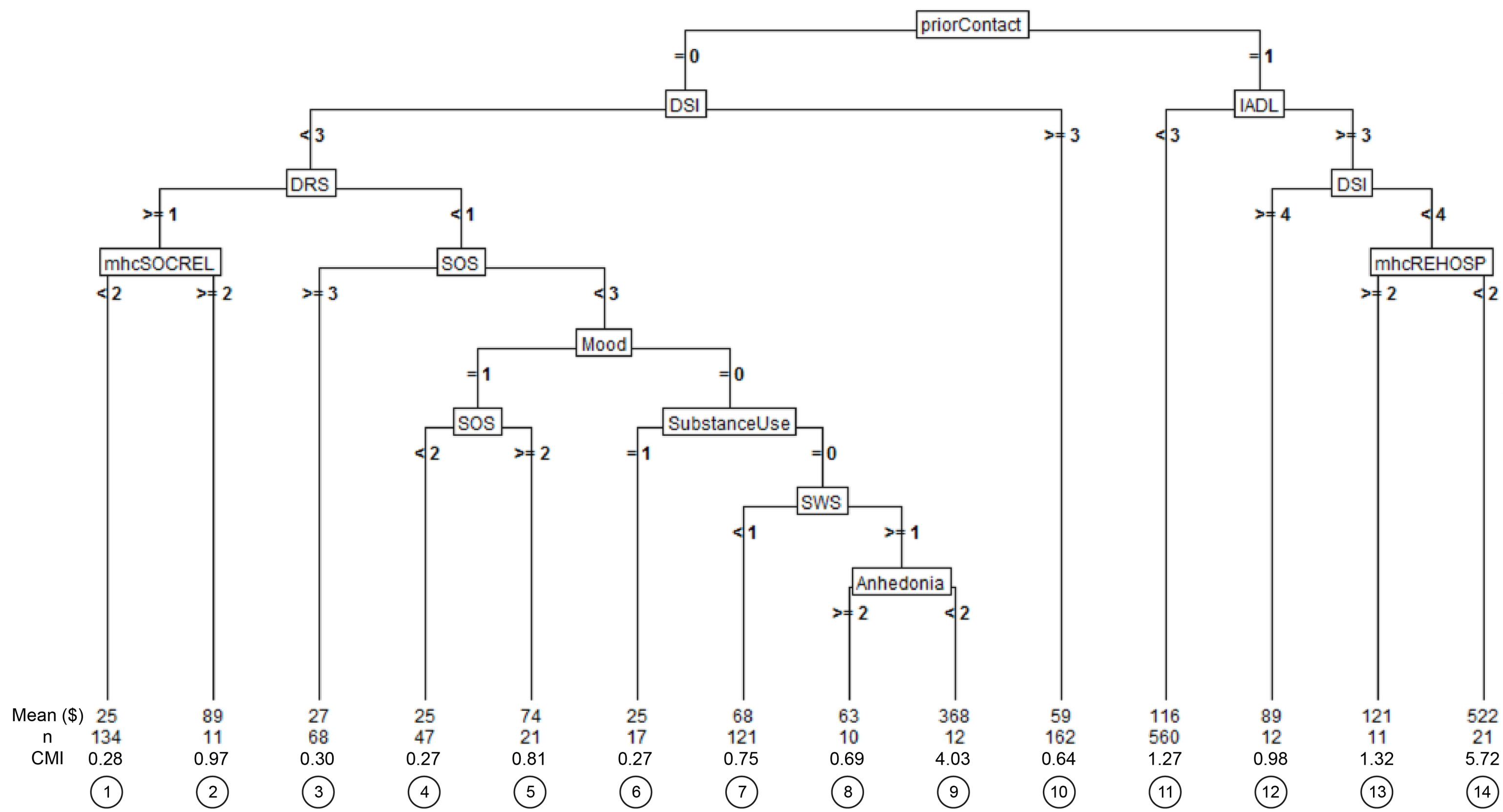

Figure 5.3 Decision tree (CART-Log) trained using prior contact with community mental health agency, diagnosis, clinical scales, and assessment protocols to predict log-transformed resource use

Abbreviations: CART, Classification and Regression Tree; CMI, Case-mix Index; DRS, Depression Rating Scale; DSI, Depressive Severity Index; IADL, Instrumental Activities of Daily Living; mhcCRIM, Criminal Activity Clinical Assessment Protocol; mhcSOCREL, Social Relationships Clinical Assessment Protocol; mhcREHOSP, Rehospitalization Clinical Assessment Protocol; SOS, Severity of Self-harm; SWS, Social Withdrawal Scale

Table 5.5 Characteristics of terminal nodes of decision tree (CART-Log) trained using prior contact with community mental health agency, diagnosis, clinical scales, and assessment protocols to predict log-transformed resource use

Group	n	Observed Resource Use ($) Mean [95% CI]	Predicted Resource Use ($)	Coefficient of Variation	Observed CMI	Covariance(Group, Residual Difference)
Group 1	134	63 [45, 80]	25	1.64	0.28	-9.83 [-15.15, -4.12]
Group 2	11	169 [58, 280]	89	0.97	0.97	-0.42 [-2.05, 1.13]
Group 3	68	61 [38, 83]	27	1.54	0.3	-5.2 [-9.13, -1.15]
Group 4	47	71 [12, 131]	25	2.84	0.27	-3.08 [-6.39, 0.23]
Group 5	21	157 [72, 242]	74	1.19	0.81	-0.75 [-2.99, 1.39]
Group 6	17	76 [13, 139]	25	1.61	0.27	-1.06 [-3.08, 0.9]
Group 7	121	223 [160, 285]	68	1.55	0.75	2.88 [-2.29, 7.43]
Group 8	10	157 [19, 294]	63	1.23	0.69	-0.27 [-1.83, 1.2]
Group 9	12	588 [308, 869]	368	0.75	4.03	0.94 [-0.77, 2.44]
Group 10	162	155 [117, 194]	59	1.58	0.64	-3.91 [-9.76, 1.81]
Group 11	560	283 [254, 312]	116	1.23	1.27	19.46 [10.96, 25.75]
Group 12	12	404 [-23, 831]	89	1.66	0.98	1.87 [0.17, 3.3]
Group 13	11	261 [71, 451]	121	1.09	1.32	0.13 [-1.5, 1.63]
Group 14	21	604 [444, 764]	522	0.58	5.72	-0.77 [-3.01, 1.37]

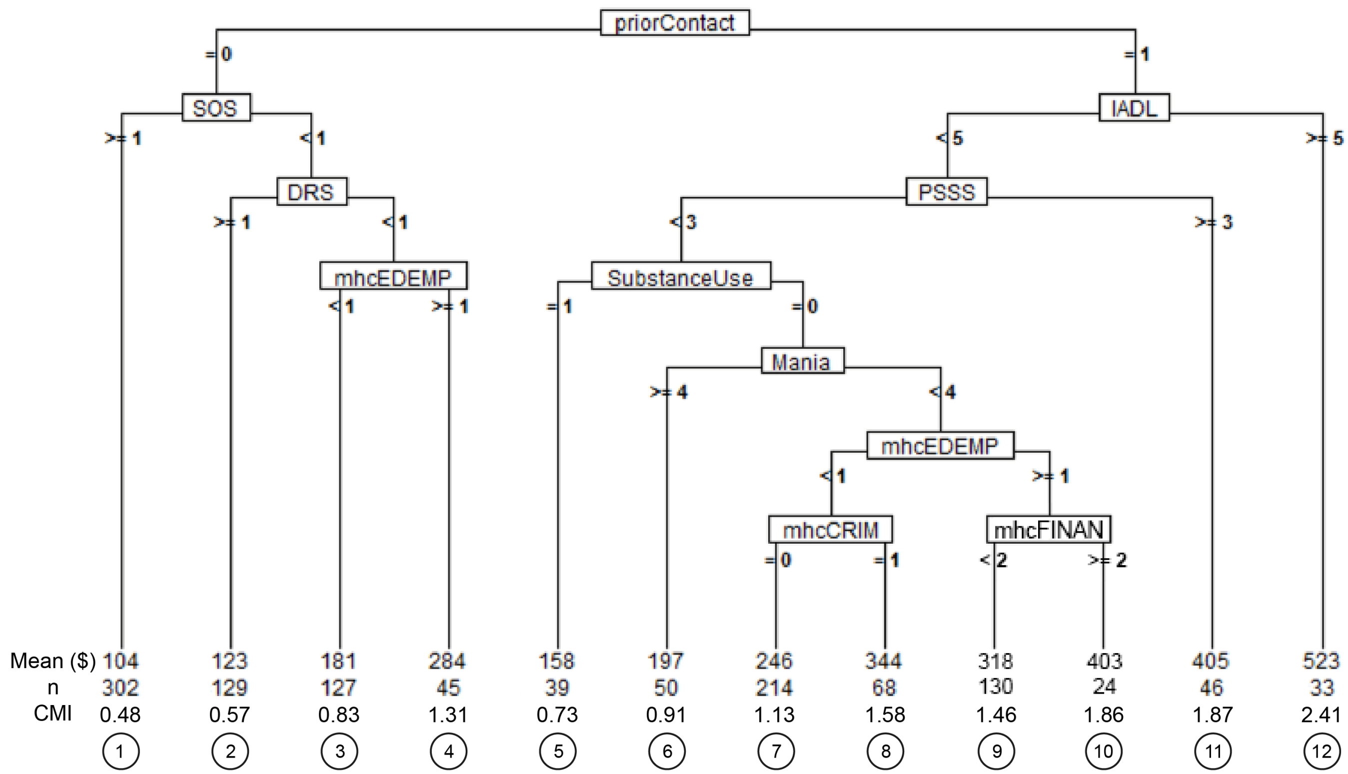

Figure 5.4 Decision tree (CART-Manual) manually built using prior contact with community mental health agency, diagnosis, clinical scales, and assessment protocols to predict resource use. Each terminal node shows the predicted resource use, number of observations of the node, case-mix index, and node number

Abbreviations: CART, Classification and Regression Tree; CMI, Case-mix Index; DRS, Depression Rating Scale; IADL, Instrumental Activities of Daily Living; mhcCRIM, Criminal Activity Clinical Assessment Protocol; mhcEDEMP, Education and Employment Clinical Assessment Protocol; mhcFINAN, Personal Finances Clinical Assessment Protocol; PSSS, Positive Symptoms Scale Short-form; SOS, Severity of Self-harm

Table 5.6 Characteristics of terminal nodes of decision tree (CART-Manual) trained using prior contact with community mental health agency, diagnosis, clinical scales, and assessment protocols to predict resource use

Group	n	Observed Resource Use ($) Mean [95% CI]	Predicted Resource Use ($)	Coefficient of Variation	Observed CMI	Covariance(Group, Residual Difference)
Group 1	302	104 [83,125]	104	1.80	0.48	0.00 [-7.36, 6.80]
Group 2	129	123 [92, 154]	123	1.45	0.57	0.00 [-5.25, 4.85]
Group 3	127	181 [124, 238]	181	1.79	0.83	0.00 [-5.21, 4.82]
Group 4	45	284 [164, 404]	284	1.40	1.31	0.00 [-3.22, 2.98]
Group 5	39	158 [89, 227]	158	1.34	0.73	0.00 [-3.00, 2.78]
Group 6	50	197 [126, 269]	197	1.27	0.91	0.00 [-3.38, 3.13]
Group 7	214	246 [210, 283]	246	1.11	1.13	0.00 [-6.49, 6.00]
Group 8	68	344 [250, 437]	344	1.12	1.58	0.00 [-3.92, 3.62]
Group 9	130	318 [249, 387]	318	1.26	1.46	0.00 [-5.26, 4.87]
Group 10	24	403 [262,544]	403	0.83	1.86	0.00 [-2.37, 2.19]
Group 11	46	405 [250,560]	405	1.29	1.87	0.00 [-3.25, 3.01]
Group 12	33	523 [345,701]	523	0.96	2.41	0.00 [-2.77, 2.56]

5.3.1.1 Calibration

All three simplest functional tree models showed sigmoidal calibration curves (Figure 5.5). Ideally, a perfectly consistent model is expected to follow the reference line ($y = x$). For observations that had lower levels of resource use, all three models over predicted. The CART and CART-Manual models appeared to be more consistent for observations in the mid- to high-range (> 50 percentile) of the resource use distribution. The CART-Log model appeared to be more consistent for observations in the mid-range of the resource use distribution (about 40-60 percentiles).

