Insights and Advancing Mental Health Care: The Utility of Administrative Health Records

by

Derek Pierce

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychiatry

University of Alberta

Abstract

This dissertation explores the role and future possibilities of Administrative Health Records (AHRs) in enriching mental health care analysis and intervention strategies. It is organized into various sections that are based off papers or potential papers, each delving into distinct aspects of AHR utilization. Initially, this dissertation delves into a narrative review and commentary on the evolution of mental healthcare. The focus then shifts to the critical balance between ensuring privacy and leveraging the potential of AHRs, highlighting the need for privacy measures in the handling of synthetic health data and the significance of collaborative efforts among governmental bodies, healthcare entities, and academic institutions for effective data utilization and enhancement.

The narrative progresses to elaborate on several case studies, illustrating AHRs' efficacy in mental health research. These papers are meant to showcase the application of machine learning (ML) to predict opioid overdose risks, the examination of developmental disorder utilization shifts within the Alberta healthcare system, and the impact of the pandemic on neurocognitive disorder trends and healthcare demands. These examples underline the diverse applications and insights AHRs can provide into mental health issues.

Looking to the future, the dissertation advocates for broadening AHRs' applications, including their potential in post-disaster mental health outcome predictions. It proposes an integration of AHRs with wearable device data, aiming to transform mental health care from a traditionally reactive approach to a proactive and preventive strategy. This forward-thinking perspective envisions a system where real-time data from wearables enriches AHRs, offering nuanced, immediate insights into individual mental health statuses.

Overall, the dissertation aims to comprehensively dissect the capacity of AHRs to revolutionize mental health care research and practice. It not only addresses the challenges of privacy and the necessity for cross-sector collaboration but also demonstrates the practical applications of AHRs in current mental health scenarios and anticipates their future role in advancing mental health care, especially in contexts affected by disasters.

This work not only highlights the present state of mental health care research but also recommends new directions for future innovations in the field.

Preface

This thesis focuses on the potential use of Administrative Health Records in managing crisis events that severely impact mental health. Some research conducted for this thesis was part of a research collaboration. The papers are collaborations with different institutions in Alberta. All secondary analyses received approval from the University of Health Ethics Board (Pro00072946).

Chapter 2 was a collaborative project with Yutong Li, Andrew Greenshaw, Bo Cao, and Tracey Bailey who are all affiliated with the University of Alberta. I designed the main study, initial review, and was the primary contributor to the writing and editing of the original manuscript.

Chapter 3 was also a collaborative project approved by the University of Alberta Health Ethics Board (Pro00072946). Yang S. Liu and I did the primary investigation, data analysis, writing, and editing of the original manuscript. Dan Metes, Lawrence Kiyang, and Mengzhe Wang from the Government of Alberta's Ministry of Health and Yipeng Song, Fernanda Talarico, Yutong Li, Jia Lin Tian, and Bo Cao from the University of Alberta aided in conceptualization, results interpretations, and editing.

Chapter 4 This paper was a collaborative project with Yipeng Song, Fernanda Talarico, Yutong Li, Jia Lin Tian, and Bo Cao who are affiliated with the University of Alberta. This was also done in collaboration with Mengzhe Wang from the Ministry of Health for the Government of Alberta. The data was collected and stored at the Ministry of Health. I conducted the main investigation, data analysis, writing and editing of the original manuscripts. This paper was based on a secondary analysis of administrative healthcare data collected by the Ministry of Health in Alberta and was conducted within the Ministry.

Chapter 5 This paper was a collaborative project with Yipeng Song, Fernanda Talarico, Huda Al-Shamal, Yutong Li, Jia Lin Tian, and Bo Cao who are affiliated with the University of Alberta. This was also done in collaboration with Mengzhe Wang from the Ministry of Health for the Government of Alberta. The data were collected and stored at the Ministry of Health. I conducted the main investigation, data analysis, writing and editing of the original manuscripts. This paper was based on a secondary analysis of administrative healthcare data collected by the Ministry of Health in Alberta and was conducted within the Ministry.

To my wife and daughter, whose love, understanding, and encouragement have been the driving force behind my pursuit of knowledge and the completion of this thesis.

Together.

Acknowledgements

I express my gratitude to the University of Alberta, Ministry of Health, and HiMARC for providing opportunities and support throughout my degree. I am humbled to have worked with amazing researchers from these esteemed institutions.

I also extend my appreciation to my supervisors, Cloud, Mengzhe, and Andy, for their faith in me, allowing me to grow as a student, researcher, and as a clinician. I am thankful for their mentorship and guidance. I would also like to acknowledge Allen Chan, Tristin Hamilton, Phillip Sevigny, Jake Hayward, Yanbo Zhang, and Suzette Bremault-Phillips for their personal mentorship and support. Their time and effort made this journey more meaningful and fulfilling. Furthermore, I extend my thanks to the talented Fernanda Talarico, Yipeng Song, and Yang Liu for their friendship, patience, and support throughout my degree.

I am indebted to my wife and daughter, Stephanie and Jaimee Pierce, for providing me with the time and support to excel in my studies and career. This thesis is dedicated to both of them. I am particularly grateful to my wife, who always supported my pursuits and let me be my true self.

I am grateful to my brother, David Greekas, for all his support over the years to both me and my family. Our walks together provided not only stress relief but also a cherished source of happiness with challenging conversations that kept me grounded throughout my studies.

I would also like to express my heartfelt appreciation to MaryAnn Korolyk and her family, who played an extraordinary role in my life during a time of immense hardship. Following the passing of my mother and the absence of my father, they opened their hearts and home, adopting me as their own. MaryAnn's unwavering love, care, and support provided me with a stable and nurturing environment that allowed me to heal and grow. Her selflessness and dedication have forever impacted my life, reminding me of the immeasurable kindness that exists in the world. I am forever grateful for Maryann and the impact she had on shaping the person I am today.

Additionally, I want to extend my appreciation to my Uncle Ray Greekas, Aunt Ingrid Piecha, Grandma Marian Hanneman, and my grandfather the late Wilfred Hanneman. Their presence, love, and support have been a constant source of comfort and strength for me, especially during the challenging times. Despite the profound loss we have experienced, their guidance and unwavering support continue to uplift me, reminding me of the enduring power of family bonds. Their presence in my life has been a true blessing, and I am deeply grateful for their love, guidance, and belief in me.

Finally, I am filled with bittersweet appreciation for my mother, Lynette Greekas, whose absence in witnessing my growth and accomplishments leaves an indescribable void. Yet, her unconditional love and unwavering belief in me as a child continue to shape my journey, reminding me to challenge myself and make her proud of the person I have become.

Contents

List of Tables

Primary Tables

Supplemental Tables

List of Figures

DESTITUTUS VENTIS, REMOS ADHIBE

“If the wind will not serve, take to the oars.”

-Latin Proverb

Chapter 1
Narrative Review and Commentary on the Evolution of Mental Health Care: Leveraging Machine Learning and Administrative Health Records

This section delves into the integration of Machine Learning (ML) algorithms with Administrative Health Records (AHRs) and the potential for advancements in mental health care, a topic that has gained relevance in the wake of escalating global mental health challenges, magnified by the continued fallout of the COVID-19 pandemic lockdowns. It underscores the novelty of merging ML with AHRs to address persistent challenges in psychiatric care, promising directions for enhanced diagnostic precision, research in personalized treatment plans, and improved patient outcomes. It is important to note that while statistical methods are capable of identifying associations between risk factors and disease outcomes, the advantage of ML algorithms lies in their ability to process and learn from vast amounts of data in complex, non-linear ways, offering potentially greater precision in predicting individual disease risks and enhancing the personalization of healthcare interventions.

This exploration is rooted in an examination of the historical evolution of psychiatric practices, diagnostic challenges, and the potential of ML to revolutionize mental health services.

Critiques on the Section

The narrative review in this chapter does not explicitly articulate critiques within its discourse. However, in the broader discussion of ML and AHRs in mental health care there are considerations and challenges such as data robustness, ethical implications of data use, and the need for interdisciplinary collaboration. These areas, while not directly criticized, represent fields of ongoing debate and concern within the integration of technology and health care, suggesting an underlying challenge of current practices and the need for careful consideration of these issues.

Future Research Directions

The chapter identifies several future research directions that could leverage the foundational work presented in this review. It suggests expanding the predictive capabilities of ML models by incorporating data from wearable devices, which could provide real-time physiological and behavioral insights, enhancing the precision of mental health diagnoses and interventions. Additionally, it proposes exploring the application of these models in post-disaster scenarios to better understand and address the unique mental health challenges arising from such events.

The contents and insights provided within this chapter could be further developed into an academic paper, potentially suitable for submission to a journal specializing in computational psychiatry, health informatics, or public health. Such a paper could contribute to the academic discourse on leveraging technology to improve mental health outcomes, highlighting the innovative approaches, challenges, and ethical considerations inherent in this interdisciplinary field. However, its primary intent is to provide a backdrop for the subsequent chapters, many of which are in some stage of being published.

1.0 Leveraging Machine Learning and Administrative Health Records

1.1 Abstract

This narrative review examines the integration of ML and AHRs in mental health care, focusing on recent advancements and the impact of ML on improving diagnostic accuracy and treatment outcomes. By selecting and analyzing relevant studies, this review aims to provide a clear understanding of how the merging of these domains is leading to new scientific frontiers of research, providing the reader with a thorough overview of current insights and emerging trends.

1.2 Introduction

In the evolving landscape of psychiatry, integrating ML with AHRs presents a novel approach to overcoming longstanding challenges in mental health care. This combination offers promising solutions for enhancing diagnostic accuracy, personalizing treatment plans, and ultimately, improving patient outcomes. This review delves into the intersection of psychiatry, public health, and ML, spotlighting how these advancements could transform mental health care. We discuss the historical evolution of psychiatric practices, the diagnostic challenges inherent in current classification systems, and the potential of ML to address these issues effectively. This chapter lays the groundwork for a comprehensive exploration of ML's role in advancing psychiatric care, aiming to contribute to a future where mental health services are more tailored, efficient, and responsive to individual needs.

1.3 Defining Administrative Health Records and Their Distinction from Other Health Data

Before moving ahead, it's essential to clarify the concept of Administrative Health Records (AHRs) and how they differ from other health data forms, especially Electronic Health Records (EHRs) and various health data sources. Given the variability in literature regarding these terms, this section aims to clearly delineate these distinctions for the context of this paper.

Administrative Health Records refer to the collection of data primarily generated through patient interactions with billing and administrative functions within the healthcare system. AHRs typically include information on patient demographics, billing details, diagnoses coded using standardized systems (like ICD codes), treatment codes, and other services rendered. This type of record is often used for administrative purposes, health services research, and population health management.

In contrast, Electronic Health Records are digital versions of patients' paper charts and are designed to be a more comprehensive record of patient care. EHRs include detailed medical histories, medications, lab results, imaging reports, and notes from healthcare providers. The primary aim of EHRs is to support ongoing medical care for patients by providing clinicians with detailed and up-to-date information.

Another term often mentioned in healthcare discussions is health data, which encompasses a broader range of data types, including both clinical data from EHRs and administrative data from AHRs, as well as data from wearable devices, patient-reported outcomes, and social determinants of health. Health data can be used for a wide range of applications, from clinical decision support to population health management and research.

The distinction between AHRs and EHRs lies in their primary purposes and content: AHRs are more focused on administrative, billing, and population-level information, whereas EHRs are centered around detailed clinical information to support patient care. Understanding these differences is crucial when discussing the integration of ML in mental health care, as the source and type of data can significantly influence the insights gained and the applications developed.

1.4 Historical Overview: Psychiatry and Public Health Challenges

The field of psychiatry has witnessed a series of transformative shifts, each contributing to its evolving landscape. One of the pivotal figures in this transformation is Emil Kraepelin, whose work in the late 19th and early 20th centuries laid the foundation for modern psychiatric diagnosis (Kendler, 2016). Kraepelin was instrumental in differentiating between manic-depressive illness—now termed bipolar disorder—and dementia praecox, currently known as schizophrenia. His emphasis on observable symptomatology and disease course marked a significant departure from the psychoanalytic paradigms of Sigmund Freud (Shorter, 1997).

In the mid-20th century, the American Psychiatric Association introduced the first edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association , 1952). Concurrently, the World Health Organization launched the International Classification of Diseases (ICD), which included a section on mental disorders (World Health Organization, 1948). These classification systems, although groundbreaking, were not without limitations. Issues related to the reliability and validity of symptom-based diagnostic criteria emerged, particularly because of overlapping symptoms across different disorders and the lack of consideration for underlying biological and genetic factors (Clark, Cuthbert, Lewis-Fernández, Narrow, & Reed, 2017).

The advent of big data has resulted in a shift towards personalized, precision medicine in psychiatry (Wang & Krishnan, 2014). By leveraging large-scale datasets and ML algorithms, researchers have been able to identify intricate patterns across brain activity, behavior, and genetic markers (Monteith, Glenn, Geddes, Whybrow, & Bauer, 2016). This has opened avenues for a biologically-grounded reclassification of major psychiatric disorders, enabling individualized early diagnosis, treatment selection, and dosage adjustments (Fernandes, et al., 2017).

Further, big data has expanded the scope of psychiatric research to include population-level studies across diverse geographies and contexts, including disasters (Coppersmith, Dredze, & Harman, 2014). For example, ML has been employed to study mental health outcomes among immigrants, refugees, and racial and ethnic minorities, offering insights that inform targeted interventions (Fazel, Hoagwood, Stephan, & Ford, 2014). Additionally, real-time data analytics have been used to assess the impact of crises, such as the COVID-19 pandemic, on mental health indicators (Holmes, et al., 2020)

The integration of big data into psychiatric research has also facilitated comparative studies across different regions, revealing geographical variations in the prevalence and risk factors of mental disorders (Mathers & Loncar, 2006). This has implications for resource allocation in mental health services. Moreover, population-level analyses have provided nuanced insights into the societal and economic ramifications of mental disorders, thereby informing public health strategies (Patel, et al., 2018).

Despite these advancements, the field continues to navigate the tension between the need for standardized diagnostic criteria and the inherently complex nature of mental disorders (Zachar & Kendler, 2017). The evolution of psychiatric diagnostics, from Kraepelin's era to the present, encapsulates the ongoing endeavor to comprehend and effectively manage mental disorders.

1.5 Diagnostic Dilemmas: The DSM and ICD in Mental Health

The history of psychiatric diagnosis in Canada is a complex interplay between two major classification systems: the DSM and the ICD. The DSM, published by the American Psychiatric Association, is the primary tool for diagnosing mental health disorders in clinical settings across Canada. It places an emphasis on observable signs and symptoms and undergoes periodic updates to incorporate the latest psychiatric research (American Psychiatric Association , 2017). Conversely, the ICD, published by the World Health Organization, is predominantly used for administrative functions such as billing and health services management (World Health Organization, 2018). Canada formally adopted the ICD system in 1979 for physician billing claims, thereby standardizing data collection both nationally and internationally (Garies, et al., 2022). Canada later introduced its own adaptation, with several versions in use, like the ICD-10-CA remaining in use (Canadian Institute for Health Information, 2023).