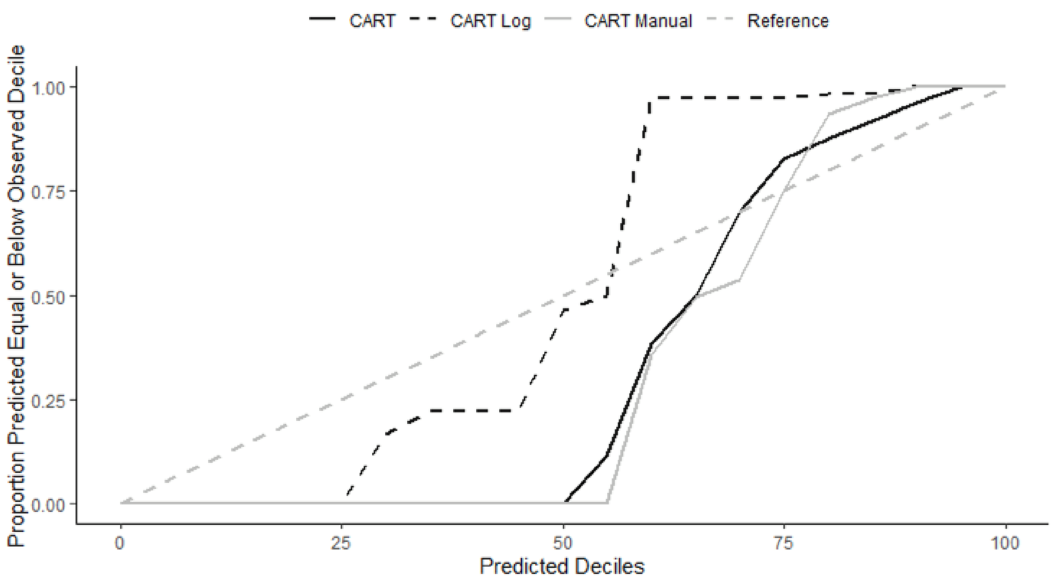

Figure 5.5 Calibration plot of decision trees trained using resource use and log-transformed resource use

5.3.1.2 Relative Resource Intensity as an Indicator of Usage of Community Mental Health Services Post-Discharge

Finally, the association between the relative resource intensity (CMI) and usage of community mental health services post-discharge were examined. Although the simplest functional models (CART, CART-Log, and CART-Manual) were developed using only the data from discharges with subsequent usage of community services, they were used to predict the CMI for all eligible discharges in this analysis. The results showed that the predicted CMI of three models were positively associated with usage of community services post-discharge, for all follow up periods between 30 to 180 days. In other words, those who were predicted to be at high resource intensity of community services were more likely to use community service services post-discharge (Table 5.7).

Table 5.7 Odds ratios in a logistic regression model of usage post-discharge within different time windows and sample selections, their corresponding 95% confidence intervals, and p-values < 0.05 indicated by *

	Usage within 30 days (True = 1504, False = 2649)	Usage within 60 days (True = 1682, False = 2481)	Usage within 90 days (True = 1787, False = 2366)	Usage within 30 days (True = 1953, False = 2200)
Observed CMI (CART)	1.80 [1.63-1.99]*	1.84 [1.67-2.04]*	1.85 [1.67-2.06]*	1.79 [1.62-1.99]*
AUROC	0.64	0.64	0.64	0.63
	Usage within 30 days (True = 1504, False = 2649)	Usage within 60 days (True = 1682, False = 2481)	Usage within 90 days (True = 1787, False = 2366)	Usage within 30 days (True = 1953, False = 2200)
Observed CMI (CART-Log)	1.39 [1.28-1.50]*	1.42 [1.31-1.54]*	1.44 [1.33-1.57]*	1.43 [1.32-1.56]*
AUROC	0.67	0.67	0.66	0.66
	Usage within 30 days (True = 1504, False = 2649)	Usage within 60 days (True = 1682, False = 2481)	Usage within 90 days (True = 1787, False = 2366)	Usage within 30 days (True = 1953, False = 2200)
Observed CMI (CART-Manual)	3.45 [2.98-4.00]*	3.51 [3.03-4.08]*	3.62 [3.11-4.22]*	3.48 [2.99-4.06]*
AUROC	0.66	0.66	0.66	0.65

5.4 Discussion

Using linked clinical data from a psychiatric hospital and resource use data from a local community mental health agency, this study was able to build models using machine learning to predict community mental health resource use with predictive performance comparable or higher than other classification systems found in the literature [113]. In addition to machine learning-based models, a model as manually developed based on insights from machine learning approach. The models were able to separate the population into relatively homogeneous groups in terms of resource intensity. The relatively resource intensity indicated by the models were also positively associated to usage of community services post-discharge (Table 5.7).

In a previous study (Chapter 4), a different classification system used for psychiatric inpatients (System for Classification of In-Patient Psychiatry) achieved an explained variance of 11.9% of resource use, and 14.1% of log-transformed resource use of community mental health services, when prior contact with the community mental health agency within a month before inpatient admission was included. Without prior contact, SCIPP achieved only 6% of explained variance.

In this study, two machine learning-based simplest decision tree models (Figures 5.2, 5.3) were able to explain 22.4% of variance of resource use, and 16.4% of variance of log-transformed resource use on the same data set used in the previous study, which is the training data for this study. The performance from the cross-validation experiments were lowered but relatively similar to SCIPP, 11.2% of variance in resource use and 11.8% of variance in log-transformed resource use (Table 5.3). An attempt to manually create a simple model achieved 11.1% of explained variance on the training data.

The models produced in this study were also simpler than SCIPP, with 30, 14, and 12 terminal groups respectively, compared to 47 groups in SCIPP. The coefficient of variation also suggested that the within-group variation of SCIPP was also more varied (mean = 2.27, min = 1.15, median = 2.26, max = 4.98) than all three simplest models in this study (Tables 5.4, 5.5, 5.6).

Although SCIPP was shown to be predictive of community mental health resource use, the model is not entire compatible for use in community settings. Specifically, SCIPP uses days of inpatient stay as one of its input variables and as an indicator of resource intensity. In a previous study (Chapter 4), there are significant differences observed in the pattern of service use and resource consumption between the inpatient and community settings, therefore the days of inpatient stay is less likely to be relevant to community settings.

Machine learning helped to produce high performing models in predicting resource use. However, the cross-validation experiments suggested that the estimated explained variance on unseen observations may be closer to 12% (Table 5.3). Although more complex algorithms, such as random and xgboost, were able to achieve very high performance on the training data, the performance in cross-validation experiments were similar or lower than that of simpler models, such as CART and linear regression. This behavior is an indication that more complex models, both in learning algorithms and input variables, may have learned a large portion of the noise in the data as well as the signal, which may have reduced their generalizability. Another possible explanation was that the training data were too small and heterogeneous. The effect of holding out a small subset of the data for cross-validation may have removed a lot of signals, which led to high sensitivity to training data.

Tree-based learning algorithms all aim to maximize the variance reduction at every split. While this strategy may result in a tree with only minimum number of splits needed to achieve good prediction, it may not necessarily be optimal clinically and risks ignoring other sequential combinations of variables that may yield better result. For example, it may be advantageous to first divide the population using clinically relevant variables (such as diagnosis) that may not necessarily yield the best variance reduction in earlier splits, but could possibly yield better variance reduction subsequently.

This study examined several subsets of input variables. The simplest subset of variables included those are almost universally available throughout the health care system: age, sex, and diagnosis, which achieved the lowest explained variance (Table 5.3). On the other hand, using all variables available did not improve explained variance in cross-validation compared to using only the clinical scales and assessment protocols. The use of cormorbidity related variables also did not have a major impact on explained variance. One possible explanation was that cormorbidity did not add additional information beyond the clinical scales and assessment protocols. For example, cormorbidity may result in more physical or cognitive impairment, which may have already been capture by the scales such as ADL, IADL for physical domains, and CPS, PSSS for cognitive domains. Overall, the results suggested that the clinical scales and assessment protocols were able to capture and summarize the overall clinical characteristics that are relevant to resource use.

Prior contact was observed to be the most relevant predictor of resource use. This is an indication that clinical information available at discharge from inpatient psychiatry alone although informative, it is not enough on its own. There are two factors that could have influenced the usefulness of clinical information from the discharge assessment. First, in a previous study (Chapter 4), it was observed that the duration between discharge and community service initiation could play a role in the decrease in explained variance, possibly due to the changes in clinical characteristics between discharge and service initiation. Second, it is also possible that persons with prior contact with the community mental health agency might already have a care plan ready to start. Therefore, their service use post-discharge was more predictable than persons who did not have prior contact or a care plan created.

Future research should aim to use clinical characteristics measured closer to or at service initiation, preferably done by the community mental health agency, to address the limitation of this study in using historical clinical measurements done prior to service initiation. Although clinical measurements from the inpatient setting alone may not be enough to predict community mental health resource use, it still has values as a differentiator of community mental health resource use, illustrated in both their abilities to predict resource use and whether a person would use community mental health services post-discharge. A classification system that could integrate both inpatient and community data may be more robust than using data solely from just one setting. On the other hand, the strength of using clinical data and resource use data from two organizations enhanced the validity of the findings because the data were measured independently without incentives to match the clinical characteristics to the resource use.

Although a case-mix classification system provides a connection between clinical characteristics and expected levels of resource use, it is one of many tools available to policy makers to achieve health outcome objectives. Differences in clinical practice, legal requirements, clinician-client fit, and preferences of the clients are also important determinants of health outcomes that are not included in the case-mix classification. Therefore, the pricing component of a funding formula should also be designed to take into account these determinants, and reward or penalize accordingly if desired health outcomes are not met. In other words, the case-mix classification system is only one component of a funding formula, and it may not be enough to achieve policy objectives by itself. The design of a pricing component is currently beyond the scope of this research.

There are several important considerations in using machine learning as a primary method for developing classification systems. The use of machine learning made an underlying assumption that the observed resource use was the same as the ideal resource that should have been allocated. There are cases where this assumption may need further consideration. For example, in a previous study (Chapter 3), it was observed that readmission to inpatient psychiatry was associated lowered usage of community mental health services. The same was observed in the CART-Log model between groups 13 and 14 (Figure 5.3). Although machine learning captured the correct observed relationship between readmission risk and community resource use, this relationship may need to be reversed by deploying more resources to those at higher risk if the goal is to break the cycle of readmission. The CART-Manual attempted to mitigate this issue by choosing to split higher clinical severity in the same direction as the higher observed resource use whenever possible. Future research needs to also consider the goals of care into the classification systems.

This study only considered direct costs in the resource use measure. There may be some client-specific indirect costs that could be partially driven by clinical severity or complexity, such as documentation and case review. Depending on the needs of a client, these indirect costs could be quite significant. Since these activities are done without the client being present, adding the variance of these costs could also be driven by differences in clinical practice among facilities and clinicians. These costs could also be adjusted for in the pricing component of a case-mix funding formula, so that the case-mix classification system remained focused on the relationship between the direct costs and clinical needs.

This study provided some preliminary evidence that machine learning may be beneficial in development a case-mix classification for community mental health services using observed data. Clinical scales and assessment protocols were shown to be informative in predicting resource use. However, a major limitation of this study was the time lag between clinical assessment and service initiation that could have reduced the predictability of the clinical measurements. Simpler models were shown to have similar predictive performance than more complex models in cross-validation. Taking into account the practical considerations, simple models such as decision tree may be the most appropriate for implementation for daily use. While machine learning can detect relevant relationships between clinical characteristics and resource use, human experts are still needed to modify observed relationships that do not reflect the goals of the health care system.

Chapter 6
Roadmap for Future Research

6.1 Summary of Contributions

This dissertation showed that community mental health care is a complex delivery setting. There is a wide range of community mental health services provided collaboratively by different organizations, each with very specific foci and target populations (Chapter 3). A centralized intake pathway provides single simplified coordinated access for all agencies in the region, and enabled this research to capture the population-level pattern of usage at the transition from inpatient psychiatry to community mental health.

About half of those discharged from the inpatient psychiatry were observed to subsequently use community mental health services within six months. The majority of cases received one in-person service appointment on a weekly or bi-weekly basis for a period of seven to eight weeks, which is significantly less that the resource consumption typically observed in other care settings like home care [31]. This research also explored different methods to capture the nature of infrequent and intermittent service use of community mental health in chapter 4. Constructing a community episode of care that sensibly summarizes the observed resource use is also unavoidable for any future community mental health case-mix studies. An episode of 90 days from the first contact with the community mental health agency post-discharge appeared to be the most appropriate approach given the data.

Since the clinical characteristics measured by the RAI-MH at discharge were observed to be closely linked with subsequent community service use, having an integrated information system with interoperable clinical assessments in both care settings can improve coordination across the continuum of care. For example, the embedded readmission clinical assessment protocol from the RAI-MH was shown to be predictive of 30-day same hospital admission. This protocol therefore can be used for care planning at the community mental health agency to break the cycle of readmission.

The association between clinical characteristics and observed community resource use was also more nuanced. In chapter 3, higher clinical severity measured by the RAI-MH at discharge was observed to be associated with usage and high usage of community mental health services. However, higher clinical severity was not always associated with higher usage if the service delivery was through group therapy or by specialized services mandated by the justice system.

To examine whether individuals with shared clinical characteristics would have relatively similar levels of resource use, chapter 4 tested two high performing case-mix classification systems for their potential to predict community mental health service use: the Systems for Classification of In-Patient Psychiatry (SCIPP) and the Australian Mental Health Case-mix Classification (AMHCC). The results found that SCIPP is better than diagnosis alone in explaining the variance in resource use of community mental health services. SCIPP was also observed to be equitable in allocating resources across its classification groups. However, the biggest differentiator of resource use was whether someone had prior contact with the community mental health agency before their inpatient psychiatric episode, which also enhanced the performance of SCIPP when used together. The relative resource intensity weights of SCIPP appeared to be not well calibrated to predict resource use observed in community settings. That said, the study data were not ideal for the evaluation of SCIPP because there was a gap in resource use data for addictions and senior mental health services. This attenuation of the distribution of SCIPP categories could have underestimated the effectiveness of SCIPP. The AMHCC was not immediately operationalizable outside of the Australian context because most of its explained variance depended on a subjective classification of the needs of care and goal of care (referred to as phases of care) by a clinician at the beginning of a community episode.

In addition, there are also fundamental differences between inpatient psychiatry and community mental health that may also contributed to the limited of transferability of SCIPP. The needs that the two care settings address share some similarities but also differ in several ways. Inpatient psychiatry has a historical origin in the model that aimed to cure illnesses [8]. The trend of deinstitutionalization contributed to a shift toward crisis management and stabilization, so that a person can be discharged for outpatient or community services [8]. On the other hand, community mental health is rooted in the recovery model, which aims to support and improve a person's ability to function outside of the health care system [131]. Therefore, the scope of needs that community mental health services addresses can be broader than inpatient psychiatry. By extension, the pattern of resource use of the two settings may be different.

Empirically, inpatient psychiatry was shown to have much higher resource use intensity than community mental health [27]. A narrower range of observed resource use in community mental health may also contribute to the difficulty in explaining variance when the variance is not as wide as observed in inpatient psychiatry.

In chapter 5, machine learning algorithms were used to experiment with building alternative case-mix classification systems that could be used for predicting community mental health resource use. All models were able to reduce the variance in observed resource use, but the limit of the current data set appeared to be at about 12% of explained variance. Three exploratory decision tree models, two were machine learning-based and one was manually developed. These were proposed because they were able to achieve similar explained variance as more complex models, and more user-friendly to clinicians. While machine learning was able to produce models that described the resource use patterns in the observed data, more research is needed to align the models with the goals of the health care system. For example, in chapter 3, it was observed that persons at risk of self-harm may not necessarily be more resource intensive if they were provided services in a group setting. Similarly, the risk of readmission can reduce the observed resource intensity of community services, therefore the observed data may not reflect the true resource intensity of readmitted cases. Clinical and health services expertise are still required in development of future case-mix classification systems. The results pointed toward the need for more contemporaneous clinical data that are closer to the community mental health service initiation in future development work.

Classification of community mental health services remained a difficult problem. There are several reasons why the observed explained variance for community mental health services is often lower than in other types of health care services, such as inpatient psychiatry or home care. The scope of care of community mental health can be much broader than other types of care, such as inpatient psychiatry or home care services. Community mental health services has its roots in providing recovery-oriented services [8]. As a result, there are also other factors beyond the interactions between a person and the health care system that may affect resource use and outcomes, such as environmental triggers, employment opportunities, or discrimination experienced by the client. In comparison, the scope of inpatient psychiatry focuses more on stabilization of symptoms and crisis management. Other community-based health care services like home care also has a narrower scope, in which services are more related to daily functional tasks such as getting dressed, meal preparation, and mobility.

Although clinical characteristics have some explanatory power for resource use, there are also other important drivers of mental health outcomes that may not exist in other types of health care services, such as fit between clinicians and clients, and preferences of clients.

The sample of this research included only a portion of the population who were previously discharged from inpatient psychiatry, which could have higher clinical needs than the general population. Therefore, the distribution of clinical needs was likely attenuated that could result in low explained variance.

The use of the clinical characteristics measured at the time of assessment, such as in this research, only provided a static snapshot of a person's clinical needs. While clinical severity or complexity are important considerations in resource allocation, clinical chronicity is also important, especially in the tiered frameworks for planning substance use service delivery systems [132]. In addition to performing frequent re-assessment, a case-mix classification system could also use the changes between assessments to take into account the trajectory and nuanced dynamic of the clinical changes over time. For example, an ideal system would combine clinical information from the point of discharge from hospital with current clinical assessment at the time of initiation of community mental health services (with on-going re-assessment if the person has a prolonged episode of community-based service use).

A case-mix classification system compatible with community mental health services is essential to the implementation of Ontario's proposed funding reform. More research is needed to develop case-mix classification systems for community mental health settings (Chapter 2) because even the best available systems are very limited when externally validated in the context of this research (Chapter 4). While data at the point of discharge from inpatient psychiatry were relevant to the pattern of resource use in community mental health services, they were insufficient to provide a robust estimate of resource use on their own.

This research showed that it is possible to separate a population into relatively homogeneous groups in term of community mental health resource use with individual-level clinical characteristics. Despite using a study sample from one of the largest community mental health agencies in Canada, more high quality and contemporaneous clinical data is needed to improve the explained variance of resource use. Data standards that support interoperability across multiple mental care settings can improve clinical meaningfulness and predictive performance of case-mix classification of community mental health services. The need for standards also applies to management information systems that capture resource use at community mental health agencies. These insights support a larger effort to develop a case-mix classification system for community mental health services.

On the other hand, the generalizability of this research is limited by the available data. The sample used in this research only represented people who were previously discharged from inpatient psychiatry and subsequently used community mental health services. Missing from this sample are the child/youth population, geriatric mental health services or primary care provided by external agencies, and adults who were not previously admitted the inpatient psychiatry unit of this research. The sample of this research only represents a portion of the population who are likely to be at higher clinical severity and complexity compared to the population of community mental health service clients. The observed relationships between clinical characteristics and resource use from this study may not be generalizable beyond the service pathway examined by this research, which was from inpatient psychiatry to community mental health services. Therefore, development of a case-mix based funding formula for community mental health will require additional research.

6.2 Roadmap for Development

The development of a case-mix classification system requires two essential sources of individual-level data that must be brought together: the clinical data and resource use data. In this research, these two sources of data were successfully brought together for only one community mental health agency. Data from one agency, despite being one of the largest in Canada, may not be representative of all cases observed in the province or the country. This was indeed observed in Chapters 3 and 4, in which the clinical characteristics at discharge of the study sample were not similar to the rest of the province.

6.2.1 Representativeness of Study Sample

Recruiting every community mental health agencies in the province or the country for a case-mix development study may not be possible or necessary due to the complexities, cost, time, and labor required. Most case-mix classification systems were developed using a sample of their respective jurisdictions. For example, the SCIPP was developed using a sample of psychiatric units across Canada [95]. The Canadian Case-Mix Groups+, used for non-psychiatric inpatient care, was also developed using a sample of hospitals from only 3 provinces [22].

The number of agencies recruited should be large but is not the only criterion. The agencies recruited should represent a wide range of mental health services offered. As observed in this research, within a jurisdiction, there may be several agencies offering publicly-funded community mental health services and each may be specialized in a specific area. The agency that was examined in this research offered a wide range of services; however, they did not offer addiction services, and senior mental health services were offered in partnership with another agency. The limited availability of standardized clinical assessment data also narrowed the study sample to only persons who were previously discharged from inpatient psychiatry.

Classification of eligible community mental health services is a major difficulty. As observed in this research, services offered by joint venture of more than one organizations may pose a barrier to gather the necessary data for case-mix classification development. Future developers will need to negotiate across organizational boundaries to obtain the data required. Cross-organizational issues will also arise for mental health services delivered through primary care or integrated care model [133]. For example, services delivered by psychiatrists can be funded based on fee-for-service similar to family physicians, and they may have their own practice that is not part of a community mental health agency.

Additionally, this research only captured the service pathway from inpatient psychiatry to community mental health services, and not other service pathways, such as to primary care or private counseling services. The mix of eligible services offered is also likely to be different for across jurisdictions. For example, housing, employment, and education have been suggested to be an important part in mental health recovery and treatment [80]. However, funding for assistant services for housing, employment, and education are often separated from funding for health care. Therefore, the types of services included in the development of a community mental health case-mix classification are also dependent on the design of the local health care systems.

Future developers should perform a comprehensive review of community mental health services. Then, the future developers should assemble a panel of stakeholders that can provide inputs into the inclusion and exclusion criteria of services. This panel of stakeholders should be an ongoing advisory panel for both the initial development and ongoing maintenance of the case-mix classification systems. For example, the Diagnosis Related Groups used by the US Medicare system for inpatient care relies on their Medical Payment Advisory Commission for their annual review and adjustment [134]. The advisory panel should include experienced clinicians, funding authority administrators, statisticians, and health services researchers.

Although the services that may be classified as community mental health services will be subject to debate, the future developers should focus on the reaching a consensus, rather than completeness. It should be emphasized that the inclusion of services will evolve over time and the case-mix classification does not have to be static. For example, the Resource Utilization Groups that is used in long-term care has gone through many updates and revisions throughout the years [32].

One potential solution to reconcile the differences in eligible services may be to identify of set of core services that are commonly eligible for health care funding in most jurisdictions. Through the advisory panel, services that were not included in the early versions may be added in future revisions if there is wider acceptance.

Another approach may be to focus on the inclusion of the type of cases that the advisory panel aims to target. For example, this research only focus on one small segment of all community mental health cases, which was individuals previously discharged from inpatient psychiatry. Alternatively, a more inclusive target population may be to include all individuals who were served by the central community mental health services intake (Here 24/7) in the Waterloo-Wellington region.