This dual usage of the DSM and ICD has both historical roots and practical implications. Clinicians in Canada primarily rely on the DSM for diagnostic purposes, while the ICD serves as the coding system for administrative and billing procedures. However, this bifurcated approach is not without complications. For instance, the DSM and ICD, although largely congruent, have subtle differences in the classification of certain conditions. A clinician diagnosing a patient with Generalized Anxiety Disorder based on the DSM criteria may encounter difficulties in finding an exactly corresponding ICD code for billing purposes. This discrepancy can lead to data quality issues and potential misrepresentations in healthcare statistics.

Moreover, the DSM employs a categorical approach to diagnosis, where a disorder is either present or absent. This is in contrast to the ICD, which adopts a dimensional approach that considers the severity and duration of symptoms (First, Reed, Hyman, & Saxena, 2015). Such divergence can result in diagnostic and billing inconsistencies. For example, a patient diagnosed with Major Depressive Disorder based on the DSM criteria may initially receive an ICD code reflecting moderate severity. However, if the patient's symptoms escalate, the DSM diagnosis remains static, while a modification of the ICD code to indicate increased severity becomes necessary (Zimmerman, Ellison, Young, Chelminski, & Dalrymple, 2015).

As psychiatry advances, the imperative to address the complexities arising from the concurrent use of the DSM and ICD systems gains importance. This is essential not only for diagnostic precision but also for the appropriate allocation of healthcare resources through accurate billing. While both the DSM and ICD are subject to ongoing revisions aimed at resolving these challenges, vigilance will be required to ensure changes don’t inadvertently alter cause errors in interpretations built on the duality of this system. Specifically, the Canadian healthcare system should adopt proactive measures, such as the formulation of crosswalks or unified diagnostic algorithms, to facilitate the translation of DSM-based diagnoses into corresponding ICD codes, and vice versa. However, even with such tools they will need to continually be monitored as these two systems change over time. Such initiatives would not only elevate the standard of psychiatric care but also enhance the reliability and validity of healthcare data, a critical factor in not only informed policy-making and resource distribution but also for use in ML.

1.6 Current Landscape: Morbidity, Mortality, and Economic Impact of Mental Health

Building on the complexities of psychiatric diagnosis and the dual use of the DSM and ICD systems in Canada, it becomes necessary to delve into the broader societal and economic implications of mental health disorders as we look at precision health for individuals. These disorders have emerged as a leading cause of disability in developed countries, a trend that is projected to escalate significantly in the coming decades (World Health Organization, 2020). The societal ramifications are multi-faceted, extending beyond the individual to affect communities and economies. These include not only the direct costs of treatment but also indirect costs such as productivity losses and a diminished quality of life for those affected (Thornicroft, et al., 2017) (World Health Organization, 2020).

Despite the availability of effective treatments for many mental health disorders, a significant gap persists in healthcare access, which further exacerbates the economic and societal impact of these conditions (Thornicroft, et al., 2017). This lack of access triggers a cascade of economic difficulties, including unemployment, housing instability, and other forms of financial distress, thereby amplifying the societal costs (Thornicroft, et al., 2017)

In some jurisdictions, the economic burden of mental health has even eclipsed that of unemployment, emerging as the costliest social issue (Knapp, McDaid, & Parsonage, 2011). The economic repercussions are intricate and extend to various sectors. Workforce productivity is adversely affected, leading to decreased overall performance and output (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015) (Knapp, McDaid, & Parsonage, 2011). The financial burden also permeates the healthcare system, encompassing costs related to medical care, pharmacotherapy, and ancillary health services (Thornicroft, et al., 2017). Moreover, mental health conditions can impose a 'career penalty,' restricting job opportunities, impeding career progression, and thereby reducing individual earning potential (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015).

At the societal level, individuals with mental health disorders are more likely to be dependent on government-funded assistance programs, disability benefits, and other social services, further inflating the associated expenditures. The cumulative impact of these factors—lower employment rates, reduced productivity, and diminished incomes—can exert a downward pressure on a nation's overall economic growth and competitiveness (Knapp, McDaid, & Parsonage, 2011).

Ultimately, the economic and societal toll of mental health disorders is substantial and manifests through various channels, including workforce impact, healthcare expenditures, reduced earning potential, and increased reliance on social welfare programs. Addressing these issues necessitates a multi-pronged approach that spans both individual-level interventions and broad policy initiatives. Such comprehensive strategies are essential for fostering economic prosperity and enhancing societal well-being, thereby connecting back to the need for accurate and consistent psychiatric diagnosis as discussed in the previous section (World Health Organization, 2020; Thornicroft, et al., 2017).

1.7 Machine Learning: A New Frontier in Psychiatry

Expanding upon the previous discussions on the complexities of psychiatric diagnosis and the economic impact of mental health disorders, the integration of ML in psychiatry emerges as a transformative frontier. This technological evolution, often encapsulated within the emerging field of computational psychiatry, promises to refine psychiatric diagnosis, treatment, and prognosis, thereby potentially alleviating some of the challenges discussed earlier. The surge in data availability, along with computational advancements, has facilitated the development of algorithms adept at deciphering intricate patterns within complex mental health data (Bzdok & Meyer-Lindenberg, 2018; Topol, 2019).

ML models, a cornerstone of computational psychiatry, employ a data-driven methodology, analyzing both historical and real-time data from diverse sources such as electronic health records, neuroimaging, genomics, smart devices, and even social media interactions (Dwyer, Falkai, & Koutsouleris, 2018; Bzdok & Meyer-Lindenberg, 2018). These models offer insights into the biological, psychological, and social determinants of mental health disorders, thereby contributing to the formulation of patient-specific treatment plans (Dwyer, Falkai, & Koutsouleris, 2018; Thornicroft, et al., 2017).

In terms of diagnosis, ML can mitigate subjectivity and enhance the accuracy and consistency of assessments, aligning with the need for precise diagnosis as discussed in earlier sections (Koutsouleris, et al., 2016; Thornicroft, et al., 2017). For example, algorithms have been developed to identify early signs of schizophrenia and bipolar disorder, enabling timely interventions (Koutsouleris, et al., 2016). These techniques also show promise in differentiating between disorders with similar symptomatic profiles, which is potentially the start to addressing the longstanding issue of symptom overlap in psychiatry (Koutsouleris, et al., 2016; Thornicroft, et al., 2017).

On the therapeutic front, ML facilitates a shift towards precision medicine, aligning treatments with individual biological markers and behavioral patterns (Chekroud, et al., 2016; Dwyer, Falkai, & Koutsouleris, 2018). Predictive algorithms can assist in determining optimal drug dosages, identifying likely treatment responders, and even predicting potential side effects (Chekroud, et al., 2016; Topol, 2019).

Moreover, ML is increasingly utilized in the continuous monitoring and management of mental health disorders. Technologies such as remote sensing, wearable devices, and smartphone applications collect real-time data on mental health parameters, offering immediate insights into patients' well-being and treatment responsiveness (J Torous, 2017; Dwyer, Falkai, & Koutsouleris, 2018).

However, the integration of ML in psychiatry, and specifically within the context of computational psychiatry, is not devoid of challenges, including data privacy, algorithmic bias, and the need for multidisciplinary collaboration (Topol, 2019; Bzdok & Meyer-Lindenberg, 2018). Addressing these issues necessitates rigorous evaluation, standardized methodologies, and guidelines that align with both clinical and technological standards (Topol, 2019; Thornicroft, et al., 2017).

Taken together, ML offers a dynamic and promising avenue in psychiatry, with the potential to revolutionize the understanding, diagnosis, treatment, and management of mental health disorders. By bridging the gap between extensive data sets and individualized care, ML is beginning to contribute to a future in psychiatry that is more precise, evidence-based, and patient-centered (Bzdok & Meyer-Lindenberg, 2018; Topol, 2019; Thornicroft, et al., 2017).

1.8 Computational Psychiatry and Disasters

The emerging field of computational psychiatry seeks to employ mathematical models, ML, and computational techniques to understand, predict, and influence mental health outcomes (Wang & Krishnan, 2014). At the intersection of this computational methodology and the escalating concern of mental health in the face of disasters, lies a promising domain for collaborative research and intervention.

Disasters present complex mental health challenges that are multifaceted in nature, involving immediate and long-term psychological effects. The traditional methods of mental health assessment and treatment may fall short in capturing and addressing these complexities (Shultz, et al., 2014). Computational psychiatry offers the potential to enhance the understanding of disaster-related mental health issues by employing data-driven, algorithmic approaches to model the underlying mechanisms of psychological responses to traumatic events.

By integrating large-scale data from various sources, such as public health records, social media, and self-report assessments, computational models can facilitate the identification of patterns, risk factors, and predictors of mental health conditions in the aftermath of disasters (Calhoun, et al., 2012). This approach may enable targeted interventions and personalized treatments that consider individual variability in susceptibility, resilience, and recovery.

Machine learning techniques provide further potential by allowing for the development of predictive models that can identify at-risk individuals or communities prior to or in the immediate aftermath of a disaster (Dwyer, Falkai, & Koutsouleris, 2018). Such early identification can guide mental health interventions, possibly preventing or mitigating severe mental health outcomes.

Additionally, computational psychiatry can facilitate the evaluation of intervention effectiveness, monitoring of recovery trajectories, and optimization of resource allocation in disaster mental health response (Friston, Stephan, Montague, & Dolan, 2014). This alignment of computational methods with traditional psychiatric practices may enhance the adaptability and efficiency of mental health services in rapidly changing disaster scenarios.

However, the integration of computational psychiatry into disaster mental health response is not without challenges. Again, ethical considerations, data privacy, and security concerns, along with the need for multidisciplinary collaboration and adequate training in both computational and mental health fields, are essential elements that must be addressed (Montague, Dolan, Friston, & Dayan, 2012).

While computational psychiatry is still working through these issues, the merger of computational psychiatry with disaster mental health presents a promising and integrative approach for addressing the complex mental health needs arising from catastrophic events. By leveraging the strengths of computational methods and data-driven analysis, this integrative approach has the potential to advance the understanding, prediction, and treatment of disaster-related mental health issues, opening new avenues for timely and effective interventions. This coming at a time when we are seeing a surge in disasters.

1.9 The Intersection of Precision Health, Public Health Data, and Machine Learning

Looking at the complexities of psychiatric diagnosis, the economic impact of mental health disorders, and the transformative role of ML, the intersection of precision health, public health data, and ML emerges as an important shift towards a more integrated and personalized healthcare system. This convergence draws upon the strengths of each domain, aiming to enhance healthcare outcomes at both individual and population levels.

Precision health is fundamentally rooted in the concept of individualized care, focusing on the biological, environmental, and lifestyle uniqueness of each patient (Ashley, 2016; Topol, 2019). It employs a tailored approach to prevention, diagnosis, and treatment strategies, leveraging genetic and other biometric data to meet the specific needs of the individual (Ashley, 2016; Obermeyer & Emanuel, 2016). In contrast, public health adopts a more macroscopic lens, targeting broader communities with the aim of identifying trends, implementing preventive measures, and managing health resources effectively (David W Bates 1, 2014; Char, Shah, & Magnus, 2018).

Machine learning serves as a potential intersection point, processing vast and diverse datasets to find patterns and correlations that may elude human analysis (Rajkomar, Dean, & Kohane, 2019; Topol, 2019). It bridges the gap between the individual-centric focus of precision health and the population-based objectives of public health, thereby harmonizing personalized and public healthcare (Rajkomar, Dean, & Kohane, 2019; Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014).

In the domain of diagnosis, this intersection allows for a nuanced analysis of individual risk factors, incorporating both personal and population-level data (Topol, 2019; Rajkomar, Dean, & Kohane, 2019). ML algorithms can forecast the risk of specific diseases with heightened accuracy, thereby facilitating early interventions and the crafting of tailored preventive measures (Topol, 2019; Obermeyer & Emanuel, 2016).

Furthermore, the collaboration between precision health, public health, and ML enables optimal resource allocation (Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014; Char, Shah, & Magnus, 2018). By understanding the specific needs and risks associated with different population segments, healthcare systems can more effectively and equitably allocate interventions and resources (Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014; Obermeyer & Emanuel, 2016).

Treatment strategies also benefit from this convergence. Machine learning models, informed by individual patient data and broader population trends, can predict responses to various therapeutic interventions (Obermeyer & Emanuel, 2016; Topol, 2019). This data-driven approach can inform the development of individualized treatment plans, thereby elevating success rates and minimizing potential side effects (Obermeyer & Emanuel, 2016; Ashley, 2016).

However, this confluence is not without its ethical and logistical challenges. The management of sensitive health data necessitates robust safeguards, with stringent considerations for privacy, consent, and potential biases (Char, Shah, & Magnus, 2018; Rajkomar, Dean, & Kohane, 2019). Ensuring that the data also adequately represents diverse population groups is paramount to avoid exacerbating healthcare disparities (Char, Shah, & Magnus, 2018; Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014). In parallel, the economic ramifications of merging precision health, public health data, and ML in mental health care are significant. The strategic application of ML to refine diagnostic procedures and customize treatment protocols presents an opportunity to diminish the financial strain associated with mental health conditions. Achieving this requires streamlining treatment methodologies, curtailing unnecessary hospitalizations, and enhancing the efficiency of resource deployment. Further, utilizing public health data to devise and implement preventive strategies for populations at elevated risk could pre-empt the development of mental health issues, further alleviating economic pressures on healthcare infrastructures. However, capitalizing on these economic advantages mandates a thoughtful navigation of the ethical and logistical complexities tied to data privacy and the fair distribution of healthcare resources.

In summary, the intersection of precision health, public health data, and ML heralds a transformative shift in healthcare practice. It integrates the individual-centric approach of precision medicine with the broader, population-focused objectives of public health, employing machine learning and economics as a catalytic agent for this fusion (Ashley, 2016; Topol, 2019; Rajkomar, Dean, & Kohane, 2019). While the collaboration promises a future where healthcare is both personalized and population-centric, significant ethical and logistical hurdles remain to be addressed for its full potential to be realized (Char, Shah, & Magnus, 2018; Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014).

1.10 The Escalating Issue: Mental Health in the Face of Disasters

Continuing from the transformative potential of ML and precision health in psychiatry, another pressing concern that warrants attention and research is the escalating issue of mental health in the context of disasters. The increasing frequency and severity of disasters, both natural and human caused, have highlighted the importance to comprehend and mitigate their psychological ramifications. Such events can inflict enduring psychological trauma, aggravate pre-existing physical and mental health conditions, and give rise to new disorders, particularly when they overwhelm the coping mechanisms of individuals and communities (Goldmann & Galea, 2014; Norris, et al., 2002).

Historically, the focus on the immediate physical health consequences of disasters has often eclipsed the equally significant mental health repercussions. However, the growing acknowledgment of mental health as an integral facet of overall well-being has catalyzed a shift towards a more balanced understanding of the psychological impacts of catastrophic events (Norris, et al., 2002; Goldmann & Galea, 2014).

Disasters exert a profound impact on the social and psychological fabric of communities, leading to a myriad of mental health disturbances such as acute stress, anxiety, depression, and post-traumatic stress disorder (PTSD) (Neria, Nandi, & Galea, 2007; Shultz, et al., 2014). These disruptions disproportionately affect vulnerable populations, including children, the elderly, and individuals with pre-existing mental health conditions (Felix, et al., 2015; Neria, Nandi, & Galea, 2007).