6.2.2 Individual-level Clinical Data

The role of the individual-level clinical data is to describe the individuals who use community mental health services at the individual level. In the development of case-mix classification, it is used to separate the population into clinically homogeneous groups, when used together with the resource use data during the development of case-mix classification systems. The individual-level data must also be standardized across the recruited agencies. At a minimum, the clinical data must allow crosswalk (the mapping of data equivalence across two or more data standards) across agencies (also known as interoperability).

In many data standards, the record of an individual may also contain data elements that describe the providers and processes (services or treatment provided). In chapter 2, the appropriateness of variables used as input of the case-mix classification system was discussed extensively. In summary, variables that describe the clinical needs of an individual are desirable for case-mix classification systems. These are variables that directly drive the resource use, such as: diagnosis, clinical status, or clinical severity. Variables describing the provider should be avoided because they account for the variance in observed costs across facilities. The observed variance may be due to systematic inequalities across the facilities not related to the clinical needs of a person, therefore relying on these variables may lead to reinforcing those inequalities. Variables describing the services or treatments should also be avoided because they can have the effect of incentivizing service or treatment volume for financial gains. Historical variables describing a prior usage of health care services are not ideal because they are not modifiable and change with current health care needs. Historical variables describing an individual can be useful if they are relevance to current health care needs (e.g., past history of abuse or violence).

While there may be numerous data standards available for community mental health, the choice of data standards will have different implications for the health care system. The simplest and most universally available may be diagnosis and demographics (age, sex) information. In the review of chapter 2, although the use of diagnosis as the sole variable of differentiating levels resource use was common, the high performing case-mix classification systems relied on additional clinical variables beyond diagnosis. Similarly, it was observed in chapters 4 and 5 that diagnosis alone accounted for smaller amount of variance in resource use compared to when used in addition with other clinical variables.

Age and sex are sometimes used as predictors of resource use but they should avoided if possible. The observed associations of demographic variables with resource use may be the reflection of other causes, either social or clinical [135]. Therefore, there is a risk of reinforcing existing biases that could exclude demographic groups from accessing mental health care if demographics variables are used. Replacing demographic variables with variables indicating clinical severity or complexity during the case-mix classification development could alleviate this issue. Additionally, in chapter 5, using solely diagnosis and demographics was shown to be less predictive of community mental health resource use than using diagnosis in combination with variables indicating clinical severity or complexity.

There is also opportunity for the future developers to play a role in influencing clinical practice through the choice of data standard. For example, to promote integration of care, the future developers may consider a data standards that is designed for care integration and compatible across settings, as suggested by chapter 3. In Ontario, the RAI-MH has been the data standard for inpatient psychiatry, therefore a sensible option that could promote care integration is the interRAI CMH, which belong to the same integrated suite of mental health assessments as the RAI-MH [52].

Promoting integration of care through data standards can also benefit the development of case-mix classification system. Although chapters 4 and 5 showed that the predictive utility of clinical characteristics measured at discharge could be masked by whether a person was known to the community mental health agency and delay between discharge and service initiation, the discharge assessment still offer some predictive utility. Therefore, a case-mix classification system that can leverage information from both the inpatient and community settings by simply using an integrated data standard could see gains in both predictive utility and interoperability.

6.2.3 Resource Use Data

The role of the resource use data is to capture the cost of health care resources that a person was provided for their care. The resource use data is then used in conjunction with individual-level clinical data to partition the population into groups of homogeneous clinical characteristics and resource use.

Not all costs incurred by a provider should be included in the resource use data for the purpose of developing a case-mix classification system. The costs of care used for case-mix classification should be viewed from the perspective of the persons receiving health care services, hence referred to as resource use. A barrier to getting of full picture of relevant costs for this study was the limited data available. Future developers will encounter this challenge in negotiating which costs are relevant for case-mix classification.

Accounting principles broadly classify costs incurred by the providers into direct and indirect costs [136]. Direct costs can be considered as the costs required to provide care for one additional person. Indirect costs are costs that are required to operate an health care organization, shared by many activities, and cannot be attributed to the care provided for one additional person. Only the direct costs should be included in the resource use data because the goal is to capture the costs that are driven by an individual's clinical characteristics.

For some patient-specific indirect costs (such as documentation, or case review), since they are often performed without a client present, considerations should be given to potential variance related practice patterns across facilities and clinicians that can also be added if these costs are included. Alternatively, these costs can also be taken into account outside of the case-mix classification system through the pricing policy mechanisms, so that the case-mix classification stays focused on the connection between direct costs and clinical needs.

Future developers have two approaches in measuring the resource use: bottom-up or top-down. The bottom-up approach is similar to the approach used in this research, in which the resource use for an episode was aggregated from costs of each direct service events. Cost of each direct service was simply the staff time spent providing service multiplied by the median wage for the staff position at the agency. While using the actual wage may add precision, there is also a risk that inequality in wages can contribute to variance in observed resource. Additionally, the staff time spent providing service was recorded for administrative purposes. While a more rigorous approach is to measure the staff time directly via observation as done in the SCIPP development study [27], it may not be necessary if services are often time-limited office visits.

The top-down approach is to estimate the resource use by allocating costs of functional centers to the individuals who received care equally. Functional centers are smaller units within an organization, such as departments [136]. This approach should be avoided because the differentiator of resource intensity is whether someone received services within a certain accounting period and how many others also received services. Therefore, this approach cannot differentiate resource intensity that is driven by an individual's clinical needs.

For jurisdictions that are currently using fee-for-service for community mental health instead of global budget, it may also be possible to use historical claim costs to estimate resource use. However, claims tend to be higher than the actual costs (both direct and indirect) incurred by the providers in order for the providers to be financially viable. Additionally, claims are not likely to represent the actual resource use of individuals receiving care, but rather an average over large amount of individuals [27].

This research examined the observed resource use data and determined that the most appropriate and practical method of constructing an episode is a period of 90 days from the first contact with the community mental health agency. The immediate implication for clinical practice is that, upon completion of an episode, the community mental health agency should perform a new assessment to determine whether the expected resource use of the person has changed. Future developers should also review the episode construction and its implication with the advisory panel for feedback and support.

6.2.4 Silent Deployment

Although it is possible to develop a model with excellent generalizability on the first attempt, it may be unlikely due to several reasons. First, if case-mix funding can provide incentives towards more efficient care as suggested by the literature, it is expected that the pattern of observed resource use across the health care system may change over time. Therefore, older models cannot be expected to perform after the overall pattern of resource use has shifted, but at the same time the clinical characteristic distributions have not. Second, the nature of community mental health is still evolving. New services are added over the years. For example, Assertive Community Treatment used to be experimental, but became more common over time. As deinstitutionalization continues to evolve, more intensive services may also be shifted to the community settings. Development of case-mix classification system, therefore, should be an ongoing commitment.

The use of algorithms in automating decisions in health care are subject to greater scrutiny recently due to the potential of reinforcing undesirable biases, which is also partly also due to the emerging abundance and ease of use of artificial intelligence [137, 138]. Health care resource allocation is not immune from this, given the recent finding that a popular resource allocation algorithm was observed to be racially biased [119].

Silent deployment, in which a candidate model is used prospectively and makes predictions in real-time visible only to a selected number of clinical experts but not acted upon, was suggested as a responsible approach in applying algorithms in clinical practice [139]. Silent deployment can be used as a prospective validation study at a smaller scale when it allows the model developers to review predictions and errors immediately, consult with the advisory panel, and make corrections if necessary. For example, expected association between risk of self-harm and high resource use were not observed immediately in chapter 3 due to the services delivered via group therapy. Such unexpected relationships require domain knowledge from the advisory panel to uncover and adjust.

6.2.5 Pilot Studies and Broad Consultation

Once the developers gained confidence of the candidate model, they should seek support from a broad range of stakeholders in their targeted jurisdictions. The best way to showcase a new product is to let the providers try out the candidate model via pilot studies. Pilots studies are different than silent deployment because the predictions are acted upon at small scale. These studies will allow the providers to experience the new reality, while still able to influence the final form of the funding formula.

The providers are not the only stakeholders. Others may include staff from funding authorities, clinicians and their unions, information technology specialists, finance specialists, and health services researchers. Support from a broad range of stakeholders is required for broad adoption of a case-mix classification system [22]. Therefore, a broad consultation beyond the advisory panel is needed during the validation studies. For example, one model of broad consultation is the mental health case-mix development in Australia. It follows a process that both tests their proposed model at pilot sites and collects feedback [46].

6.2.6 Maintenance

One indicator that could hint at the need for adjustment of the case-mix classification system is the stability of the observed relative resource intensity of the terminal groups, or CMI [22]. On the other hand, the more frequently the adjustments are made, the more instability is introduced. For hospital funding, it is estimated that the detection of changes in resource use pattern due to an adjustment can lag between two to three years after the introduction [22].

Results obtained from both the silent deployment and pilot studies have one weakness in common, which is that they both reflect the resource use patterns prior or during the adoption of the case-mix funding. As mentioned, the expected beneficial gain in cost efficiency may lag behind the introduction of the new funding model [22], which may not be observed in the pilot studies. Additionally, there are potential changes in the resource use pattern that may not be observable during the pilot studies, such as: a real change in costs due to practice innovations or technology change, changes in methods of capturing resource use, higher signal to noise ratio due to larger sample size, or improved in data quality over time after the initial training [22].

There may be also jurisdictional specific considerations that require adjustments. For example, the first-ever case-mix classification system, Diagnosis Related Groups [59], has been maintained and continuously developed in many jurisdictions beyond its origin. In Canada, it was modified to become the Case Mix Groups+, the Australian Refined-DRG in Australia, NordDRG in the Nordic countries, German-DRG in Germany, and Health Resource Group in the United Kingdom.

Adjustments to the case-mix classification should also be performed in relation to desired health outcomes. If a classification group is observed to have worsening outcome metrics, it may be an indicator that more funding is needed to improve health outcome for that particular group. In cases that have sufficient supporting evidence, one option is to keep the case-mix classification structure the same, but readjust the CMI to divert more resource as deemed appropriate.

There may also be a need for adjustments due to differences in health care systems, culture, or practice patterns in different jurisdictions. On the other hand, adjustments to a case-mix system should be carefully considered to avoid degrading the integrity of the funding formula. First, proposed changes to a case-mix system should be based on valid reasons, which are supported by independent evaluation of the observed data. Another strategy that could discourage unnecessary changes that are intended to game the system is to keep the structure of the case-mix classification the same, and only allowing changes to the case-mix index values.

6.3 Implications for Health System Policy

A case-mix classification system may only be one component of the health care funding formula. However, it changes the relationships between the funders and the providers [22]. The funders used to be responsible for the costs of health care services incurred by the providers under global budget or fee-for-service. Under case-mix funding, the role of the funders became more of a purchaser of services that makes payment based on the clinical needs of the individuals that use health care services.

In Ontario, this transition has already taken placed for the hospitals. The approach has been to slowly reduce the portion of funding received via global budget, and increase the portion of funding received via case-mix classification until they reach the target mix [26]. Feedback and lessons learned from the hospital funding transition should be incorporated in the planning of future transition of community mental health funding, such as: the target mix of global budget and case-mix funding (if any), and the length of the transition period.

Substantial training will be required for administrators of both the funders and the providers [22]. The funders will be relied upon for clear and comprehensive guidance, therefore they must be well-versed in the data requirements of both individual clinical data and resource use data. The providers will be required to adhere to data and information system standards in order to receive funding. However, the success of a case-mix classification also requires support from partners beyond the funders and providers, such as: interRAI and Canadian Institute for Health Information.

6.3.1 Role of interRAI

The choice of the individual-level clinical data to be used as the input a community mental health case-mix classification has not been decided. The interRAI Community Mental Health is a potential candidate to support such as system. However, the adoption of this assessment instrument requires a broad consensus across Ontario. interRAI can play an advocacy role for the adoption in two ways.