The intersection between mental health and disasters is intricate, necessitating a dual-pronged approach to intervention. Immediate mental health care and support are essential for mitigating acute distress, while long-term strategies are required for the prevention and management of chronic mental health conditions (Shultz, et al., 2014; Reifels, et al., 2013).

The role of mental health professionals in this context is indispensable, and their responsibilities span assessment, immediate intervention, and ongoing care. Effective disaster response necessitates interdisciplinary collaboration, involving sectors such as public health, social services, private industry, and emergency management, to holistically address the complex needs of affected individuals and communities (Reifels, et al., 2013; Paton & Johnston, 2001).

Emerging technologies like telemedicine and online counseling platforms offer innovative modalities for delivering mental health support in disaster-stricken areas, facilitating timely interventions even in geographically isolated or severely impacted regions (Wind, Rijkeboer, Andersson, & Riper, 2020; Shultz, et al., 2014). However, these technological solutions must be seamlessly integrated with traditional, community-based mental health services to ensure comprehensive and effective care (Wind, Rijkeboer, Andersson, & Riper, 2020; Reifels, et al., 2013).

Moreover, preventive measures and preparedness are integral components of a comprehensive strategy to minimize the mental health impact of disasters. Initiatives aimed at building community resilience, fortifying social support networks, leveraging new technologies and devices, and implementing early intervention strategies could hopefully mitigate the psychological toll of these events (Paton & Johnston, 2001; Felix, et al., 2015).

Ultimately, the burgeoning issue of mental health in the context of disasters necessitates a multi-level, integrated approach. This involves the incorporation of mental health considerations into disaster response planning, the utilization of innovative technologies, the strengthening of community resilience, and the implementation of preventive measures. These concerted efforts are critical for mitigating the profound and enduring psychological impacts of catastrophic events (Goldmann & Galea, 2014; Norris, et al., 2002; Paton & Johnston, 2001).

1.11 From Innovation to Ethical Application: Navigating Privacy in Mental Health Advancements

As we navigate the intricate interplay between ML advancements and the utilization of AHRs, it's important to address the new privacy and collaboration challenges that come with this vast amount of private data. This transition underscores the necessity of balancing innovation with strong ethical practice of sensitive data, paving the way for newer and nuanced explorations of implementing privacy measures that ensure the responsible use of technology in mental health care. The progression into the subsequent chapter reflects this critical juncture, emphasizing the importance of safeguarding patient confidentiality while leveraging data to enhance mental health outcomes. While this thesis primarily explores the potential of AHRs in managing crisis events that significantly impact mental health, we begin with an in-depth examination of the privacy concerns associated with AHRs and propose potential solutions to these issues, setting a foundational understanding before delving into the specific applications of AHRs in crisis situations.

References

American Psychiatric Association . (1952). *Diagnostic and Statistical Manual: Mental Disorders.*

American Psychiatric Association . (2017). *DSM 5 Diagnostic And Statistical Manual Of Mental Disorders 5 Ed.* Washington, DC: American Psychiatric Publishing.

Ashley, E. A. (2016). Towards precision medicine. *Nature Reviews Genetics*, 17, 507–522.

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*, https://doi.org/10.1377/hlthaff.2014.0041.

Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Cognitive Neuroscience and Neuroimaging*, https://doi.org/10.1016/j.bpsc.2017.11.007.

Calhoun, V. D., Sui, J., Kiehl, K., Turner, J., Allen, E., & Pearlson, G. (2012). Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder. *Frontiers of Psychiatry*, https://doi.org/10.3389/fpsyt.2011.00075.

Canadian Institute for Health Information. (2023). *Canadian Institute for Health Information*. Retrieved from Codes and Classifications: https://www.cihi.ca/en/submit-data-and-view-standards/codes-and-classifications

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp1714229.

Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., . . . Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(15)00471-x.

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three Approaches to Understanding and Classifying Mental Disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest*, 72-145. https://doi.org/10.1177/1529100617727266.

Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying Mental Health Signals in Twitter. *Workshop on Computational Linguistics and Clinical Psychology*, https://doi.org/10.3115/v1/w14-3207.

David W Bates 1, S. S.-M. (2014). Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*, https://doi.org/10.1377/hlthaff.2014.0041.

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, https://doi.org/10.1146/annurev-clinpsy-032816-045037.

Fazel, M., Hoagwood, K., Stephan, S., & Ford, T. (2014). Mental health interventions in schools 1: Mental health interventions in schools in high-income countries. *Lancet Psychiatry*, 377-387. https://doi.org/10.1016/s2215-0366(14)70312-8.

Felix, E., Afifi, T., Kia-Keating, M., Brown, L., Afifi, W., & Reyes, G. (2015). Family functioning and posttraumatic growth among parents and youth following wildfire disasters. *American Journal of Orthopsychiatry*, https://doi.org/10.1037/ort0000054.

Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F., & Berk, M. (2017). The new field of 'precision psychiatry'. *BMC Medicine*, https://doi.org/10.1186/s12916-017-0849-x.

First, M. B., Reed, G. M., Hyman, S. E., & Saxena, S. (2015). The development of the ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. *World Psychiatry*, https://doi.org/10.1002/wps.20189.

Friston, K. J., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(14)70275-5.

Garies, S., Ng, P., Dickinson, J. A., McDonald, T., O'Beirne, M., McBrien, K. A., . . . Quan, H. (2022). Leaving the Walkman and ICD-9 Behind: Modernizing the Disease Classification System Used by Canadian Physicians. *Health Policy*, https://doi.org/10.12927/hcpol.2022.26907.

Goldmann, E., & Galea, S. (2014). Mental Health Consequences of Disasters. *Annual Review of Public Health*, https://doi.org/10.1146/annurev-publhealth-032013-182435.

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *Journal of Clinical Psychiatry*, https://doi.org/10.4088/jcp.14m09298.

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., . . . Madan, I. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(20)30168-1.

J Torous, J.-P. O. (2017). New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. *Translational Psychiatry*, https://doi.org/10.1038/tp.2017.25.

Kendler, K. S. (2016). The Phenomenology of Major Depression and the Representativeness and Nature of DSM Criteria. *The American Journal of Psychiatry*, 771-780. https://doi.org/10.1176/appi.ajp.2016.15121509.

Knapp, M., McDaid, D., & Parsonage, M. (2011). *Mentalhealthpromotion andmentalillness prevention: The economic case.* London: Department of Health.

Koutsouleris, N., Kahn, R. S., Chekroud, A. M., Leucht, S., Falkai, P., Wobrock, T., . . . Hasan, A. (2016). Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(16)30171-7.

Mathers, C. D., & Loncar, D. (2006). Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLOS Medicine*, https://doi.org/10.1371/journal.pmed.0030442.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Science*, https://doi.org/10.1016/j.tics.2011.11.018.

Monteith, S., Glenn, T., Geddes, J., Whybrow, P. C., & Bauer, M. (2016). Big data for bipolar disorder. *Int J Bipolar Disord*, https://doi.org/10.1186/s40345-016-0051-7.

Neria, Y., Nandi, A., & Galea, S. (2007). Post-traumatic stress disorder following disasters: a systematic review. *Psychological Medicine*, https://doi.org/10.1017/s0033291707001353.

Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry: Interpersonal and Biological Processes*, https://doi.org/10.1521/psyc.65.3.207.20173.

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp1606181.

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., . . . Jordans, M. J. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet*, https://doi.org/10.1016/S0140-6736(18)31612-X.

Paton, D., & Johnston, D. (2001). Disasters and communities: vulnerability, resilience and preparedness. *Disaster Prevention and Management: An International Journal*, https://doi.org/10.1108/eum0000000005930.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, https://doi.org/10.1056/nejmra1814259.

Reifels, L., Pietrantoni, L., Prati, G., Kim, Y., Kilpatrick, D. G., Dyb, G., . . . O'Donnell, M. (2013). Lessons learned about psychosocial responses to disaster and mass trauma: an international perspective. *European Journal of Psychotraumatology*, https://doi.org/10.3402/ejpt.v4i0.22897.

Shorter, E. (1997). *A history of psychiatry: From the era of the asylum to the age of Prozac.* John Wiley & Sons.

Shultz, J. M., Thoresen, S., Flynn, B. W., Muschertl, G. W., Shaw, J. A., Espinel, Z., . . . Cohen, A. M. (2014). Multiple vantage points on the mental health effects of mass shootings. *Current Psychiatry Reports*, https://doi.org/10.1007/s11920-014-0469-5.

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., . . . Bunting, B. (2017). Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry*, https://doi.org/10.1192/bjp.bp.116.188078.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

Wang, W., & Krishnan, E. (2014). Big data and clinicians: a review on the state of the science. *JMIR Medical Informatics*, https://doi.org/10.2196/medinform.2913.

Wind, T. R., Rijkeboer, M., Andersson, G., & Riper, H. (2020). The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. *Internet Interventions*, https://doi.org/10.1016/j.invent.2020.100317.

World Health Organization. (1948). *International Classification of Diseases.*

World Health Organization. (2018). *International Classification of Diseases (11th ed.).* Geneva, Switzerland: WHO.

World Health Organization. (2020). *Mental health and psychosocial considerations during the COVID-19 outbreak*. Retrieved from World Health Organization: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. Annual Review of Clinical Psychology. *Annual Review of Clinical Psychology*, 49-71. https://doi.org/10.1146/annurev-clinpsy-032816-045020.

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensice Psychiatry*, https://doi.org/10.1016/j.comppsych.2014.09.007.

Chapter 2
Privacy, Collaboration and Implementation in AHR Utilization

AHRs have become invaluable assets in mental health research due to their extensive and varied data sets (Kisely, et al., 2009). The availability of longitudinal data allows for deep explorations into patient history, offering insights that would otherwise be unavailable. Yet, the richness of AHRs poses a dilemma as it simultaneously makes them a hotbed for potential breaches in patient confidentiality (Giuffrè & Shung, 2023). Luckily technologies like synthetic data are starting to help address some of these problems. Synthetic data is a technological solution engineered to create data sets that mimic original data structures while de-identifying individual records (Emam, et al., 2012). The intention here is that synthetic data offer a viable alternative to direct data sharing by preserving the utility of AHRs while minimizing privacy risks.

While existing privacy mechanisms like cell-size measure and k-anonymity offer a semblance of data protection, their applicability in the context of AHRs is fraught with challenges. The cell-size measure, although endorsed by Health Canada and the Treasury Board of Canada, is still subject to methodological inconsistencies, as different entities adopt different cell sizes ranging from 5 to 20, depending on the sensitivity of the data. This inconsistency calls into question the uniform applicability of the measure across different datasets and jurisdictions.

Moreover, even sophisticated techniques like the addition of Gaussian noise or cryptographic measures can sometimes introduce a level of complexity that obstructs quick and efficient data utilization. Additionally, these techniques may degrade the quality of data, making it challenging to make patient-level predictions or analyze outliers in the healthcare system efficiently. This points to an overarching issue: the trade-off between privacy and data utility, which remains a persistent concern in the realm of AHRs.

Despite existing limitations, there are promising avenues for advancing the ethical and practical utilization of AHRs in mental health research. For instance, more granular methods of data anonymization, such as differential privacy, are currently being researched and could offer more robust privacy protections without sacrificing data utility (Dwork & Roth, 2020). Moreover, the concept of "reasonable risk" could be standardized across jurisdictions, thus offering a more unified approach to data privacy.

Above all, there is a compelling need for a multidisciplinary collaboration involving lawmakers, clinicians, and academics to reassess and potentially redesign the legislative frameworks that govern the use of AHRs. Such a collaborative effort could lead to more nuanced and adaptable laws that align with the ever-evolving landscape of mental health research and its inherent data complexities.

This chapter aims to shine a light on the intricate issues surrounding the utilization of AHRs in mental health research. While synthetic data presents a promising avenue for mitigating privacy concerns, there are notable gaps and limitations in existing privacy measures. As we move forward, the balancing act between data utility and patient privacy will require a multidisciplinary approach to bring about legislative reforms that are both pragmatic and ethical. Thus, the onus is on a collective effort from all stakeholders to ensure that AHRs can be utilized effectively and responsibly in the field of mental health.

The following paper that sets out to explore the issues with privacy, as it relates to personal data, a critical component of AHRs, is intended to be published, and is currently awaiting journal acceptance.

2.0 Practical Steps in Implementing Privacy Measures with Synthetic Health Data

2.1 Abstract

Privacy concerns related to use of sensitive personal healthcare information are a consistent concern for innovators in both academic and industrial sectors as a barrier to healthcare data access. Synthetic data (new data generated from the original data) is becoming one of the approaches that innovators use to reduce privacy concerns while conducting research or building translational tools. Synthetic data serve to replicate the patterns within the original data, without containing the personal information of "real" participants. In this article, we discuss potential barriers to implementation of synthetic data from a legal and practical perspective.

2.2 Introduction

Health data utilization is important for improving health outcomes but its effective sharing poses challenges due to patient privacy, and confidentiality obligations. Privacy refers to an individual's right to withhold personal information, while confidentiality involves the recipient's duty to protect shared information from unauthorized access (Alpert, 2003; Jr & Clayton, 1996). The current data-sharing processes in healthcare, such as sharing electronic health records (EHRs), require stringent protection of patient confidentiality. To enable patient confidentiality, the process of data storage often modifies data structures, revalues data fields, converts data types, and eliminates potential identifiers to uphold privacy, consequently removing details vital for developing healthcare applications (Office of the Privacy Commissioner of Canada, 2020). One such application is predicting individual patient outcomes, which is an essential aspect of precision medicine, a new field seeking to optimize medical decisions for individual patients using various forms of patient data including demographics and biological parameters. (Gonzales, Guruswamy, & Smith, 2023; Horgan, et al., 2015; MacEachern & Forkert, 2021). The modification of data to enable patient confidentiality causes a barrier in achieving individual-level predictions. Hence, we must explore effective ways to balance data usability with privacy and confidentiality.

Synthetic data may be a way for us to share the data without patient confidentiality concerns, and compromising data structure for making individualized predictions. Synthetic data may offer a more protective approach to sharing information compared to common methods such as de-identification, where 18 types of protected health information like names, medical record numbers, and biometric identifiers are removed from the health data (Chevrier, Foufi, Gaudet-Blavignac, Robert, & Lovis, 2019; Office for Civil Rights, 2022). While these traditional methods remove or replace identifiable information, they can still leave data vulnerable to re-identification through techniques where other data records (i.e. voter list) can be linked to the de-identified health data (Benitez & Malin, 2010). Although synthetic data does not completely eliminate the risk of re-identification, it significantly reduces it by creating entirely new data that mimics the structure and statistical properties of the original data, without directly corresponding to any real individual. Hence, synthetic data can be used in many ways to improve patient care, as researchers can use synthetic data for precision healthcare to improve patient diagnosis and treatment reducing the chances of compromising patient privacy (Chen, Lu, Chen, Williamson, & Mahmood, 2021).

To illustrate the potential of synthetic data to develop healthcare applications for individualized predictions, the US Department of Veteran Affairs utilized the synthetic data of veterans to identify risk factors for chronic illnesses and suicide in the veteran population and reduce the incidence of these events (Purnell, 2020). However, this use case represents an internal tool within US Veteran Affairs. As a result, barriers still exist to accessing the veteran data by external researchers due to patient confidentiality of the individuals within the data. Working towards a more standardized approach for data usage within existing legal frameworks may alleviate confidentiality concerns and allow for increased access to healthcare data and sharing with external researchers. A major concern with this data is privacy-related, where it is crucial to ensure that there are proven standardized methods to balance maintenance of privacy and confidentiality, while preserving the critical information used for individual-level predictions. Another problem with the utilization of synthetic data is de-identification for rare events (e.g. rare healthcare conditions) to ensure that individuals cannot be identified based on rare identifiers. In this commentary, we will discuss the current legislation for the use of synthetic data, practical considerations for creating new legislation or amending legislation to enable the use of synthetic data, and expand on future solutions to improve access to healthcare data.

2.3 Defining the Problems with Synthetic Data Usage

Currently, there are gaps in addressing privacy-related North American legislation on governance of synthetic data use, and the processes to create synthetic data. At this time, various Canadian entities are beginning to act on related privacy and access concerns in health, including the Office of the Privacy Commissioner (OPC) of Canada, which has identified the current laws as a challenge to implementing synthetic data (Office of the Privacy Commissioner of Canada, 2020). Recommendations included increasing the flexibility of PIPEDA (Personal Information Protection and Electronic Documents Act) and changing the current federal legislation to provide additional protection for personal information that has been de-identified and authorized for use for certain purposes (e.g. establishing prohibitions on data matching/re-identification where synthetic data could lead to tracing back to specific individual identities) (Office of the Privacy Commissioner of Canada, 2020).

In addition to the recommendations set by OPC, individual Canadian provinces have also set out their recommendations and steps for using synthetic healthcare data. The Information and Privacy Commissioner of Ontario, for example, put forward guidelines for the de-identification of structured data as early as 2016 (Information and Privacy Commissioner of Ontario, 2016). In 2019, the Winnipeg Regional Health Authority took steps to modernize a more flexible system, enabling more public-facing data to be shared with privacy protecting mechanisms for rare cases (Wilkinson, Green, Nowicki, & Von Schindler, 2020). Alberta is also exploring synthetic data projects related to health data utilization in 2021 to put more information into the hands of researchers and private innovators (Global Newswire, 2021).

Like Canada, the United States does not currently have specific legislation targeted towards the governance and usage of synthetic data for healthcare applications. Existing legislation, like HIPAA (The Health Insurance Portability and Accountability Act of 1996) governs private patient data but does not fully address the complexities of synthetic data usage (Price, 2021). Currently, HIPAA governs identifiable health data, containing 18 types of identifiers, like zip code and demographic information (Price, 2021; Office for Civil Rights, 2022). To allow for the use of patient data in developing precision health tools, the identifiable data types must be removed, or patient consent must be obtained (Price, 2021). However, removing identifiable data types may remove information like demographic information that may be important for developing precision health tools for different groups of people. In this context, synthetic data would be ideal for use in its stead.

Despite the effort in defining and identifying the current regulations surrounding the usage of synthetic data, there still exist challenges surrounding widespread implementation of new regulations for regulating synthetic data. One such challenge is conveying an understanding of the techniques surrounding how synthetic data are generated, whether the synthetic data maintains the same properties as the original data, and if individual identities are protected.

2.4 Practical Considerations

Within the legal framework surrounding data privacy, it is clear that there must be legislation defining both how we evaluate synthetic data to protect the confidentiality of the individuals who are the subjects of the data, and how synthetic data may be used by researchers, governments, and private companies. An example of a challenge related to synthetic data use was the complaint filed under PIPEDA against Facebook/Meta regarding the mishandling of personal contact data on the Facebook social media site (CanLII, 2018). Specifically, when individuals downloaded their personal contact information using the "DYI" or Download Your Information tool, individuals downloaded personal contact information not imported by other users, raising concerns regarding how individual persons' data were used by Meta. The error within the DYI tool was likely due to the inadequacy regarding how Meta tested their new tools with synthetic data; the new tools functioned differently with synthetic data compared to real data. Hence, as the synthetic data were not representative of real data, the bug within the DYI tool was not detected. While Facebook submits that the storage error could not have been detected easily via testing, that is not to say that the error could not have been detected with the appropriate measures in place. This may reflect limitations in the synthetic data used for the testing, or the means by which it was compiled. This indicates the importance of understanding how to evaluate synthetic data in relation to matters including when, where and how such data may be used to support appropriate amendments to legislation. Furthermore, there should be regulations on how synthetic data must be evaluated in terms of whether it functions similarly to real data for creating data tools or testing existing tools. If Meta as the custodian kept a separate system to verify their findings and determine if their processes were working appropriately, this could have avoided the public concerns that the synthetic data use caused.

Despite the need for legislation surrounding synthetic data, one of the biggest barriers to changing our current data privacy legislation to regulate synthetic data use is the ‘disconnect’ between synthetic data research and policy aimed at appropriate legislative amendments. Researchers who study how best to create or test synthetic data often focus on a complex battery of tests including Chi-square, the Kolmogorov-Smirnov test, and simulated cyber-attacks to see whether and how easily they are able to reconstruct the original data from the synthetic data. They use other complex evaluation metrics to demonstrate the probability of revealing individual identities within a given health data set (Choi, et al., 2018; Hayes, Melis, Danezis, & De Cristofaro, 2017). This approach is based on trying to find the optimal method to maximize privacy, while maintaining the details and patterns within the data set so that the data can be used to develop precision health tools. The challenge with such techniques is that they are always changing and their combined use requires an advanced understanding of statistics and other analytical methods. This is not typically an area of expertise for policy and lawmakers. Lawmakers who lack a sufficient understanding of synthetic data creation and use, and the related legal and practical issues associated with the use of synthetic data in applied situations, represent a significant barrier for the creation of sound legislative options.

Due to the innovative nature of synthetic data, another significant barrier to sound legislative changes is the dearth of precedent in law around the use of techniques applied to ensure that synthetic data mimics the original data while maintaining privacy. This underlines the need for policymakers and legal experts to consult with multiple external domain experts who can compare these methods against other privacy-enhancing technologies. This is not only time-consuming and costly but can also be inhibited by a domain expert's inability to articulate the complexities of the issue and their possible biases. Further, external domain experts may have biases and doubts about newer methods, putting at risk the positive appraisal of synthetic data metrics.

In addition to the practical considerations for protecting the privacy of the individuals within healthcare data, we also need to consider legislation that governs how the synthetic data must be used. Hence, it is important to evaluate whether synthetic data use can benefit current projects within academia, industry, or government. While synthetic data represents a useful tool, much of the research done to date does not consider standard data pipelines, such as the CRISP-DM model used in industry (Chapman, et al., 2000). The CRISP-DM model, as seen in Figure 2.1, is a data mining methodology that is useful as a framework for synthetic data, as it provides a clear and concise framework for data preparation, modeling, and deployment. Methodologies, like CRISP-DM are an essential step in applying synthetic health data into company processes, as the companies wishing to implement synthetic data need to incorporate and deploy them within their existing structure.

The final practical issue is how the synthetic data should be evaluated in terms of how well it can protect the privacy of the individuals within the original data. For example, the Treasury Board of Canada, the federal Ministry of Health, and some regional health authorities in Canada use the aggregated cell-size measure to assess the degree of de-identification within patient data (Health Canada, 2019; Treasury Board of Canada, 2020; Wilkinson, Green, Nowicki, & Von Schindler, 2020). The primary intent of this measure is to protect privacy when releasing information about individuals to researchers or the public in a consistent manner (Health Canada, 2019). The cell-size measure looks at every observation of a patient in a dataset to determine how many other observations of the same type exist for other patients. The smaller the number of individuals with the same observation, the greater the patient's uniqueness and risk to privacy exists. To address this issue, organizations such as Health Canada, Treasury Board of Canada, and the Winnipeg Health Authority have set thresholds between 5 and 11 observations that must be the same in other patients or the data must be removed (Health Canada, 2019; Treasury Board of Canada, 2020; Wilkinson, Green, Nowicki, & Von Schindler, 2020). This helps maintain confidentiality by reducing one's ability to guess a patient's identity correctly. It is important to note here that in the realm of big data and the sharing of information, it is impossible to completely eliminate the risk of exposing someone's identity within the context of healthcare data; therefore, we ought to view the issue through the lens of reasonable risks to disclosure. Health Canada, for example, sets a high threshold; it has adopted a risk threshold of 11 patients that must have the same set of indirectly identifying

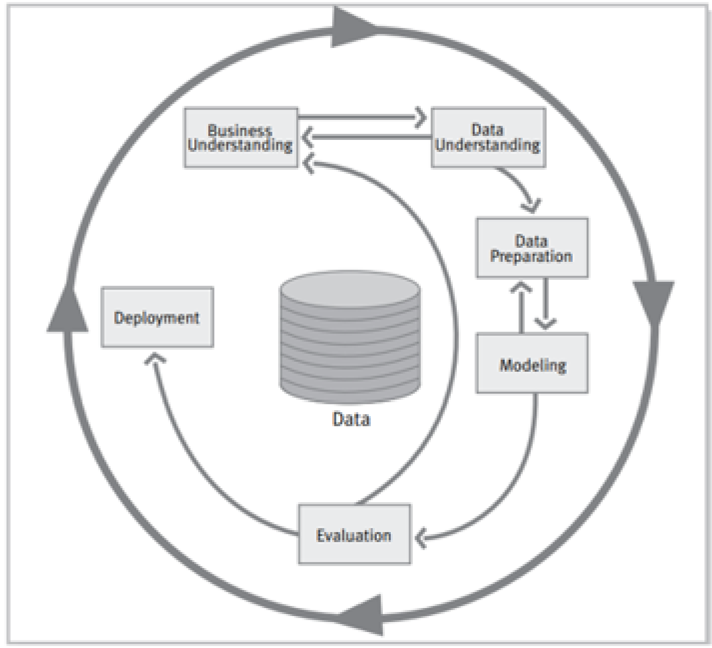

Figure 2.1: The CRISP-DM Model. Adapted from "Cross-Industry Standard Process for Data Mining (CRISP-DM) Manual" (Version 1.0)

variables to provide what they see as a reasonable level of privacy protection when releasing data that may become publicly available (Health Canada, 2019). Similarly, in the United States, HIPAA (Health Insurance Portability and Accountability Act) uses k-anonymity, which dictates that for each of the identifiers, there must be at least k-1 number of records that contain the same (El Emam & Dankar, 2008; Office for Civil Rights, 2022). Although other techniques might achieve a better level of de-identification or anonymization, they do so by adding new complexity to a situation that already has a possible improved solution, further delaying implementation. Multiple privacy enhancing techniques have been recommended as additional layers of privacy protection, including the addition of Gaussian noise, cryptography to scramble data, manual de-identification by a trusted third-party, federated analysis, etc. (Emam, Moura, Locton, Jonker, & Kardash, 2021). There is no exception for synthetic data. In fact, implementing a technique like cell-size measure prior to the generation of synthetic health data would help address a fundamental concern of privacy for policy makers. If something like a cell-size measure is first applied to the original healthcare data, then the data would already meet a common measure of privacy. Hence, any synthetic data generated from the original healthcare data would naturally have the same protection as the pre-assessed original healthcare (Emam, Moura, Locton, Jonker, & Kardash, 2021). This does not mean that the cell-size measure is perfect. Despite the establishment of privacy protection measures like cell-size and k-anonymity for assessing the degree of de-identification within health data, there are limitations for the use of these metrics for assessing privacy within health data. Specifically, for the cell-size measure, preferences for how it is implemented vary between different entities, with the cell size ranging from 5 and 20 depending on how sensitive the data is (Information and Privacy Commissioner of Ontario, 2016; Treasury Board of Canada, 2020; Wilkinson, Green, Nowicki, & Von Schindler, 2020). This is because the cell size measure can be based on a single feature in the dataset or a cluster of features containing sensitive information. Furthermore, for health data for small populations that have unique attributes, the application of cell-size anonymity or k-1 anonymity may not work well for these data sets, as important details within the data may be removed, leading to information loss (El Emam & Dankar, 2008). While other standard techniques can be used, they achieve disclosure risks with synthetic data that are substantially lower than the typical cell size measure of 0.09 and exist outside the legally accepted framework of cell-size (Mosquera, et al., 2023). While techniques like these may be better at accounting for both count size and attribute groupings, they often either degrade the structure of that data, harming the validity of any prediction thereafter, or they become too generalized to make patient-level predictions. In the case of Mosquera et al. (2023), the model was limited where individuals with the greatest number of interactions within the healthcare system were not modelled, emphasizing how techniques like this still struggle to highlight the impacts of important and costly outliers in the healthcare system. In these examples, the cell size measure and k-anonymity still allow for a more reasonable range of variance on count and grouping of features, while still maintaining a well understood privacy measure that is recommended by government agencies (Treasury Board of Canada, 2020).

2.5 Conclusion

The healthcare industry is at a crucial juncture when it comes to the use of sensitive personal healthcare information. While innovators in both the academic and industrial sectors are exploring various methods to reduce privacy concerns, there is a risk that these efforts could remain fragmented if a call to action is not made. If stakeholders continue to work independently on their own business problems or research ideas, it will be challenging to identify and overcome potential barriers to the adoption of health-related synthetic data. To ensure that progress is made in this area, a combined effort between academic and industrial sectors is necessary. First, legal experts should be encouraged to work with innovators to ensure that the use of health-related synthetic data complies with current regulations related to data privacy and protection. Furthermore, effort from law makers, academics and industry specialists should be invested in harmonizing the inconsistent lines of current health related synthetic data research and development of future legislation surrounding data usage and assessment. Second, researchers and industry specialists must engage in open dialogue with stakeholders. This will help to address concerns and increase acceptance of the use of synthetic health data among potential investors or other organizations that may want to work with or develop synthetic health data technologies. As it stands, privacy of healthcare data is a scary topic of discussion amongst many industry specialists. By involving key players in discussions about synthetic data and other privacy-preserving technologies, we can work together to ensure that these approaches are effectively implemented and widely adopted. Finally, investments should be made in research to improve the quality of synthetic health data technologies, in addition to the platforms used to interact or work with the synthetic data. This will help to overcome any skepticism that stakeholders may have about the use of synthetic data with health data. In all, finding solutions to privacy concerns related to sensitive personal healthcare information is essential to facilitating access to valuable healthcare data while protecting patient privacy. By taking a collaborative and comprehensive approach, we can make significant strides in this area and ensure that healthcare data is used to its full potential.

2.6 Bridging Data Privacy and Prediction: Toward a Holistic Approach

This chapter's discussion on the challenges and potential solutions surrounding the privacy of sensitive health information sets the stage for the responsible application of ML within the realm of prediction and research. This also underscores the necessity of a multidisciplinary approach that not only leverages the power of ML algorithms and AHRs but also adheres to stringent privacy standards to protect patient information.