First, this research showed that the many clinical measurements available from discharge of inpatient psychiatry, which are also available from interRAI CMH, were closely linked to community mental health resource use. However, the value proposition of the assessment is more than just an administrative data collection tool for case-mix classification, and should be promoted by interRAI. Specifically, the assessment is primarily intended to be used for care planning, and secondarily as outcome measures and quality indicators. Data that are used beyond case-mix application alone can offset the administrative burden, and the effect of up-coding clinical characteristics for financial gain [18]. Second, the funders and providers of community mental health services will require training in data entry, usage, and interpretation of the assessment. interRAI should support this training through inputs to future case-mix developers regarding training curricula.

6.3.2 Role of Canadian Institute for Health Information

The Canadian Institute for Health Information (CIHI) is the organization responsible "the development and maintenance of comprehensive and integrated health information" across Canada [140]. Their support for a proposed case-mix classification system is crucial because CIHI is also a data standards organization that set national standards for health system data across Canada. This includes both essential data inputs of case-mix classification system - individual level clinical data and resource use data.

For ongoing maintenance, support required from CIHI will be similar as for existing case-mix classification systems as the technical stewards for data standards, storage, and independent analysis. Funders and providers will look towards CIHI for objective evidence of explained variance of the case-mix classification system, equitable allocation, expected cost efficiency gain, and overall recommendation for implementation.

6.4 Conclusions

This research supported the effort of health care funding reform in Ontario by exploring potential solutions for a case-mix classification system that is one component of a case-mix funding formula. The findings suggest that a case-mix system based on similar data standards as the inpatient psychiatry can both promote integration of care across settings and offer predictive utility of community mental health resource use. The findings suggest that a case-mix system based on similar data standards as the inpatient psychiatry can both promote integration of care across settings and predict community resource use. The extent of this research was limited by the available data. However, the results suggested that the approach of this research is appropriate and sensible, and there is a need for substantial collaboration of stakeholders across the health care system for future development work.

References

[1] D. Vigo, G. Thornicroft, and R. Atun, "Estimating the true global burden of mental illness," *The Lancet Psychiatry*, vol. 3, no. 2, pp. 171–178, 2016.

[2] E. Hewlett and V. Moran, *Making Mental Health Count: The Social and Economic Costs of Neglecting Mental Health Care*. 2014.

[3] P. Smetanin, C. Briante, D. Stiff, S. Ahmad, and M. Khan, *The life and economic impact of major mental illnesses in Canada*. Mental Health Commission of Canada, 2015.

[4] K. L. Lim, P. Jacobs, A. Ohinmaa, D. Schopflocher, and C. S. Dewa, "A new population-based measure of the economic burden of mental illness in canada," *Chronic Dis Can*, vol. 28, no. 3, pp. 92–98, 2008.

[5] Public Health Agency of Canada, "Economic burden of illness in canada custom report." `http://cost-illness.canada.ca/custom-personnalise/national.php`. Accessed: 2018-11-19.

[6] World Health Organization, "Promoting mental health: concepts, emerging evidence, practice (summary report)," 2004.

[7] S. Galderisi, A. Heinz, M. Kastrup, J. Beezhold, and N. Sartorius, "Toward a new definition of mental health," *World Psychiatry*, vol. 14, no. 2, pp. 231–22, 2015.

[8] E. Lin, Z. Or, M. Coldefy, K. Urbanoski, D. Seitz, C. Carlisle, P. Szatmari, and P. Kurdyak, "Medical practice variations in mental health and addictions care," *Medical Practice Variations*, pp. 1–41, 2016.

[9] E. Shorter *et al.*, *A historical dictionary of psychiatry*. Oxford University Press, 2005.

[10] D. Mechanic and D. A. Rochefort, "Deinstitutionalization: An appraisal of reform," *Annual Review of Sociology*, vol. 16, no. 1, pp. 301–327, 1990.

[11] G. N. Grob, "Community mental health policy in america: Lessons learned," *Mental Health Reform and a Window of Opportunity for Community Mental Health*, vol. 53, no. 1, p. 6, 2016.

[12] E. Shorter, *A history of psychiatry: from the era of the asylum to the age of Prozac*. Wiley, 2005.

[13] W. H. Organization *et al.*, "Policies and practices for mental health in europe: meeting the challenges," 2008.

[14] Mental Health Commission of Canada, "Changing directions, changing lives: The mental health strategy for canada," 2012.

[15] M. Knapp, "Economic evaluation and mental health: sparse past... fertile future?," *The journal of mental health policy and economics*, vol. 2, no. 4, pp. 163–167, 1999.

[16] Canadian Mental Health Association - Waterloo Wellington, "Annual report 2017/2018," 2018.

[17] Waterloo Wellington Local Health Integration Network, "Service accountability agreement with canadian mental health association - waterloo wellington," 2018.

[18] M. L. James, B. E. Fries, M. Goodell, and N. Wellens, "Design principles for HCBS case mix: a primer," *Ann Arbor, MI: University of Michigan, Institute of Gerontology*, 2015.

[19] J. M. Sutherland, N. Repin, and R. T. Crump, *Reviewing the potential roles of financial incentives for funding healthcare in Canada.* Canadian Foundation for Healthcare Improvement Canada, 2012.

[20] Ontario Health Insurance Plan, "OHIP schedule of benefits and fees." `http://www.health.gov.on.ca/en/pro/programs/ohip/sob/`.

[21] S. A. Schroeder and W. Frist, "Phasing out fee-for-service payment," *The New England journal of medicine*, vol. 368, no. 21, p. 2029, 2013.

[22] S. J. Duckett, T. Jackson, J. Hatcher, H. Richards, and K. Murphy, *The why, the what and the how of activity-based funding in Canada: a resource for health system funders and hospital managers.* Canadian Institute for Health Information, 2014.

[23] A. P. Costa, J. W. Poss, and I. McKillop, "Contemplating case mix: A primer on case mix classification and management," in *Healthcare management forum*, vol. 28, pp. 12–15, SAGE Publications Sage CA: Los Angeles, CA, 2015.

[24] M. B. Rosenthal and R. G. Frank, "What is the empirical basis for paying for quality in health care?," *Medical Care Research and Review*, vol. 63, no. 2, pp. 135–157, 2006.

[25] M. Kirby, "The health of canadians: The federal role. final report. volume six: Recommendations for reform," 2002.

[26] Ministry of Health and Long Term Care, *A Guide to Patient-Based Funding.* Government of Ontario, 2012.

[27] J. Hirdes, B. Fries, C. Botz, C. Ensley, M. Marhaba, and E. Pérez, "The system for classification of in-patient psychiatry (SCIPP): a new case-mix methodology for mental health," *University of Waterloo*, 2002.

[28] H. Sanderson, L. Mountney, and P. Lees, *Casemix for all.* CRC Press, 2018.

[29] M. A. Björkgren, B. E. Fries, and L. R. Shugarman, "A RUG-III case-mix system for home care," *Canadian Journal on Aging/La Revue Canadienne du Vieillissement*, vol. 19, no. S2, pp. 106–125, 2000.

[30] B. E. Fries, D. P. Schneider, W. J. Foley, M. Gavazzi, R. Burke, and E. Cornelius, "Refining a case-mix measure for nursing homes: Resource utilization groups (RUG-III)," *Medical care*, pp. 668–685, 1994.

[31] J. W. Poss, J. P. Hirdes, B. E. Fries, I. McKillop, and M. Chase, "Validation of resource utilization groups version III for home care (RUG-III/HC): evidence from a canadian home care jurisdiction," *Medical care*, pp. 380–387, 2008.

[32] L. A. Turcotte, J. Poss, B. Fries, and J. P. Hirdes, "An overview of international staff time measurement validation studies of the RUG-III case-mix system," *Health services insights*, vol. 12, p. 1178632919827926, 2019.

[33] C. for Medicare & Medicaid Services, "OASIS user manuals," 2014.

[34] M. C. Hornbrook, "Hospital case mix: its definition, measurement and use: Part i. the conceptual framework," *Medical Care Review*, vol. 39, no. 1, pp. 1–43, 1982.

[35] B. Buckingham, P. Burgess, S. Solomon, J. Pirkis, and K. Eagar, *Developing a casemix classification for mental health services: summary.* Commonwealth Department of Health and Family Services, 1998.

[36] J. Jones, F. Amaddeo, C. Barbui, and M. Tansella, "Predicting costs of mental health care: a critical literature review," *Psychological Medicine*, vol. 37, no. 4, pp. 467–477, 2007.

[37] R. C. Hermann, C. K. Rollins, and J. A. Chan, "Risk-adjusting outcomes of mental health and substance-related care: a review of the literature," *Harvard review of psychiatry*, vol. 15, no. 2, pp. 52–69, 2007.

[38] A. Mason, M. Goddard, *et al.*, "Payment by results in mental health: a review of the international literature and an economic assessment of the approach in the english nhs," tech. rep., 2009.

[39] M. Harris, N. Legge, S. Diminic, G. Carstensen, G. McKeon, D. Siskind, P. Burgess, G. Stewart, and H. Whiteford, "Mental health service cost drivers: an international literature review. final report for stage b of the definition and cost drivers for mental health services project. volume 2," 2013.

[40] H. Arksey and L. O'Malley, "Scoping studies: towards a methodological framework," *International journal of social research methodology*, vol. 8, no. 1, pp. 19–32, 2005.

[41] A. C. Tricco, E. Lillie, W. Zarin, K. K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D. Peters, T. Horsley, L. Weeks, *et al.*, "PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation," *Annals of internal medicine*, vol. 169, no. 7, pp. 467–473, 2018.

[42] World Health Organization, "The world health report 2000 - health systems: improving performance," 2000.

[43] G. Dunn, M. Mirandola, F. Amaddeo, and M. Tansella, "Describing, explaining or predicting mental health care costs: a guide to regression models: methodological review," *The British journal of psychiatry*, vol. 183, no. 5, pp. 398–404, 2003.

[44] M. Mirandola, F. Amaddeo, G. Dunn, and M. Tansella, "The effect of previous psychiatric history on the cost of care: a comparison of various regression models," *Acta Psychiatrica Scandinavica*, vol. 109, no. 2, pp. 132–139, 2004.

[45] P. Gaines, A. Bower, B. Buckingham, K. Eagar, P. Burgess, and J. Green, "New zealand mental health classification and outcomes study: Final report," 2003.

[46] I. H. P. Authority, "Development of the australian mental health care classification – public consultation paper 2," 2015.

[47] R. F. Averill, N. I. Goldfield, M. E. Wynn, T. E. McGuire, R. L. Mullin, L. W. Gregg, and J. A. Bender, "Design of a prospective payment patient classification system for ambulatory care," *Health care financing review*, vol. 15, no. 1, p. 71, 1993.

[48] D. L. Leslie, R. Rosenheck, and W. D. White, "Capitated payments for mental health patients: a comparison of potential approaches in a public sector population," *The journal of mental health policy and economics*, vol. 3, no. 1, pp. 35–44, 2000.

[49] A. K. Rosen, S. A. Loveland, J. J. Anderson, C. S. Hankin, J. N. Breckenridge, and D. R. Berlowitz, "Diagnostic cost groups (DCGs) and concurrent utilization among patients with substance abuse disorders," *Health services research*, vol. 37, no. 4, pp. 1079–1103, 2002.

[50] K. L. Sloan, M. E. Montez-Rath, A. Spiro III, C. L. Christiansen, S. Loveland, P. Shokeen, L. Herz, S. Eisen, J. N. Breckenridge, and A. K. Rosen, "Development and validation of a psychiatric casemix system," *Medical care*, pp. 568–580, 2006.

[51] Ontario Joint Policy and Planning Committee and Mental Health Technical Working Group, "Grouper and weighting methodology for adult inpatient mental health care in ontario a technical report," 2008.