When considering predictive models developed for opioid overdose in the next chapter, it is important to keep in mind the privacy safeguards discussed in here. Ensuring the ethical use of AHRs in such predictive endeavors requires a collaborative effort among data scientists, healthcare professionals, and policymakers to implement privacy measures effectively. Research alone will not help overcome these barriers. Moreover, this intersection opens avenues for future research to refine these models further, incorporating a broader spectrum of data while maintaining the integrity and confidentiality of individual health records.

References

Alpert, S. A. (2003). Protecting Medical Privacy: Challenges in the Age of Genetic Information. *Journal of Social Issues*, https://doi.org/10.1111/1540-4560.00066.

Benitez, K., & Malin, B. (2010). Evaluating re-identification risks with respect to the HIPAA privacy rule. *Journal of the American Medical Informatics Association*, https://doi.org/10.1136/jamia.2009.000026.

CanLII. (2018). *Facebook agrees to stop using non-users' personal information in users' address books*. Retrieved from https://canlii.ca/t/hvs0v

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. Retrieved from https://www.semanticscholar.org/paper/CRISP-DM-1.0%3A-Step-by-step-data-mining-guide-Chapman-Clinton/54bad20bbc7938991bf34f86dde0babfbd2d5a72

Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, https://doi.org/10.1038/s41551-021-00751-8.

Chevrier, R., Foufi, V., Gaudet-Blavignac, C., Robert, A., & Lovis, C. (2019). Use and Understanding of Anonymization and De-Identification in the Biomedical Literature: Scoping Review. *Journal of Medical Internet Research*, https://doi.org/10.2196/13484.

Choi, E., Biswal, S., Malin, B., Duke, J., Stewart, W. F., & Sun, J. (2018). Generating Multi-label Discrete Patient Records using Generative Adversarial Networks (Version 3). *arXiv*, https://doi.org/10.48550/ARXIV.1703.06490.

El Emam, K., & Dankar, F. K. (2008). Protecting Privacy Using k-Anonymity. *Journal of the American Medical Informatics Association*, https://doi.org/10.1197/jamia.m2716.

Emam, K. E., Arbuckle, L., Koru, G., Eze, B., Gaudette, L., Neri, E., . . . Gluck, J. (2012). De-identification Methods for Open Health Data: The Case of the Heritage Health Prize Claims Dataset. *Journal of Medical Internet Research*, https://doi.org/10.2196/jmir.2001.

Emam, K. E., Moura, P., Locton, V., Jonker, E., & Kardash, A. (2021). Practices for Generating Non-identifiable Data: FINAL REPORT . *OFFICE OF THE PRIVACY COMMISSIONER OF CANADA CONTRIBUTIONS PROGRAM* , https://www.priv.gc.ca/en/opc-actions-and-decisions/research/funding-for-privacy-research-and-knowledge-translation/completed-contributions-program-projects/2020-2021/p_2020-21_10/.

Giuffrè, M., & Shung, D. L. (2023). Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. *npj Digital Medicine* , https://doi.org/10.1038/s41746-023-00927-3.

Global Newswire. (2021). *Alberta Project Explores Use of Synthetic Data as Alternative to Health Data Utilization*.

Gonzales, A., Guruswamy, G., & Smith, S. R. (2023). Synthetic data in health care: A narrative review. *PLOS Digital Health*, https://doi.org/10.1371/journal.pdig.0000082.

Hayes, J., Melis, L., Danezis, G., & De Cristofaro, E. (2017). LOGAN: Membership Inference Attacks Against Generative Models. *arXiv*, https://doi.org/10.48550/ARXIV.1705.07663.

Health Canada. (2019). *Guidance document on Public Release of Clinical Information*. Retrieved from https://www.canada.ca/en/health-canada/services/drug-health-product-review-approval/profile-public-release-clinical-information-guidance.html

Horgan, D., Paradiso, A., McVie, G., Banks, I., Wal, T. V., Brand, A., & Lawler, M. (2015). Is precision medicine the route to a healthy world? *The Lancet*, https://doi.org/10.1016/s0140-6736(15)61404-0.

Information and Privacy Commissioner of Ontario. (2016). *De-identification Guidelines for Structured Data*. Retrieved from https://www.ipc.on.ca/resource/de-identification-guidelines-for-structured-data/

Jr, R. C., & Clayton, P. D. (1996). Privacy, confidentiality, and electronic medical records. *Journal of the American Medical Informatics Association*, https://doi.org/10.1136/jamia.1996.96236282.

Kisely, S., Lin, E., Lesage, A., Gilbert, C., Smith, M., Campbell, L. A., & Vasiliadis, H.-M. (2009). Use of administrative data for the surveillance of mental disorders in 5 provinces. *Canadian Journal of Psychiatry*, https://doi.org/10.1177/070674370905400810.

MacEachern, S. J., & Forkert, N. D. (2021). Machinelearningforprecisionmedicine. *Genome* , https://doi.org/10.1139/gen-2020-0131.

Mosquera, L., Emam, K. E., Ding, L., Sharma, V., Zhang, X. H., Kababji, S. E., . . . Eurich, D. T. (2023). A method for generating synthetic longitudinal health data. *BMC Medical Research Methodology*, https://doi.org/10.1186/s12874-023-01869-w.

Office for Civil Rights. (2022). *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. Retrieved from U.S. Department of Health and Human Services: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Office of the Privacy Commissioner of Canada. (2020). *Consultation on the OPC's Proposals for ensuring appropriate regulation of artificial intelligence*. Retrieved from https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/pos_ai_202001/ (accessed 10 May 2022).

Office of the Privacy Commissioner of Canada. (2020). *Office of the Privacy Commissioner of Canada*. Retrieved from https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/pos_ai_202001

Price, N. (2021). Problematic Interactions Between AI and Health Privacy Problematic Interactions Between AI and Health Privacy. *Utah Law Review Utah Law Review* , https://ssrn.com/abstract=3797161.

Purnell, A. (2020). *How synthetic data will improve Veteran health and care*. Retrieved from VA News: https://news.va.gov/81908/synthetic-data-improve-veteran-care/

Treasury Board of Canada. (2020). *Privacy Implementation Notice 2020-03: Protecting privacy when releasing information about a small number of individuals*. Retrieved from https://www.canada.ca/en/treasury-board-secretariat/services/access-information-privacy/access-information-privacy-notices/2020-03-protecting-privacy-releasing-information-about-small-number-individuals.html

Wilkinson, K., Green, C., Nowicki, D., & Von Schindler, C. (2020). Less than five is less than ideal: replacing the "less than 5 cell size" rule with a risk-based data disclosure protocol in a public health setting. *Canadian Journal of Public Health*, https://doi.org/10.17269/s41997-020-00303-8.

Chapter 3
Opioid Overdose Prediction using Machine Learning and AHRs

This section specifically focuses on opioid overdose prediction through ML algorithms applied to AHRs. This study is both timely and relevant given the alarming surge in opioid overdose cases, particularly accentuated by the COVID-19 pandemic.

The opioid crisis represents not merely an isolated public health concern but an issue intricately connected with mental health. Opioid dependence and abuse have been linked with a myriad of mental health disorders including depression and, highlighting the need for an interdisciplinary, data-driven approach in understanding and mitigating its impact. This case study capitalizes on ML techniques to discern patterns and risk factors that traditional statistical methods might overlook. By using AHRs, which offer a rich, longitudinal treasure trove of healthcare encounters, prescriptions, and outcomes, the model endeavors to predict the likelihood of opioid overdose with improved accuracy and granularity.

Critiques of the Paper

While the study makes significant strides in the application of AHRs for predicting opioid overdose risks, several limitations and areas for improvement warrant discussion. Firstly, ML models are only as robust as the data they are trained on. Given that AHRs are administrative in nature, they may lack certain clinical or psychosocial variables crucial for a more holistic understanding of opioid overdose risks. The omission of these variables could inadvertently introduce bias or limit the model's predictive power a challenge that was already noted in Chapter 1.

Secondly, the study does not address the ethical implications tied to the utilization of AHRs in the realm of mental health research comprehensively, as discussed in Chapter 2. While AHRs offer an invaluable resource, they also come fraught with privacy concerns. The operationalization of privacy safeguards, which the dissertation covers in Chapter 2, should be revisited and explicitly linked to the methodologies employed in this particular case study.

Future Research Directions

As for future avenues of research, this case study can serve as a stepping stone for more expansive, multimodal predictive models. Given that mental health and opioid use disorders are multifactorial in nature, incorporating data from wearable devices -which capture real-time physiological and behavioral markers -could add another layer of precision to the predictive algorithms.

Furthermore, it would be insightful to extend the ML models to post-disaster scenarios. As the final section of this dissertation suggests, AHRs have untapped potential in forecasting mental health outcomes in such contexts. By refining the algorithms to account for the unique stressors and healthcare utilization patterns that emerge post-disaster, a more nuanced understanding of opioid overdose risks within these high-stress environments could be achieved.

In conclusion, this section of the dissertation showcases the efficacy of leveraging ML algorithms and AHRs to predict opioid overdose risks, thereby contributing to the broader agenda of optimizing mental health care through data-driven methodologies. However, it also underscores the need for interdisciplinary collaboration and ethical scrutiny to fully realize the transformative potential of AHRs in mental health research.

Please note that this manuscript is in the process of submission to academic journals and awaits final confirmations from all contributing authors.

3.0 Population-level individualized prospective prediction of opioid overdose using machine learning and administrative health data

3.1 Abstract:

The opioid overdose epidemic has rapidly expanded in North America, with rates accelerating during the COVID-19 pandemic. This study aims to develop and validate a population-level individualized prospective prediction model of opioid overdose (OpOD) using ML and administrative health data in Alberta, Canada. Our analysis of cohort data from 2018 to 2020, approximately 4 million people, showed that the model achieved a balanced accuracy of 83.7%, 81.6%, and 85.0% in each year, respectively, with corresponding area under the receiver operating characteristic curve (AUC) values of 89.2%, 86.9%, and 89.8%. Drug/alcohol abuse/dependence, all substance use claims, depression, neurotic/anxiety/obsessive compulsive disorder and superficial skin injury/contusion/non-serious burns derived from variables by the Canadian Institute for Health Information (CIHI), and physician claims of depression are top predictors for OpOD. This study provides a foundation for a sustainable and adaptive modeling framework for individual-level prospective OpOD prediction at a population-level, which can inform targeted interventions, policy, and treatment planning.

3.2 Introduction

Opioid overdose is a rapidly growing epidemic in North America. In Canada, 30,843 people have died from apparent opioid-related overdose, between January 2016 and March 2022 (Government of Canada, 2023; Vojtila et al., 2020). During the COVID-19 pandemic, overdose rates have only accelerated, with a 24% increase in opioid-related poisonings in the first half of 2022 alone (Government of Canada, 2023). The United States is facing a similar crisis reporting 212,892 opioid-related deaths between 2017 and 2020 (National Institute on Drug Abuse, 2022). Given the new realities of the COVID-19 pandemic and the acceleration of the size and scope of the opioid crisis in North America, we need to explore all potential risks related to drug overdose and death and identify actionable factors for intervention to help patients and their communities.

Rising research interests in the clinical applications of OpOD prediction have emerged due to the increased availability of cross-linked population-level administrative health data and Electronic Health Records (EHR) data (Tseregounis & Henry, 2021). The use of advanced machine-learning algorithms further enables the creation of generalizable models capable of making individual-level predictions, demonstrated by prior work of successful prediction of opioid use disorder and risks of adverse outcomes (Liu et al., 2022; Sharma et al., 2022). One challenge of existing studies is the lack of population-level representative data, a problem exacerbated by non-universal or stratified access to health insurance. For example, in the United States in 2018, public health plans covered only 34.4% of the population, and 8.5% were not insured at all (Berchick et al., 2019). Non-representative data could introduce bias into the studies, affecting the generalizability and reliability of the predictions.

Predicting opioid overdoses with ML using data and algorithms can enable targeted interventions to prevent such overdoses from occurring at both the individual and community levels. Current clinical applications of overdose predictions mainly involve cross-linking population-level Electronic Health Records (EHR) data with other forms of administrative data, focusing on a broad, community-based approach rather than individual-level analysis.

Advanced machine-learning algorithms could enable precision models with the capability for individual-level predictions (Tseregounis & Henry, 2021). This can involve analyzing various types of data, such as medical records, demographic data, location-based variables, and other health-related information, to identify patterns and trends that may indicate an increased risk of overdose or death. For example, Ellis et al. (2019) found that they could train a machine learning model to predict patients receiving a diagnosis of substance dependence using EHR patient data. The model was able to predict substance dependent patients correctly 92% of the time and non-substance dependent patients correctly 76% of the time, with pain related symptoms and mental health issues as key factors for predicting OpOD (Ellis et al., 2019).

Other researchers have looked at neighborhood-level variables related to opioid overdose deaths in their efforts to develop intervention resources. Schell et al. (2022) looked at 206 neighborhood-level demographic variables derived from the US Census Bureau's American Community Survey from 2016–2019. The model examined injuries in neighbourhoods, regardless of residence within that neighbourhood, and then examined specific factors of the community. Factors include: rates of education, income, employment, social isolation, educational attainment, income, disability, employment, and age of housing (Schell et al., 2022). This model was able to identify 40 predictors that together explained approximately 17% of the variance in fatal overdose rates for those individuals. Using a model like this can help hospitals and policy makers make more informed decisions on the opioid crisis based on communities.

In the current study, we developed an administrative health dataset based on population-level administrative health data from the Government of Alberta, Canada. Administrative health data were already collected and updated routinely, providing a sustainable data source. We applied a machine-learning algorithm validated first using a held-out sample, and longitudinally validated using subsequent cohorts. For example, for the 2017 cohort, predicting variables were developed based on information available in 2017, and a model was trained using this feature set to predict OpOD outcome in 2018. We then temporally validated this model by providing the model with predicting variables based on 2018, to predict OpOD outcomes in 2019. This approach aimed to assess the model's ability to generalize and perform accurately on new, unseen data within a different time frame. The model's performance was similar and promising across the two timeframes, suggesting that establishing an adaptive modeling framework with continually updated models could be a viable approach for maintaining its accuracy and generalizability on individualized OpOD prediction. In this study, we aim to validate a novel sustainable and adaptive modeling framework for individual-level prospective OpOD prediction.

3.3 Methods

3.3.1 Data source

This study was a retrospective study where we included individual-level information of different types (e.g., demographic, socio-economic, health utilization) that were collected by the Alberta Ministry of Health and cross-linked and analyzed by the authors. The linked administrative health data were prepared based on Alberta Health Care Insurance Plan (AHCIP) Practitioner Claims, the National Ambulatory Care Reporting System (NACRS), the Canadian Institute of Health Information Discharge Abstract Database (DAD), the AHCIP Population Registry Database, Alberta Pharmaceutical Information Network (PIN) database, data from the Canadian Institute of Health Information (CIHI), and Alberta Health Services Drug Supplement Plan database (AHSDSP). All databases accessed are universal within the Alberta population.