[52] J. P. Hirdes, C. V. Everdingen, J. Ferris, M. A. Franco, B. E. Fries, J. Heikkila, A. Hirdes, R. Hoffman, M. L. James, L. Martin, *et al.*, "The interrai suite of mental health assessment instruments: An integrated system for the continuum of care," *Frontiers in Psychiatry*, vol. 10, p. 926, 2020.

[53] T. Trauer, R. A. Duckmanton, and E. Chiu, "Economic grand rounds: Estimation of costs of public psychiatric treatment," *Psychiatric Services*, vol. 49, no. 4, pp. 440–442, 1998.

[54] J. Samuels, *Risky Business: Classifying Mental Health Consumers for Capitation Reimbursement Under Managed Care*. PhD thesis, New York University, 1996.

[55] J. K. Wing, R. Curtis, and A. Beevor, "HoNOS: Health of the nation outcome scales: report on research and development, july 1993-december 1995," 1996.

[56] E. Cohen-Tovée, "Care packages and pathways/payment by results for mental health services for adults, implications for psychological services," *Leicester: British Psychological Society*, 2012.

[57] The Casemix Service, "End stage report - mental health casemix classification development," tech. rep., 2006.

[58] W. D. Wood and D. F. Beardmore, "Prospective payment for outpatient mental health services: evaluation of diagnosis-related groups," *Community mental health journal*, vol. 22, no. 4, pp. 286–291, 1986.

[59] R. B. Fetter, Y. Shin, J. L. Freeman, R. F. Averill, and J. D. Thompson, "Case mix definition by diagnosis-related groups," *Medical care*, vol. 18, no. 2, pp. i–53, 1980.

[60] L. Wittman and Y. Lerner, "Classification system for the chronic mentally ill outpatient.," *Israel Journal of Psychiatry and Related Sciences*, 1990.

[61] S. Barker, N. Barron, B. H. McFarland, and D. A. Bigelow, "A community ability scale for chronically mentally ill consumers: Part i. reliability and validity," *Community mental health journal*, vol. 30, no. 4, pp. 363–383, 1994.

[62] E. S. Uehara, M. Smukler, and F. L. Newman, "Linking resource use to consumer level of need: Field test of the level of need-care assessment (LONCA) method.," *Journal of Consulting and Clinical Psychology*, vol. 62, no. 4, p. 695, 1994.

[63] S. L. Ettner and E. H. Notman, "How well do ambulatory care groups predict expenditures on mental health and substance abuse patients?," *Administration and Policy in Mental Health and Mental Health Services Research*, vol. 24, no. 4, pp. 339–357, 1997.

[64] J. P. Weiner, B. H. Starfield, D. M. Steinwachs, and L. M. Mumford, "Development and application of a population-oriented measure of ambulatory care case-mix," *Medical care*, pp. 452–472, 1991.

[65] S. L. Ettner, R. G. Frank, T. G. McGuire, J. P. Newhouse, and E. H. Notman, "Risk adjustment of mental health and substance abuse payments," *Inquiry*, pp. 223–239, 1998.

[66] A. Rosen, D. Hadzi-Pavlovic, and G. Parker, "The life skills profile: a measure assessing function and disability in schizophrenia," *Schizophrenia Bulletin*, vol. 15, no. 2, pp. 325–337, 1989.

[67] T. Trauer, R. A. Duckmanton, and E. Chiu, "The life skills profile: a study of its psychometric properties," *Australian and New Zealand Journal of Psychiatry*, vol. 29, no. 3, pp. 492–499, 1995.

[68] T. Coombs, "Australian mental health outcomes and classification network: Training manual adult ambulatory," 2005.

[69] S. G. Gowers, R. C. Harrington, A. Whitton, P. Lelliott, A. Beevor, J. Wing, and R. Jezzard, "Brief scale for measuring the outcomes of emotional and behavioural disorders in children: Health of the nation outcome scales for children and adolescents (HoNOSCA)," *The British Journal of Psychiatry*, vol. 174, no. 5, pp. 413–416, 1999.

[70] D. Shaffer, M. S. Gould, J. Brasic, P. Ambrosini, P. Fisher, H. Bird, and S. Aluwahlia, "A children's global assessment scale (cgas)," *Archives of General psychiatry*, vol. 40, no. 11, pp. 1228–1231, 1983.

[71] Department of Health and Ageing, Quality and Effectiveness Section Mental Health and Suicide Prevention Branch, "Mental health national outcomes and casemix collection: Overview of clinician rated and consumer self-report measures and data items version 1.50," 2003.

[72] J. Endicott, R. L. Spitzer, J. L. Fleiss, and J. Cohen, "The global assessment scale: A procedure for measuring overall severity of psychiatric disturbance," *Archives of general psychiatry*, vol. 33, no. 6, pp. 766–771, 1976.

[73] R. DeLiberty, F. L. Newman, and E. O. Ward, "Risk adjustment in the hoosier assurance plan: Impact on providers," *The journal of behavioral health services & research*, vol. 28, no. 3, pp. 301–318, 2001.

[74] A. J. Selim, B. G. Fincke, X. S. Ren, A. Lee, W. H. Rogers, D. R. Miller, K. M. Skinner, M. Linzer, and L. E. Kazis, "Comorbidity assessments based on patient report: results from the veterans health study," *The Journal of ambulatory care management*, vol. 27, no. 3, pp. 281–295, 2004.

[75] R. Kronick, T. Gilmer, T. Dreyfus, and L. Lee, "Improving health-based payment for medicaid beneficiaries: CDPS," *Health care financing review*, vol. 21, no. 3, p. 29, 2000.

[76] P. Martin, R. Davies, A. Macdougall, B. Ritchie, P. Vostanis, A. Whale, and M. Wolpert, "Developing a case mix classification for child and adolescent mental health services: the influence of presenting problems, complexity factors and service providers on number of appointments," *Journal of Mental Health*, pp. 1–8, 2017.

[77] National Institute for Health and Care Excellence, *Guidance List - Clinical guidelines*, 2015.

[78] V. Patel, A. J. Flisher, S. Hetrick, and P. McGorry, "Mental health of young people: a global public-health challenge," *The Lancet*, vol. 369, no. 9569, pp. 1302–1313, 2007.

[79] M. W. Browne, "Cross-validation methods," *Journal of mathematical psychology*, vol. 44, no. 1, pp. 108–132, 2000.

[80] Mental Health Commission of Canada, "Why investing in mental health will contribute to canada's economic prosperity and to the sustainability of our health care system," 2014.

[81] Institute for Health Metrics and Evaluation, "Global burden of diseases, injuries, and risk factors study, 2013," 2014.

[82] Mental Health Commission of Canada, "Changing directions, changing lives the mental health strategy for canada. calgary: Mental health commission of canada," 2012.

[83] Mental Health Commission of Canada, "Advancing the mental health strategy for canada: a framework for action (2017-2022)," 2016.

[84] The Change Foundation, "Who is the puzzle maker? patient/caregiver perspectives on navigating the health services in ontario," 2008.

[85] T. K. Gandhi, "Fumbled handoffs: one dropped ball after another," *Annals of internal medicine*, vol. 142, no. 5, pp. 352–358, 2005.

[86] M. D. Naylor, "A decade of transitional care research with vulnerable elders," *Journal of Cardiovascular Nursing*, vol. 14, no. 3, pp. 1–14, 2000.

[87] T. Stephens and N. Joubert, "The economic burden of mental health problems in canada," *Chronic Dis Can*, vol. 22, no. 1, pp. 18–23, 2001.

[88] S. Ratnasingham, J. Cairney, J. Rehm, H. Manson, and P. A. Kurdyak, "Opening eyes, opening minds: the ontario burden of mental illness and addictions report," *Toronto, Ontario: An ICES/PHO Report*, 2012.

[89] S. Brien, L. Grenier, M. Kapral, P. Kurdyak, and S. Vigod, *Taking stock: A report on the quality of mental health and addictions services in Ontario.* Health Quality Ontario and the Institute for Clinical Evaluative Sciences, 2015.

[90] S. Vanderbent, "Strategies for transition planning in ontario's local health integration networks," *Healthc Q*, vol. 8, no. 3, pp. 78–81, 2005.

[91] M. D. Naylor, L. H. Aiken, E. T. Kurtzman, D. M. Olds, and K. B. Hirschman, "The importance of transitional care in achieving health reform," *Health affairs*, vol. 30, no. 4, pp. 746–754, 2011.

[92] E. Piraino, G. Heckman, C. Glenny, and P. Stolee, "Transitional care programs: who is left behind? a systematic review," *International journal of integrated care*, vol. 12, 2012.

[93] S. Kripalani, F. LeFevre, C. O. Phillips, M. V. Williams, P. Basaviah, and D. W. Baker, "Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care," *Jama*, vol. 297, no. 8, pp. 831–841, 2007.

[94] C. McWilliam and J. Sangster, "Managing patient discharge to home: the challenges of achieving quality of care," *International Journal for Quality in Health Care*, vol. 6, no. 2, pp. 147–161, 1994.

[95] J. P. Hirdes, M. Marhaba, T. F. Smith, L. Clyburn, L. Mitchell, R. Lemick, N. C. Telegdi, E. Perez, P. Prendergast, T. Rabinowitz, *et al.*, "Development of the resident assessment instrument–mental health (RAI-MH)," *Hosp Q*, vol. 4, no. 2, pp. 44–51, 2000.

[96] J. Hirdes, N. Curtin-Telegdi, K. Mathias, C. Perlman, T. Saarela, H. Kolbeinsson, R. Valdimarsdottir, J. Morris, B. Fries, T. Rabinowitz, *et al.*, "interRAI mental health clinical assessment protocols (CAPs) for use with community and hospital-based mental health assessment instruments. version 9.1," 2011.

[97] M. P. Lawton and E. M. Brody, "Assessment of older people: self-maintaining and instrumental activities of daily living," *The gerontologist*, vol. 9, pp. 179–186, 1969.

[98] J. N. Morris, B. E. Fries, D. R. Mehr, C. Hawes, C. Phillips, V. Mor, and L. A. Lipsitz, "MDS cognitive performance scale ©," *Journal of gerontology*, vol. 49, no. 4, pp. M174–M182, 1994.

[99] C. M. Perlman and J. P. Hirdes, "The aggressive behavior scale: a new scale to measure aggression based on the minimum data set," *Journal of the American Geriatrics Society*, vol. 56, no. 12, pp. 2298–2303, 2008.

[100] E. Neufeld, C. M. Perlman, and J. P. Hirdes, "Predicting inpatient aggression using the interRAI risk of harm to others clinical assessment protocol," *The journal of behavioral health services & research*, vol. 39, no. 4, pp. 472–480, 2012.

[101] C. M. Perlman, J. P. Hirdes, and S. Vigod, "Psychiatric rehospitalization: development of a person-level indicator for care planning and quality assurance," *The primary care companion for CNS disorders*, vol. 17, no. 4, 2015.

[102] G. R. Bond, R. E. Drake, K. T. Mueser, and E. Latimer, "Assertive community treatment for people with severe mental illness," *Disease management and health outcomes*, vol. 9, no. 3, pp. 141–159, 2001.

[103] L. I. Stein and A. B. Santos, *Assertive community treatment of persons with severe mental illness.* WW Norton & Co, 1998.

[104] B. J. Burns, A. B. Santos, *et al.*, "Assertive community treatment: an update of randomized trials," *Psychiatric Services*, vol. 46, no. 7, pp. 669–675, 1995.

[105] C. D. Twomey, D. S. Baldwin, M. Hopfe, and A. Cieza, "A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK," *[A-Z][A-Z]+?J open*, vol. 5, no. 7, p. e007575, 2015.

[106] S. Rios, "Living on the edge: the influence of marginalization on mental health needs and service use (phd thesis).," *University of Waterloo*, 2018.

[107] S. N. Vigod, P. A. Kurdyak, D. Seitz, N. Herrmann, K. Fung, E. Lin, C. Perlman, V. H. Taylor, P. A. Rochon, and A. Gruneir, "READMIT: a clinical risk index to predict 30-day readmission after discharge from acute psychiatric units," *Journal of psychiatric research*, vol. 61, pp. 205–213, 2015.