3.3.2 Study Cohort

Four cohorts of all Alberta residence (approximately 4 million) were developed based on data retrieved from fiscal year 2017, 2018, 2019, 2020 and 2021 (April 1st to March 31st of the following year). For each cohort, we included all individuals with an active AHCIP status, who had used the system in the past two years.

The prospective OpOD outcome was derived from the fiscal year following the cohort (e.g., OpOD events in 2018 for the 2017 data) based on the International Classification of Diseases, Ninth Revision (ICD-9) code 965.0 in practitioner claims data, and International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10), code T40.1, T40.2, T40.3, T40.4 and T40.6, in ambulatory, inpatient data (Lo-Ciganic et al., 2019). The OpOD status was binary coded and labeled as 1 if a patient had at least 1 incident of OpOD in the fiscal year following the cohort, and 0 if no incident was found in the administrative health records.

Candidate predicting variables or "features" for machine-learning, were developed based on the cohort's fiscal year, with a total of 368 features, including health system utilization indicators (e.g., number of family physician visits), demographics (e.g., age, sex), opioid specific indicators (e.g., opioid use disorder), substance abuse and related disorders (e.g., alcohol, nicotine), and CIHI groupers (Canadian Institute for Health Information, 2023) that identified other physical and mental health indicators (e.g., chronic pain, hepatitis, depression).

3.3.3 Data preparation and modeling pipeline

SAS 9.4 and SAS Viya Data Studio software were used for data preparation. Model features that represent frequency of occurrence and binary risk indicator had no missing data. Zero is interpreted as zero occurrence or lack of evidence.

The prepared data were then processed through a modeling pipeline (see Figure 3.1) developed using SAS Viya Model Studio software, version V.03.05. Because OpOD is a rare event in population data (e.g., 0.10% of the population in 2017), there's a severe class imbalance that would impact model building (Cartus et al., 2023). Class imbalance refers to a situation in machine learning where the number of observations in one class significantly outweighs the observations in the other class, in this case, the instances of OpOD compared to non-OpOD. This imbalance can lead to biases in the model as a model can simply predict all cases as the majority class to produce high accuracy. In our exploration of the imbalanced data we ran a number of single Gradient boost models on the data, all achieving roughly a 0% sensitivity. To address the class imbalance in our ML pipeline, we used an under-sampling technique and devised 50 subsamples that each included all subjects within the OD cohort, joined with a stratified random sample drawn from subjects with no OpOD records, matched by Age, Sex, and sample size. Each of the subsamples has a 1:1 ratio of OpOD and no OpOD subjects, allowing the ML model to learn characteristics from the OpOD patients. The 50 subsamples were further split into training and validation sets, where 70% of the data were used for model training and 30% of the data were used for validation. This split ran through Gradient boosting nodes in SAS Viya to learn base models for classifying OpOD, optimizing for logistic loss, a measure of how well the model can make correct classifications. Each of the 50 models were setup to perform auto-tuning, which performed adjustments to the following parameters: number of trees, number of inputs to consider for split, learning rates, subsample rates, L1 & L2 regularization (SAS, 2019). A Gradient boosting model is a machine learning algorithm that iteratively combines multiple weak decision trees to create a stronger, more accurate model by adjusting the weights of the trees and reducing errors in the predictions. The parameters optimization method used for our model was a grid search algorithm, with the initial values used for the baseline model set to default values provided by the SAS model (SAS, 2019). The 50 models trained and validated using the subsamples were then put through an ensemble node. The ensemble node took the average of the predicted probabilities to combine the models and determine the top contributing features. By averaging the probabilities, the ensemble model reduces the impact of individual model biases or errors to improve the prediction accuracy. A predicted probability of 0.5 was used as a threshold to classify OpOD (>= 0.5) and No OpOD (< 0.5). The entire 2018, 2019, 2020 cohorts were used as testing data to evaluate the ensemble model (see Table 3.2). The top five features of the models were evaluated based on the ranking of feature importance in the ensemble model. In SAS, relative feature importance is a metric used to quantify the contribution of each feature to the predictive performance of the model. It helps to identify the more important features driving the model predictions and is valuable for interpreting the models outputs (SAS, 2019).

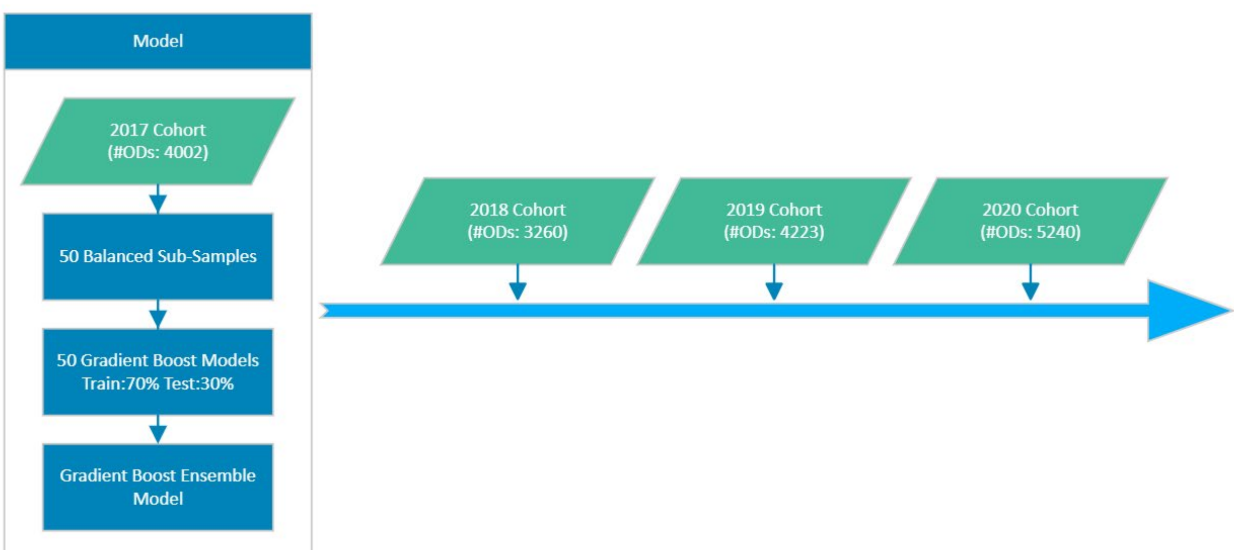

Figure 3.1. ML pipeline flow chart. Fifty subsamples with 1:1 ratio of OpOD and non-OpOD were first derived from the 2017 cohort. Fifty gradient boosting models were trained and validated for each sample, then ensembled. The ensemble model was tested using 2018, 2019, and 2020 cohort respectively.

3.4 Results

For each cohort (2017-2020), the treated prevalence rate of OpODs in the population remained under 0.2% (Table 1). Despite the large imbalance in our data the final ensemble model obtained an AUC of 89.17%, a balanced accuracy of 82.74%, an average of sensitivity of 75.8% and specificity of 89.7%, from the reserved 2017 validation data. The trained model was then applied to the full 2018-2020 cohorts (N=4,095,364-4,203,233). The sensitivity in the subsequent years 2018-2020 achieved 78.1%, 68.4%, and 77.9% respectively, while specificity was 89.3%, 94.8%, and 92.1%, respectively (Table 3.1). From the 368 features, the top five predictors were CIHI drug & alcohol abuse, substance abuse records from claims, CIHI Depression, CIHI anxiety & obsessive compulsive disorder, and physician claims of depression (Table 3.2). Claim records were stronger predictors than inpatient records, emergency department records, and ambulatory records. Although the model was trained on 2017 administrative health record data, the balanced accuracy of prospective prediction on longitudinal testing data in the entire cohorts of 2018, 2019 and 2020 were greater than 80% obtaining 83.7%, 81.6%, and 85.0% (Table 3.1).

Table 3.1. Model performance

Cohort	**N Total**	**$N_{patient}$ OpOD (%)**	**Balanced Accuracy**	**Sensitivity**	**Specificity**	**AUC**
2018	4095364	3260 (0.08)	0.837	0.758	0.893	89.2%
2019	4163304	4223 (0.10)	0.816	0.684	0.948	86.9%
2020	4203233	5240 (0.12)	0.850	0.779	0.921	89.8%

Notes. AUC denotes Area Under the Receiver Operating Characteristics Curve.

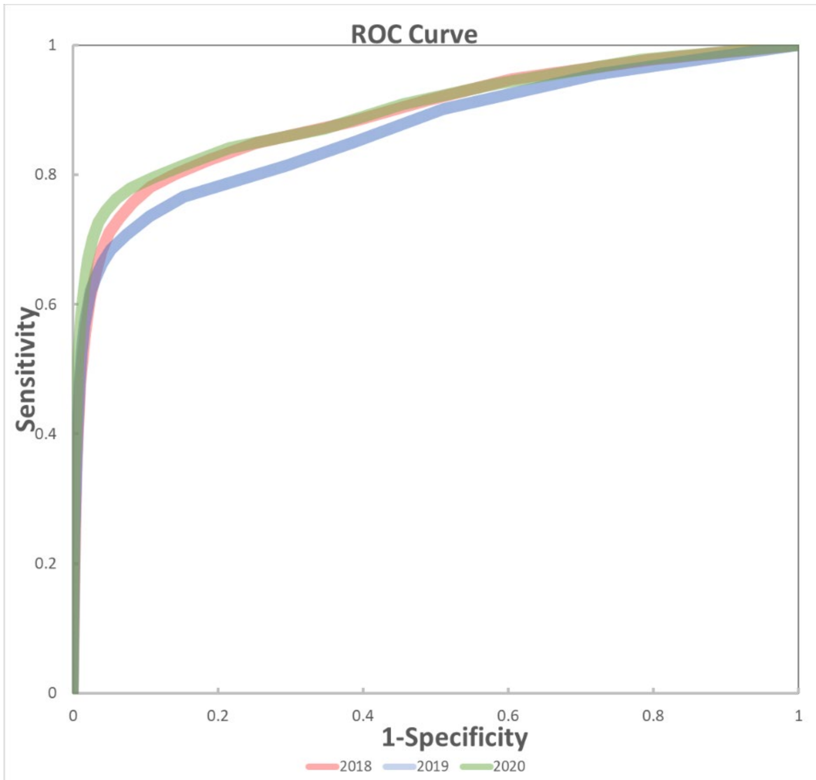

Figure 3.2. Receiver Operating Characteristics Curve (ROC) for the years 2018-2020: Model evaluation for predicting opioid overdose in a given year.

3.4.1 Top predictive features

The prospective prediction performance includes CIHI drug/alcohol abuse/dependence (Q07), CIHI depression (Q04), CIHI neurotic/anxiety/obsessive compulsive disorder (Q11), CIHI Superficial skin injury/contusion/non-serious burns (I43), and Depression from claims data. Note that depression from claims data were developed by the authors, independent from CIHI's method of developing depression indicator (Canadian Institute for Health Information, 2023). These features had relative importance of 1.00, 0.60, 0.46, 0.40, and 0.36, respectively (Table 3.2). Other top ranked features are consistent with risk factors reported in the literature, including substance abuse and substance abuse related health utilization, mood and anxiety disorder related claims, and physician health indicators such as back pain, and skin wounds.

Table 3.2. Top five feature rankings from the ensemble model based on SAS VIYA Outputs

Feature Name	Source	Relative Importance	Data Type
Drug/alcohol abuse/dependence	CIHI	1.00	BINARY (1=yes, 0=no)
Depression	CIHI	0.60	BINARY (1=yes, 0=no)
Neurotic/anxiety/obsessive compulsive disorder	CIHI	0.46	BINARY (1=yes, 0=no)
Superficial skin injury/contusion/ non-serious burn	CIHI	0.40	BINARY (1=yes, 0=no)
Depression	Claims	0.36	INTERVAL

3.5 Discussion

In this study, we developed a machine-learning model to prospectively predict OpOD in the general population. We report high model performance achieving a balanced accuracy of 83.7%, 81.6%, and 85.0% and AUC of 89.2%, 66.9%, and 89.8%, in 2018, 2019 and 2020 cohort, respectively. Our model identified a number of features that were stable across training models at a population level. The two highest ranked predictive features in the model include CIHI Drug/alcohol abuse/dependence (Q07) and CIHI Depression (Q04), suggesting that subjects with substance use disorders and depression are at high risk of OpOD.

To improve the predictive capabilities of the model, we also experimented with training the model on data from 2018. When the model was trained on the 2018 data, the predictions for the 2019 opioid overdose cases exhibited a very similar shift with higher specificity compared to other years. This observation may indicate a difference in the 2019 data due to factors that occurred in that year, data quality, or other underlying factors.

Our successful demonstration of OpOD prediction using a representative, sustainable data set suggest technical feasibility to develop population-level OpOD risk screening tool. E.g., identify individuals at high risk of OpOD may facilitate targeted preventative intervention. While our approach may be suitable for generalized risk screening in the context of public health, it is not yet practical in clinical settings at this time because physicians could increase harm by unnecessarily cutting down on pain management prescriptions dosage based on the high number of false positive predictions. However, in our exploration of the data, we suggest that depending on the use case priority, the low positive predictive value can be mitigated by selecting a higher probability threshold for classification, at the expense of missing more subjects that will OpOD in the following year. Adjusting cutoffs would likely provide more clinical utility, with various specialists benefitting from different cutoff values. For example, at cutoff values between 0.5 and 0.8 this information may be useful as a flag for risks and help guide discussions of mitigation with clients (see Figure 3.2 to visualize sensitivity and specificity trade-offs). For example, in primary care settings, a higher sensitivity may be preferred to ensure that no cases are missed, while in a population-level risk screening settings a higher specificity may be more important to avoid unnecessary interventions.

In the case of the government, our model could help to inform policy on high-risk communities, predict changes in opioid overdose rates, prioritize interventions for individuals that are at greatest risk, and help to model costs associated with opioid overdose rates as they continue to climb. We don’t know the best-cutoff value for all possible scenarios. The ability to adjust cutoffs based on clinical and policy needs may provide more utility and improve the overall effectiveness of the model in identifying and mitigating opioid overdose risks going forward. However, it remains a challenge to effectively educate clinicians and decision makers to apply cut-offs under different situations, to maximize the benefits and minimize risks.

Moreover, given the high-risk profile of individuals with substance use disorders and depression for OpOD, it may be valuable for the Ministry of Health to consider directing resources towards comprehensive and targeted public health interventions. This may include increased mental health support, substance use disorder treatment programs, and community outreach initiatives. Optimizing the classification threshold of our prediction model based on the desired balance between sensitivity and specificity for different applications could provide more utility and facilitate more targeted interventions for high-risk populations, and to explore more practical iterations for the clinical setting.