[108] J. Zhang, C. Harvey, and C. Andrew, "Factors associated with length of stay and the risk of readmission in an acute psychiatric inpatient facility: a retrospective study," *Australian and New Zealand Journal of Psychiatry*, vol. 45, no. 7, pp. 578–585, 2011.

[109] J. Durbin, E. Lin, C. Layne, and M. Teed, "Is readmission a valid indicator of the quality of inpatient psychiatric care?," *The journal of behavioral health services & research*, vol. 34, no. 2, pp. 137–150, 2007.

[110] S. N. Vigod, P. A. Kurdyak, C.-L. Dennis, T. Leszcz, V. H. Taylor, D. M. Blumberger, and D. P. Seitz, "Transitional interventions to reduce early psychiatric readmissions in adults: systematic review," *The British Journal of Psychiatry*, vol. 202, no. 3, pp. 187–194, 2013.

[111] L. Smith and H. Daughtrey, "Weaving the seamless web of care: an analysis of parents' perceptions of their needs following discharge of their child from hospital," *Journal of advanced nursing*, vol. 31, no. 4, pp. 812–820, 2000.

[112] J. Toscan, K. Mairs, S. Hinton, P. Stolee, I. R. Team, *et al.*, "Integrated transitional care: patient, informal caregiver and health care provider perspectives on care transitions for older persons with hip fracture," *International journal of integrated care*, vol. 12, 2012.

[113] N. Tran, J. W. Poss, C. Perlman, and J. P. Hirdes, "Case-mix classification for mental health care in community settings: A scoping review," *Health Services Insights*, vol. 12, p. 1178632919862248, 2019.

[114] I. H. P. Authority, "Australian mental health care classification mental health - phase of care guide," 2016.

[115] M. C. Hornbrook, A. V. Hurtado, and R. E. Johnson, "Health care episodes: definition, measurement and use," *Medical care review*, vol. 42, no. 2, pp. 163–218, 1985.

[116] I. H. P. Authority, "Activity based funding mental health care national best endeavours data set 2019-20 technical specifications for reporting," 2018.

[117] H. Lamberts and I. Hofmans-Okkes, "Episode of care: a core concept in family practice," *Journal of Family Practice*, vol. 42, no. 2, pp. 161–170, 1996.

[118] J. A. Solon, J. J. Feeney, S. H. Jones, R. D. Rigg, and C. G. Sheps, "Delineating episodes of medical care.," *American Journal of Public Health and the Nations Health*, vol. 57, no. 3, pp. 401–408, 1967.

[119] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, 2019.

[120] A. Zink and S. Rose, "Fair regression for health care spending," *Biometrics*, 2019.

[121] M. Kuhn *et al.*, "Building predictive models in r using the caret package," *Journal of statistical software*, vol. 28, no. 5, pp. 1–26, 2008.

[122] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[123] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[124] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. Feb, pp. 281–305, 2012.

[125] B. E. Fries, "Comparing case-mix systems for nursing home payment," *Health Care Financing Review*, vol. 11, no. 4, p. 103, 1990.

[126] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees.* Belmont: Wadsworth, 1984.

[127] T. Hothorn, K. Hornik, and A. Zeileis, "ctree: Conditional inference trees," *The Comprehensive R Archive Network*, pp. 1–34, 2015.

[128] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[129] T. Chen, T. He, M. Benesty, V. Khotilovich, and Y. Tang, "Xgboost: extreme gradient boosting," *R package version 0.4-2*, pp. 1–4, 2015.

[130] M. Charlson, T. P. Szatrowski, J. Peterson, and J. Gold, "Validation of a combined comorbidity index," *Journal of clinical epidemiology*, vol. 47, no. 11, pp. 1245–1251, 1994.

[131] T. Stickley and N. Wright, "The british research evidence for recovery, papers published between 2006 and 2009 (inclusive). part one: a review of the peer-reviewed literature using a systematic approach," *Journal of Psychiatric and Mental Health Nursing*, vol. 18, no. 3, pp. 247–256, 2011.

[132] B. Rush, "Tiered frameworks for planning substance use service delivery systems: Origins and key principles," *Nordic Studies on Alcohol and Drugs*, vol. 27, no. 6, pp. 617–636, 2010.

[133] W. J. Katon, E. H. Lin, M. Von Korff, P. Ciechanowski, E. J. Ludman, B. Young, D. Peterson, C. M. Rutter, M. McGregor, and D. McCulloch, "Collaborative care for patients with depression and chronic illnesses," *New England Journal of Medicine*, vol. 363, no. 27, pp. 2611–2620, 2010.

[134] G. M. Hackbarth, R. Berenson, M. E. Miller, *et al.*, "Report to the congress: Medicare payment policy," *Testimony on behalf of MedPAC, March17*, 2009.

[135] L. Hansson, "Inequality and inequity in use of mental health services.," 2003.

[136] Canadian Institute for Health Information, "Canadian patient cost database technical document MIS - patient costing methodology," 2019.

[137] C. O'neil, *Weapons of math destruction: How big data increases inequality and threatens democracy.* Broadway Books, 2016.

[138] D. R. Williams, S. A. Mohammed, J. Leavell, and C. Collins, "Race, socioeconomic status and health: Complexities, ongoing challenges and research opportunities," *Annals of the New York Academy of Sciences*, vol. 1186, p. 69, 2010.

[139] J. Wiens, S. Saria, M. Sendak, M. Ghassemi, V. X. Liu, F. Doshi-Velez, K. Jung, K. Heller, D. Kale, M. Saeed, *et al.*, "Do no harm: a roadmap for responsible machine learning for health care," *Nature medicine*, vol. 25, no. 9, pp. 1337–1340, 2019.

[140] Canadian Institute for Health Information, "About CIHI." `https://www.cihi.ca/en/about-cihi`.

Appendix A

Crosswalks of items between the RAI-MH and HoNOS and HoNOS 65+

* For each HoNOS item, the default value is zero. Increment values were added based on coding rules in RAI-MH Items column.

HoNOS	HoNOS Level	RAI-MH Items
1. Aggressive Behaviors	1	B1j Irritability = 1 OR B1dd Anger = 1 OR E1d Socially inappropriate behavior = 1 OR I2a Akathisia = 1
	2	B1j Irritability = 2 OR B1dd Anger = 2 OR E1d Socially inappropriate behavior ≥ 2 OR D2b Intimidation of others/threatened violence = 3 OR E1b Verbal abuse = 1 OR E1f Resistance to care = 1

HoNOS	HoNOS Level	RAI-MH Items
	3	B1j Irritability = 3 OR B1dd Anger = 3 OR D2b Intimidation of others/threatened violence = 4 OR D2a Violent to others $\geq$ 3 OR E1b Verbal abuse = 2 OR E1c Physical abuse $\geq$ 1 OR E1f Resistance to care = 2
	4	E1c Physical abuse = 3 OR D2b Intimidation of others/threatened violence = 5 OR D2a Violent to others $\geq$ 4 OR E1b Verbal abuse = 3 OR E1c Physical abuse = 3 OR E1f Resistance to care = 3
2. Self-injury	1	D1b Intent of any self-injurious attempt was to kill himself/herself = 1 OR D1c Considered performing a self-injurious act $\geq$ 1
	2	D1c Considered performing a self-injurious act = 3
	3	D1a Most recent self-injurious attempt = 1 OR D1c Considered performing a self-injurious act = 4 OR D1da Family, caregiver, friend or staff expresses concern that person is at risk for self-injury = 1 OR D1db Suicide plan = 1 OR K4 Intentional misuse of medication = 1
	4	D1a Most recent self-injurious attempt $\geq$ 2
3. Problem drinking or drug-taking	1	C1 Number of drinks in any single sitting episode in the last 14 days = 3
	2	C4a Person felt the need or was told by others to cutdown on drinking or drug use, or others were concerned about person's substance use = 1
	3	C3 Withdrawal symptoms $\geq$ 1
	4	C3 Withdrawal symptoms = 3 OR C4d Person had to have a drink or use drugs first thing in the morning to steady nerves = 1
4. Cognitive Problems	1	F2 Cognitive skills for daily decision-making = 2 OR Q1b Delirium, dementia and amnestic and other cognitive disorders $\geq$ 1

HoNOS	HoNOS Level	RAI-MH Items
	2	F1a Short-term memory = 1 OR F1b Procedural memory = 1 OR F2 Cognitive skills for daily decision-making = 3 OR F3b Periods of altered perception or awareness of surroundings = 1 OR H3 Making self understood = 2
	3	F2 Cognitive skills for daily decision-making = 4 OR F3b Periods of altered perception or awareness of surroundings = 2 OR F3c Episodes of disorganized speech = 1 OR H3 Making self understood = 3
	4	F2 Cognitive skills for daily decision-making = 5 OR F3c Episodes of disorganized speech = 2 OR H3 Making self understood = 4
5. Physical Illness	1	G1b Walking $\geq$ 1 OR I4 Self-reported health $\geq$ 2 OR I7 Falls = 1 For 65+: G1b Walking $\geq$ 1 OR I4 Self-reported health $\geq$ 2 OR I8b Intensity of pain = 1 OR H1 Hearing = 1 OR H2 Vision = 1
	2	G1b Walking = 4 OR I6b Foot problems = 2 For 65+: G1b Walking = 4 OR I6b Foot problems = 2 OR I7a Falls = 1 OR I8a Frequently complains or shows evidence of pain in the last 3 days = 1 OR I8b Intensity of pain = 2 OR H1 Hearing = 2 OR H2 Vision = 2

HoNOS	HoNOS Level	RAI-MH Items
	3	G1b Walking = 5 OR I6b Foot problems = 3 For 65+: G1b Walking = 5 OR I6b Foot problems = 3 OR I7a Falls = 2 OR I8a Frequently complains or shows evidence of pain in the last 3 days = 2 OR I8b Intensity of pain = 3 OR H1 Hearing = 3 OR H2 Vision = 3
	4	G1b Walking = 6 OR I6b Foot problems = 4 For 65+: G1b Walking = 6 OR G1c Wheeling ≥ 1 OR I6b Foot problems = 4 OR I7a Falls = 3 OR I7b Recent falls = 1 OR I8a Frequently complains or shows evidence of pain in the last 3 days = 2 OR I8b Intensity of pain = 4 OR H2 Vision = 4 OR F2 Cognitive skills for daily decision-making = 5
6. Hallucinations and Delusions	1	B1w Delusions = 1
	2	B1u Hallucinations = 1 OR B1v Command hallucinations = 1 OR B1w Delusions = 1
	3	B1u Hallucinations + B1v Command hallucinations + B1w Delusions ≥ 2
	4	B1u Hallucinations + B1v Command hallucinations + B1w Delusions ≥ 4
7. Depressed Mood	1	B1a Facial expression ≥ 1 OR B1c Decreased energy ≥ 1 OR Q1f Mood disorders ≥ 1

HoNOS	HoNOS Level	RAI-MH Items
	2	B1e Self-deprecation ≥ 1 OR B1f Guilt/shame ≥ 1
	3	B1f Guilt/shame = 3
	4	B1f Guilt/shame = 3 AND B1g Hopelessnss = 3
8. Other Mental OR Behavioral Problems	2	B1p Fears/phobia = 1 OR B1o Anxious complaints = 1 OR B1q Obsessive thoughts = 1 OR B1r compulsive behavior = 1 OR Q1j Dissociative disorders ≥ 1 OR Q1h Somatoform disorders ≥ 1 OR N3a Any instances of binge eating, purging or bulimia = 1 OR N3b Unrealistic fear of weight gain, statements that suggest a distorted body image = 1 OR N3c Fasting or major restriction of diet = 1 OR Q1l Eating disorders ≥ 1 OR Q1m Sleep disorders ≥ 1 OR B1gg Sleep problems = 1 OR Q1k Sexual and gender identity disorders ≥ 1 OR B1k Increased sociability or hypersexuality = 1 OR E1e Inappropriate public sexual behavior = 1 OR I3 Sexual functioning = 1
	3	B1p Fears/phobias = 2 OR B1o Anxious complaints = 2 OR B1q Obsessive thoughts = 2 OR B1gg Sleep problems = 2 OR B1k Increased sociability or hypersexuality = 2 OR E1e Inappropriate public sexual behavior = 2

HoNOS	HoNOS Level	RAI-MH Items
	4	B1p Fears/phobias = 3 OR B1o Anxious complaints = 3 OR B1q Obsessive thoughts = 3 OR B1gg Sleep problems = 3 OR B1k Increased sociability or hypersexuality = 3 OR E1e Inappropriate public sexual behavior = 3
9. Relationships	2	O2a Reports having no confidant = 1 OR O2b Family/close friends report feeling overwhelmed by person's illness = 1 OR O2c Is persistently hostile towards or critical of family/friends = 1 OR O2d Is persistently hostile towards or critical of others or staff = 1 OR O2e Family/friends are persistently hostile towards or critical of person = 1 OR O2f Staff reports persistent frustration in dealing with person = 1 OR O2g Family/friends require unusual amounts of facility staff time = 1 OR O6a Participation in social activities of long-standing interest = 1 OR O6b Visit by long-standing social relation/family member = 1 OR O6c Telephone or email contact with long-standingsocial relation/family member = 1
	3	O6a Participation in social activities of long-standing interest = 2 OR O6b Visit by long-standing social relation/family member = 2 OR O6c Telephone or email contact with long-standing social relation/family member = 2
	4	O6a Participation in social activities of long-standing interest = 3 OR O6b Visit by long-standing social relation/family member = 3 OR O6c Telephone or email contact with long-standing social relation/family member = 3

HoNOS	HoNOS Level	RAI-MH Items
10. Daily Activities	1	G1a Personal hygiene = 2 OR G1b Walking = 2 OR G1c Wheeling = 2 OR G1d Toilet use = 2 OR G1e Eating = 2 OR G2a Meal preparation = 2 OR G2b Managing medications = 2 OR G2c Transportation = 2 OR G2d Managing finances = 2 OR G2e Phone use = 2
	2	G1a Personal hygiene = 3 OR G1b Walking = 3 OR G1c Wheeling = 3 OR G1d Toilet use = 3 OR G1e Eating = 3 OR G2a Meal preparation = 3 OR G2b Managing medications = 3 OR G2c Transportation = 3 OR G2d Managing finances = 3 OR G2e Phone use = 3
	3	G1a Personal hygiene ≥ 4 OR G1b Walking ≥ 4 OR G1c Wheeling ≥ 4 OR G1d Toilet use ≥ 4 OR G1e Eating ≥ 4 OR G2a Meal preparation ≥ 4 OR G2b Managing medications ≥ 4 OR G2c Transportation ≥ 4 OR G2d Managing finances ≥ 4 OR G2e Phone use ≥ 4

HoNOS	HoNOS Level	RAI-MH Items
	4	G1a Personal hygiene ≥ 6 OR G1b Walking ≥ 6 OR G1c Wheeling ≥ 6 OR G1d Toilet use ≥ 6 OR G1e Eating ≥ 6 OR G2a Meal preparation ≥ 6 OR G2b Managing medications ≥ 6 OR G2c Transportation ≥ 6 OR G2d Managing finances ≥ 6 OR G2e Phone use ≥ 6
11. Living Conditions	1	CC4b Usual residence = 2 OR P5 Discharged to = 2
	2	CC4b Usual residence = 3 OR P5 Discharged to = 3
	4	CC4b Usual residence = 8 OR P5 Discharged to = 8
12. Occupations	1	O3 Current employment status = 0 OR (O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years = 1) AND (O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years <8)

HoNOS	HoNOS Level	RAI-MH Items
	2	O3 Current employment status = 1 OR ((O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years = 2) AND (O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years <8))
	3	O3 Current employment status = 2 OR ((O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years = 3) AND (O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years <8))
	4	O3 Current employment status = 2 OR ((O4a Increase in lateness or absenteeism over the last6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years = 4) AND (O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years <8))

Appendix B

Crosswalks of items between the RAI-MH and LSP-16

* For each LSP-16 item, the default value is zero. Increment values were added based on coding rules in RAI-MH Items column.

LSP-16 Items	LSP-16 level	RAI-MH items
1. Initiating and responding to conversation	1	F3a Easily distracted = 1 OR F3b Periods of altered perception or awareness of surroundings = 1 OR F3c Episodes of disorganized speech = 1 OR F3d Periods of restlessness = 1 OR F3e Periods of lethargy = 1 OR F3f Mental function varies over the course of the day = 1
	2	F3a Easily distracted = 2 OR F3b Periods of altered perception or awareness of surroundings = 2 OR F3c Episodes of disorganized speech = 2 OR F3d Periods of restlessness = 2 OR F3e Periods of lethargy = 2 OR F3f Mental function varies over the course of the day = 2

LSP-16 Items	LSP-16 level	RAI-MH items
	3	F3a Easily distracted + F3b Periods of altered perception or awareness of surroundings + F3c Episodes of disorganized speech + F3d Periods of restlessness + F3e Periods of lethargy + F3f Mental function varies over the course of the day ≥ 7
2. Withdraw from social contact	1	B1z Loss of interest = 1 OR B1bb Reduced interaction = 1 OR O6a Participation in social activities of long-standing interest = 1 OR O6b Visit by long-standing social relation/family member = 1 OR O6c Telephone or email contact with long-standing social relation/family member = 1
	2	B1z Loss of interest = 2 OR B1bb Reduced interaction = 2 OR O6a Participation in social activities of long-standing interest = 2 OR O6b Visit by long-standing social relation/family member = 2 OR O6c Telephone or email contact with long-standing social relation/family member = 2
	3	B1z Loss of interest = 3 OR B1bb Reduced interaction = 3 OR O6a Participation in social activities of long-standing interest = 3 OR O6b Visit by long-standing social relation/family member = 3 OR O6c Telephone or email contact with long-standing social relation/family member = 3
3. Show warm to others	2	O2c Is persistently hostile towards or critical of family/friends + O2d Is persistently hostile towards or critical of others or staff = 1
	3	O2c Is persistently hostile towards or critical of family/friends + O2d Is persistently hostile towards or critical of others or staff = 2
4. Grooming	1	G1a Personal hygiene ≥ 2

LSP-16 Items	LSP-16 level	RAI-MH items
	2	G1a Personal hygiene ≥ 4
	3	G1a Personal hygiene ≥ 6
5. Wash clothes	1	B1ff Hygiene = 1
	2	B1ff Hygiene = 2
	3	B1ff Hygiene = 3
6. Neglect physical health	1	B1aa Lack of motivation = 1 OR E1f Resistance to care = 1
	2	B1aa Lack of motivation = 2 OR E1f Resistance to care = 2
	3	B1aa Lack of motivation = 3 OR E1f Resistance to care = 3
7. Violent to others	1	A5a Police intervention for violent behavior ≥ 1 OR D2a Violent to others ≥ 1 OR D2b Intimidation of others or threatened violence ≥ 1 OR D2c Violent ideation ≥ 1 OR E1c Physical abuse = 1
	2	A5a Police intervention for violent behavior ≥ 4 OR D2a Violent to others≥ 4 OR D2b Intimidation of others or threatened violence ≥ 4 OR D2c Violent ideation ≥ 4 OR E1c Physical abuse = 2
	3	A5a Police intervention for violent behavior ≥ 5 OR D2a Violent to others ≥ 5 OR D2b Intimidation of others or threatened violence ≥ 5 OR D2c Violent ideation ≥ 5 OR E1c Physical abuse = 3
8. Make/keep friendships	1	O6a Participation in social activities of long-standing interest + O6b Visit by long-standing social relation/family member + O6c Telephone or email contact with long-standing social relation/family member ≥ 1

LSP-16 Items	LSP-16 level	RAI-MH items
	2	O6a Participation in social activities of long-standing interest + O6b Visit by long-standing social relation/family member + O6c Telephone or email contact with long-standing social relation/family member ≥ 3
	3	O6a Participation in social activities of long-standing interest + O6b Visit by long-standing social relation/family member + O6c Telephone or email contact with long-standing social relation/family member ≥ 4
9. Diet	1	N2a Weight loss of 5% or more in the last 30 days or 10% or more in the last 180 days + N2b Weight gain of 5% or more in the last 30 days or 10% or more in the last 180 days + N2c Insufficient fluid — less than 1,000 cc per day or less than four 8-oz cups per day + N2d In the last 3 days, noticeable decrease in the amount of food or fluid usually consumed ≥ 1
	2	N2a Weight loss of 5% or more in the last 30 days or 10% or more in the last 180 days + N2b Weight gain of 5% or more in the last 30 days or 10% or more in the last 180 days + N2c Insufficient fluid — less than 1,000 cc per day or less than four 8-oz cups per day + N2d In the last 3 days, noticeable decrease in the amount of food or fluid usually consumed ≥ 2

LSP-16 Items	LSP-16 level	RAI-MH items
	3	N2a Weight loss of 5% or more in the last 30 days or 10% or more in the last 180 days + N2b Weight gain of 5% or more in the last 30 days or 10% or more in the last 180 days + N2c Insufficient fluid — less than 1,000 cc per day or less than four 8-oz cups per day + N2d In the last 3 days, noticeable decrease in the amount of food or fluid usually consumed ≥ 3
10. Medication adherence	1	K1 History of medication adherence = 1
	3	K1 History of medication adherence = 2
11. Medication refusal	2	K2 Medication refusal = 1
12. Cooperative with care providers	1	E1f Resistance to care = 1 OR L5 Adherence to treatments, therapies, programs = 1
	2	E1f Resistance to care = 2 OR L5 Adherence to treatments, therapies, programs = 2
	3	E1f Resistance to care = 3 OR L5 Adherence to treatments, therapies, programs = 3
13. Problems living with others	1	O1 Relationship(s) with immediate family members is disturbed or dysfunctional = 1
	2	O1 Relationship(s) with immediate family members is disturbed or dysfunctional = 2
	3	O1 Relationship(s) with immediate family members is disturbed or dysfunctional = 3
14. Offensive behavior	1	D3 History of sexual violence or assault as perpetrator = 1 OR E1e Inappropriate public sexual behavior = 1
	2	E1e Inappropriate public sexual behavior = 2

LSP-16 Items	LSP-16 level	RAI-MH items
	3	E1e Inappropriate public sexual behavior = 3
15. Behave responsibly	1	E1a Wandering = 1 OR E1b Verbal abuse = 1 OR E1c Physical abuse = 1 OR E1d Socially inappropriate/disruptive behavior = 1 OR E1e Inappropriate public sexual behavior = 1 OR E2 Extreme behavior disturbance = 1
	2	E1a Wandering = 2 OR E1b Verbal abuse = 2 OR E1c Physical abuse = 2 OR E1d Socially inappropriate/disruptive behavior = 2 OR E1e Inappropriate public sexual behavior = 2 OR E2 Extreme behavior disturbance = 2
	3	E1a Wandering = 3 OR E1b Verbal abuse = 3 OR E1c Physical abuse = 3 OR E1d Socially inappropriate/disruptive behavior = 3 OR E1e Inappropriate public sexual behavior = 3 OR E2 Extreme behavior disturbance = 3
16. Employment capability	1	O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years ≥ 1
	2	O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years ≥ 2
	3	O4a Increase in lateness or absenteeism over the last 6 months + O4b Poor productivity or disruptiveness at work/school + O4c Expresses intent to quit work/school + O4d Persistent unemployment or fluctuating work history over the last 2 years ≥ 3