To ensure the successful integration of the overdose prediction tool into clinical practice, several areas could be addressed, such as implementation strategies, clinician training, and evaluation methods. In such cases it would be wise to adopt a phased implementation approach to manage the integration of the model. Starting with pilot programs in selected clinics which could provide valuable feedback and allow for necessary adjustments before a broader rollout. Collaboration with IT departments would also be necessary to ensure the tool integrates seamlessly with existing electronic health records (EHR) systems. Addressing technical issues early on would also prevent disruptions in clinical workflows and ensure clinician engagement, data accuracy, and security. From there a comprehensive training program for clinicians would be important, especially for wide rollout. This program could focus on the operational aspects of the tool, how to interpret the results, and how to integrate the tool into existing clinical workflows. Training could include workshops, online modules, and hands-on sessions to ensure all clinical staff are competent in using the tool. Continuous education and support would also need to be provided to address any issues that arise during the initial phases of implementation. Finally, it would be important to conduct regular surveys and focus groups with clinicians and patients to assess the tool's acceptability and identify areas for improvement. Implementing continuous monitoring and evaluation metrics to measure the tool's impact on patient outcomes, clinician workflow, and overall efficiency would be the next reasonable step. This information would be able to guide further refinements and demonstrate the tool's value in clinical settings over time.

By focusing on these key areas, the overdose prediction tool could be effectively integrated into clinical practice, enhancing the ability of clinicians to predict and prevent opioid overdoses, ultimately improving patient outcomes and care quality.

3.6 Limitations

The current study presents several limitations that should be considered when interpreting the results and drawing conclusions. First, we found potential biases stemming from incomplete or inaccurate data. As the study relies on administrative health records, there might be missing or incorrect information that could affect the quality and validity of the results. This is especially true given that mental health related measures are often addressed outside the public health system, as they are generally not covered. It is essential for future research to ensure the accuracy and completeness of the data to obtain more reliable outcomes.

Second, selection bias might limit the generalizability of the study findings. The data used in this study were drawn from Alberta, which might not be representative of other populations or regions. Consequently, the model's performance and identified risk factors may not be generalizable to other settings. Further research should aim to validate the model using data from diverse populations to ensure its applicability across different contexts.

Third, the potential loss of valuable information due to under sampling is another limitation. The study employed an under-sampling technique to address class imbalance, which might have resulted in the exclusion of some relevant data. While this study did its best to mitigate this by using 50 different sampling models, this loss of information could have impacted the model's performance and ability to identify critical risk factors. Future studies should explore alternative techniques for handling class imbalance to mitigate the potential loss of information.

Additionally, the low positive predictive value (PPV) limits the model's practical clinical application. With the best model achieving a PPV of 1.32%, there is a high rate of false positive predictions, which could lead to unnecessary interventions and costs. To enhance the model's clinical utility, future research should focus on improving the PPV or adjusting cut-offs based on different clinical and policy applications.

Lastly, the large number of features used in the model may lead to overfitting or multicollinearity issues. While the custom derived features in this study are more specific towards mental health disorders and where they were recorded, there is likely multicollinearity with the CIHI groupers. Future research should address these issues by employing feature selection techniques and regularization methods to optimize the model's complexity and avoid potential problems associated with overfitting and multicollinearity.

Future research should address these limitations and explore ways to improve the model's performance and utility in both clinical and policy settings. By refining the model and addressing its current limitations, researchers can better identify high-risk individuals and inform targeted interventions to reduce opioid overdose rates.

3.7 Conclusion

This study presents a novel approach to predict individual-level prospective opioid overdose outcomes using population-level administrative health data from Alberta, Canada. Despite the challenges of class imbalance and low positive predictive value, the developed machine-learning model demonstrated reasonable performance in terms of AUC, balanced accuracy, sensitivity, and specificity. The top predictive features identified in the model highlight the importance of substance use disorders, mental health factors, and healthcare utilization patterns in predicting opioid overdose risk.

While the current model may not be directly applicable in clinical settings due to the high number of false positives, it shows potential for informing public health strategies, policy-making, and resource allocation. Adjusting the classification threshold based on the intended use case could provide more utility and facilitate targeted interventions for high-risk populations. Future research should aim to address the limitations of the current study, refine the model, and explore its practical implementation in both clinical and policy contexts.

By developing a sustainable and adaptive modeling framework for opioid overdose prediction, this study contributes to the growing body of research on opioid overdose prevention and management. Our findings emphasize the need for a multidisciplinary approach to address the complex interplay of factors contributing to opioid overdose risk and the importance of leveraging population-level data for more effective prevention strategies.

3.8 From Opioid Overdose to Developmental Disorders in the Pandemic Era

As we conclude our examination of opioid overdose prediction using advanced data analytics, we pivot to a related but distinct public health challenge magnified by the COVID-19 pandemic: the rise in developmental disorders and utilizes comparable data-driven techniques to explore the utilization changes in this vulnerable population. Both segments collectively stress the importance of leveraging technological advancements and comprehensive data analysis to inform public health strategies. The continuity between these topics lies in the shared goal of enhancing healthcare interventions through the examination of the value of AHRs with ML. This approach aims to explore targeted interventions, improve support systems, and ultimately, make recommendations to shape more resilient public health policies that can navigate the complexities of health issues in a post-pandemic world.

In moving from the topic of opioid overdose to developmental disorders, we bridge two critical areas of public health research. This progression not only highlights the adaptability and breadth of data-driven public health research but also calls attention to the ongoing need for integrated solutions that consider the wide-ranging impacts of the pandemic and other potential disasters on community health and individual well-being.

References

Berchick, E. R., Barnett, J. C., & Upton, R. D. (2019). *Health insurance coverage in the United States: 2018*. https://census.gov/content/dam/Census/library/publications/2019/demo/p60-267.pdf

Canadian Institute for Health Information. (2023). *Codes and Classifications*. https://www.cihi.ca/en/submit-data-and-view-standards/codes-and-classifications

Cartus, A. R., Samuels, E. A., Cerdá, M., & Marshall, B. D. L. (2023). Outcome class imbalance and rare events: An underappreciated complication for overdose risk prediction modeling. *Addiction*, *118*(6), 1167–1176. https://doi.org/10.1111/add.16133

Ellis, R. J., Wang, Z., Genes, N., & Ma'Ayan, A. (2019). Predicting opioid dependence from electronic health records with machine learning. *BioData Mining*. https://doi.org/10.1186/s13040-019-0193-0

Government of Canada. (2023). *Opioid- and Stimulant-related Harms in Canada*. Government of Canda. https://health-infobase.canada.ca/substance-related-harms/opioids-stimulants/

Liu, Y. S., Kiyang, L., Hayward, J., Zhang, Y., Metes, D., Wang, M., Svenson, L. W., Talarico, F., Chue, P., Li, X. M., Greiner, R., Greenshaw, A. J., & Cao, B. (2022). Individualized Prospective Prediction of Opioid Use Disorder. *Canadian Journal of Psychiatry*. https://doi.org/10.1177/07067437221114094/ASSET/IMAGES/LARGE/10.1177_07067437221114094-FIG2.JPEG

Lo-Ciganic, W.-H., Donohue, J. M., Yang, Q., Huang, J. L., Chang, C.-Y., Weiss, J. C., Guo, J., Zhang, H. H., Cochran, G., Gordon, A. J., Malone, D. C., Kwoh, C. K., Wilson, D. L., Kuza, C. C., & Gellad, W. F. (2022). Developing and validating a machine-learning algorithm to predict opioid overdose in Medicaid beneficiaries in two US states: a prognostic modelling study. *The Lancet Digital Health*, *4*(6), e455–e465. https://doi.org/10.1016/S2589-7500(22)00062-0

Lo-Ciganic, W.-H., Huang, J. L., Zhang, H. H., Weiss, J. C., Wu, Y., Kwoh, C. K., Donohue, J. M., Cochran, G., Gordon, A. J., Malone, D. C., & others. (2019). Evaluation of machine-learning algorithms for predicting opioid overdose risk among medicare beneficiaries with opioid prescriptions. *JAMA Network Open*, *2*(3), e190968–e190968.

National Institute on Drug Abuse. (2022, January 20). *Number of National Drug Overdose Deaths Involving Select Prescription and Illicit Drugs*. Https://Nida.Nih.Gov/Research-Topics/Trends-Statistics/Overdose-Death-Rates. https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates

SAS. (2019). *Exploring SAS Viya*. https://support.sas.com/content/dam/SAS/support/en/books/free-books/exploring-sas-viya-data-mining-machine-learning.pdf

Schell, R. C., Allen, B., Goedel, W. C., Hallowell, B. D., Scagos, R., Li, Y., Krieger, M. S., Neill, D. B., Marshall, B. D. L., Cerda, M., & Ahern, J. (2022). Identifying Predictors of Opioid Overdose Death at a Neighborhood Level With Machine Learning. *American Journal of Epidemiology*, *191*(3), 526–533. https://doi.org/10.1093/aje/kwab279

Sharma, V., Kulkarni, V., Jess, E., Gilani, F., Eurich, D., Simpson, S. H., Voaklander, D., Semenchuk, M., London, C., & Samanani, S. (2022). Development and Validation of a Machine Learning Model to Estimate Risk of Adverse Outcomes Within 30 Days of Opioid Dispensation. *JAMA Network Open*, *5*(12). https://doi.org/10.1001/jamanetworkopen.2022.48559

Tseregounis, I. E., & Henry, S. G. (2021). Assessing opioid overdose risk: a review of clinical prediction models utilizing patient-level data. In *Translational Research* (Vol. 234, pp. 74–87). Mosby. https://doi.org/10.1016/j.trsl.2021.03.012

Vojtila, L., Pang, M., Goldman, B., Kurdyak, P., & Fischer, B. (2020). Non-medical opioid use, harms, and interventions in Canada–a 10-year update on an unprecedented substance use-related public health crisis. In *Drugs: Education, Prevention and Policy* (Vol. 27, Issue 2). https://doi.org/10.1080/09687637.2019.1645094

Chapter 4
Treated Prevalence of Developmental Disorders in Alberta’s Health Care System

This section is devoted to investigating the complex relationship between developmental disorders and associated comorbidities, utilizing machine learning algorithms applied to AHRs. The subject matter is of particular importance, given that developmental disorders such as Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) often present with additional health issues like anxiety and depression. Understanding these comorbidities is essential for holistic patient care and effective intervention planning.

Critiques of the Paper

The study's attempt to use AHRs and machine learning to explore comorbidities in developmental disorders is valuable. However, there are limitations worth discussing, the first critique lies in the intrinsic nature of AHRs, which are generally medical-centric and designed primarily for administrative utility rather than research. Such records do not include comprehensive developmental or educational milestones, which are critical for understanding the full scope of developmental disorders. The omission of these variables may compromise the models' validity and reliability, effectively producing results that are circumscribed in their interpretability.

The second limitation involves the lack of attention paid to the ethical dimensions of using AHRs in developmental disorder research. While AHRs offer a wealth of data, they also pose concerns regarding data privacy and consent, especially considering that many subjects are minors. Thus, the ethical implications of using AHRs are non-trivial and demand an even more rigorous approach to ensure alignment with ethical standards beyond what has already been outlined in the initial sections of this dissertation.

Future Research Directions

Building upon the foundational work presented in this study, several avenues for future research can be conceptualized. Firstly, machine learning algorithms can be adapted to incorporate more diversified data types, such as standardized developmental and educational assessments. By integrating multiple data sources, one can create a multi-dimensional model that presents a more complete understanding of developmental disorders and their associated comorbidities.

Secondly, the current ML approaches can be specialized to examine age-specific markers and symptoms. Developmental disorders and their comorbidities are not static but evolve over time. Age-specific algorithms can help in identifying how comorbidities develop and change, offering insights for tailored interventions that adapt to an individual's developmental trajectory.

Incorporating real-world settings, such as school or home environments, in the future ML models would also allow for a more ecologically valid understanding of how developmental disorders interact with everyday life and other potential comorbid conditions.

As with the other work done within the Ministry of Health, this manuscript is in the process of submission to academic journals and awaits final confirmations from all contributing authors.

4.0 How Developmental Disorders Changed Before and After the Pandemic

4.1 Abstract

The COVID-19 pandemic has had a lasting impact on mental health, with lingering effects on the treated prevalence of developmental disorders, such as autism spectrum disorder, attention-deficit/hyperactivity disorder, and intellectual disabilities. This study explores changes in developmental disorder utilization in the Alberta healthcare system before and after the pandemic using administrative healthcare databases from the Ministry of Health in Alberta. Results indicate an overall increase in the treated prevalence of developmental disorders from 2018 to 2022. The ongoing impact of the pandemic on developmental disorders highlights the need for better surveillance, mental health support, and informed policy decisions to ensure individuals with developmental disorders and their families receive the necessary support and resources.

4.2 Introduction

The coronavirus disease 2019 (COVID-19) pandemic has had a significant impact on people's physical and mental health (Mahmud et al., 2022; Xiong et al., 2020), affecting people of all ages (Cielo et al., 2021; Webb et al., 2022), but young individuals seem to be the most affected (Huang et al., 2020). While much of the initial impact has subsided, lingering mental health issues remain (Rahmati et al., 2023; Xu et al., 2022), especially for individuals with developmental disorders (Breaux et al., 2021). A concerning development is the increased utilization of the Alberta healthcare system for developmental disorders (Clark et al., 2023), such as autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), and intellectual disabilities. This research is part of a collaborative effort under an interchange agreement, focused on addressing health data priorities and providing recommendations to the Ministry in Alberta.

Developmental disorders are generally more chronic in nature and less likely to resolve over time (Murphy, et al., 2005). The impact on a younger population in need of specialized support such as school aides and therapists will have ongoing impacts as these individuals age (Boulet et al., 2009). This trend raises concerns about the potential for increased costs to the healthcare system in the future and adds urgency to addressing this issue.

The sudden shift to remote learning and increased social isolation during COVID-19 may have been an aggravating factor for individuals with developmental disorders and their families (Loades et al., 2020). This potentially exacerbated their symptoms and made it more difficult for their long-term success. Previous studies have shown the significant impact of the pandemic on mental health, leading to increased rates of various disorders such as depression, anxiety, and stress (Loades et al., 2020). However, to our knowledge, little research has been done to examine the impact of the pandemic on a large-scale review of utilization rates for individuals with developmental disorders over multiple years.

In this paper, we aim to explore the changes in developmental disorder utilization before and after the COVID-19 pandemic. To our knowledge, our research is the first to shed light on this issue from a health utilization perspective for such a long period. Our hope is that this will increase awareness for these more chronic health issues and the need for better surveillance and mental health support for individuals with developmental disorders, especially in the context of the ongoing challenges this group will face and the added burden of costs. Understanding the changes in developmental disorders before and after the pandemic and analyzing the factors contributing to this trend is valuable. The loss of critical resources has already made it more difficult for students with developmental disorders to succeed in their academic and personal lives. Research like ours has the potential to inform future policy decisions, improve surveillance, and access to support and treatment for individuals with developmental disorders in the future.

4.3 Methods

4.3.1 Data Source

This study used the administrative healthcare databases of practitioner claims (CLM), emergency department visits (ED), and inpatient services (INP) in the Ministry of Health in Alberta.

4.3.2 Cohort

The study cohort included all residents of Alberta with active health insurance coverage for public healthcare services. This cohort was dynamic in nature, as individuals may enter or exit the cohort annually.

4.3.3 Outcome

Mental health-related diagnoses were selected based on the International Classification of Diseases, 9th or 10th revision (ICD-9 and ICD-10), available on the records from physician's office visits, emergency department visits, and inpatient services. Mental health services were identified from records with a primary diagnosis code of the following mental disorders: ADHD, ASD, intellectual disability, communication disorder, special learning disorder, coordination disorder, tics, and conduct disorder.

4.3.4 Data processing

We extracted the number of unique patients and events related to mental health concerns between 2016 and 2022. To normalize the data, for each year we scaled the service utilization numbers by the healthcare utilization size, i.e., the number of subjects with active Alberta Health care insurance and access to the healthcare system. The result is the number of unique patients and events related to mental health per 1000 patients for each year.

To understand the impact of the COVID-19 pandemic on health services utilization, we estimated the expected utilization numbers and the associated standard deviations for 2020, 2021, and 2022 assuming a linear increase in utilization from 2016 to 2019. We also explored the temporal trend of utilization by scaling the observed numbers with population size for each year as a comparison.

For the monthly temporal trend, we extracted the number of unique patients and events with mental health-related services for each month from January 2016 to December 2022. Similar to the yearly data, these numbers were scaled by the healthcare utilization size and by the population size in the corresponding month and year.

As of the date of this study (January 2023), it is important to note that the data for certain months in 2022 may not be entirely complete or representative. This is due to the fact that the most recent months' data are more susceptible to modifications, as they are still subject to updates, corrections, and additional information being added. Consequently, the findings from this study should be interpreted with caution, particularly when examining trends and patterns for the latter part of 2022.

4.4 Results

The data collected in this study consists of the treated prevalence of developmental disorders per 1000 patients in four groups (All, CLM, ED, and INP) from 2016 to 2022. The data was analyzed to identify various trends in the data, including quarterly, seasonal, and annual trends.

Table 4.1. Cohort characteristics

Cohort	**Sample Size**	**Mean age**	**Female (%)**
Dev	387,337	26.2	45
GP	4,960,783	39.6	50

Table 4.1. The cohort size reported in this table is derived from the sum of unique records in the system spanning from 2016 to 2022. "Dev" represents the subjects with recorded utilization of developmental disorder information during this period. "GP" refers to the general population, encompassing all other individuals without developmental disorders.

Figure 4.1 illustrates the observed results in comparison to an expected linear growth pattern based on the changes observed between 2016 and 2019. While the expected growth pattern follows a consistent upward trajectory, the actual observed results deviate from this trend, particularly when considering looking at all sources combined. This is contrasted with the sharp dip in ED visits in 2020 when compared to previous ED trends and observations from the other data.

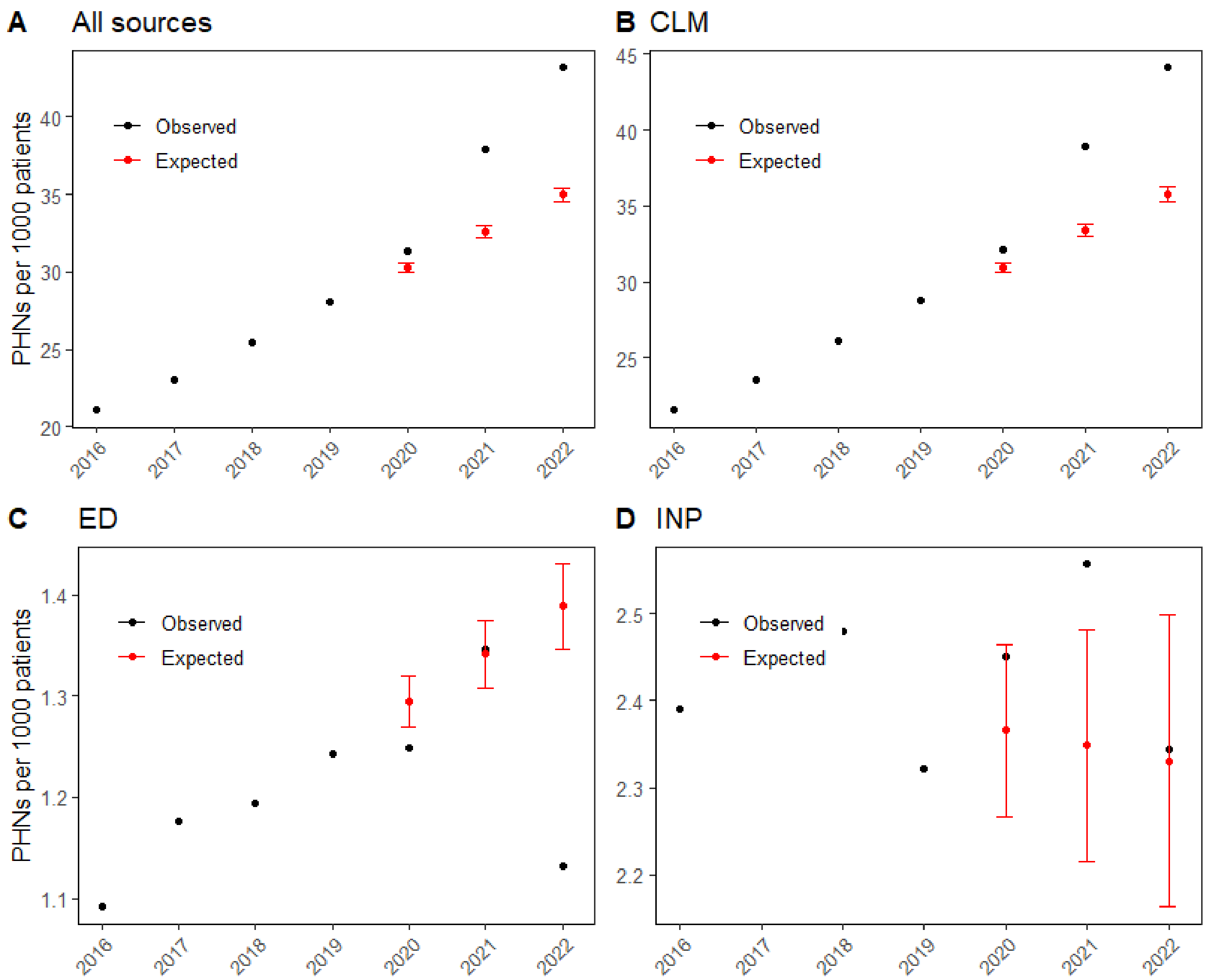

Figure 4.1. Trend of the observed and expected number of patients with development disorders scaled for 1000 patients along different years. 'All sources' is the combination of physician's office visits (CLM), emergency department (ED) visits, and inpatient (INP) data.

An increasing trend in the treated prevalence of developmental disorders was observed for the combination of all sources and CLM, while a decrease in ED visits and INP was seen over the five-year period (Figure 4.2). The overall treated prevalence in the 'All' group increased from 23.3 patients per 1000 patients in the first quarter (Q1) of 2018 to 41.6 in 2022 quarter 3 (Q3), indicating an overall increase of 78.5%. The CLM group experienced a similar increase, from 24.1 patients per 1000 patients in 2018-Q1 to 43.5 in 2022-Q3, marking a 80.5% increase. The ED group's treated prevalence rose from 1.06 patients per 1000 patients in 2018-Q1 to 0.84 in 2022-Q3, reflecting a decrease of 20.8%. The INP group saw an increase from 2.73 patients per 1000 patients in 2018-Q1 to 2.03 in 2022-Q3, representing a 25.6% decrease (Figure 4.2).

There were fluctuations in the treated prevalence of developmental disorders within the groups when comparing quarters. In general, the 'All' and CLM groups experienced an increase in treated prevalence in the second and fourth quarters of each year (Figure 4.2). The ED group showed variability across quarters, with a notable decrease in treated prevalence in 2022-Q3 (0.84) compared to 2018-Q1 (1.06). The INP group exhibited a general decrease in treated prevalence over time, with the lowest rate in 2022-Q3 (2.03).

When it comes to trends on a seasonal basis, the ‘All’ and CLM groups displayed a higher treated prevalence of developmental disorders in the summer (second quarter) and winter (fourth quarter) seasons (Figure 4.2). This pattern does not seem to be influenced by the COVID-19 pandemic (Figure 4.2). In contrast, the ED group showed no clear seasonal pattern, while the INP group demonstrated a decline in treated prevalence throughout the study period, with the lowest treated prevalence in the fall (third quarter) of 2022 (Figure 4.2).

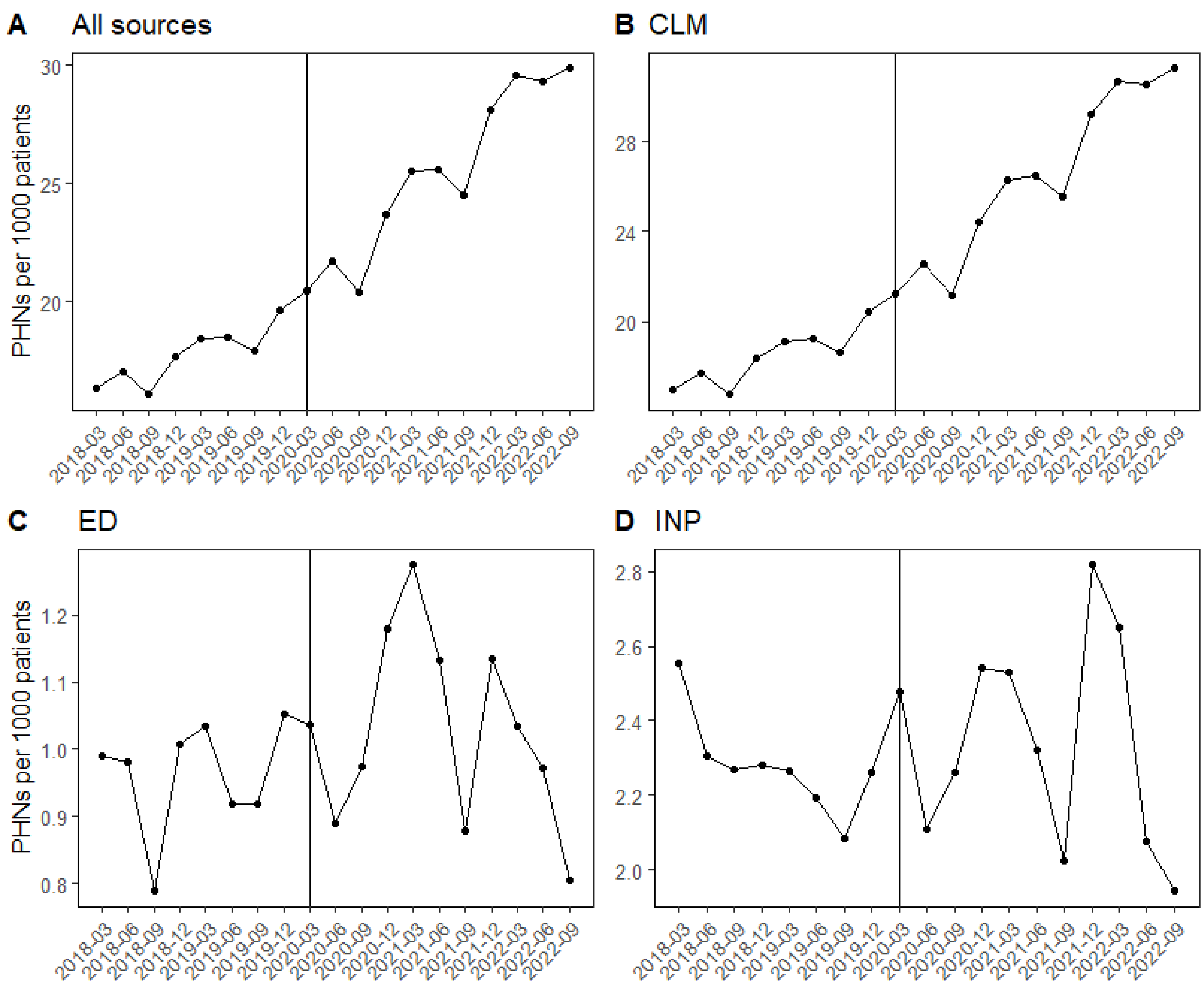

Figure 4.2: The number of patients with development disorders scaled for 1000 patients along different quarters. 'All sources' is the combination of physician's office visits (CLM), emergency department (ED) visits, and inpatient (INP) data.

Other trends of interest include a steep annual increase in the treated prevalence of developmental disorders for the 'All' group between 2020-Q4 (36.9) and 2021-Q1 (40.5), with an increase of 3.6. Similarly, the CLM group experienced its steepest annual increase between 2020-Q4 (37.9) and 2021-Q1 (41.6), with an increase of 3.7; The largest quarterly change in the 'All' group occurred between 2020-Q2 (33.9) and 2020-Q3 (30.8), with a decrease of 3.1. The CLM group experienced its largest quarterly change between 2020-Q2 (35.1) and 2020-Q3 (31.9), with a decrease of 3.2. The ED group's largest quarterly change took place between 2020-Q4 (1.29) and 2021-Q1 (1.37), with an increase of 0.08. For the

INP group, the largest quarterly change was between 2021-Q4 (3.06) and 2022-Q1 (2.74), with a decrease of 0.32. Conversely, there were also periods of stability in the treated prevalence of developmental disorders in the ED group between 2018-Q1 (1.06) and 2018-Q2 (1.06), and 2020-Q2 (1) and 2020-Q3 (1.05), with only slight changes during these periods.

Our results indicate an overall increase in the treated prevalence of developmental disorders in the 'All' and CLM groups from 2018 to 2022. The ED group displayed a decrease in treated prevalence, while the INP group demonstrated a decline over time. Seasonal trends were evident in the 'All' and CLM groups, with higher treated prevalence in summer and winter seasons.

4.5 Discussion

The results of our study provide evidence that the events surrounding the COVID-19 pandemic has had a profound and sustained impact on utilization rates in the Alberta healthcare system for developmental disorders. Developmental disorders, such as ASD and ADHD, predominantly affect children and adolescents, impacting their cognitive, emotional, and social development. The escalation in healthcare utilization for these disorders during the pandemic suggests an increase in symptom severity possibly due to disrupted care routines and reduced access to therapeutic and educational supports. This period of isolation and change could exacerbate existing conditions, emphasizing the need for healthcare systems to enhance their capacity to deliver services remotely and flexibly during crises. While previous research has shown the pandemic has impacted mental health (Penninx et al., 2022), developmental disorders rates remain high relative to other disorders that have reduced to a utilization rate closer to their pre-COVID norms. Given the chronic nature of these conditions and the young age of the population affected, our team believes it is important to explore this further. This underlines the importance of targeted policy interventions and the adaptation of service delivery models to ensure continuous care for vulnerable populations, especially during global health emergencies.

Several factors may have contributed to this increase in developmental disorders. The sudden shift to remote learning and increased social isolation significantly disrupted the daily routines of everyone (Sibley et al., 2021; Zhang et al., 2022); this included many of the support systems individuals and their families relied on for those with special needs (Currie et al., 2022). Additionally, the reduced mobility of auxiliary support staff in schools made it more difficult for these individuals to access the support they needed to succeed academically and personally during the COVID-19 lockdowns in Alberta.

In light of our study's findings, we believe that there are potential implications for the healthcare system and policy decisions related to developmental disorders. The observed increase in the treated prevalence of these conditions suggests an enhanced surveillance approach for individuals with developmental disorders, as well as potential increased access to treatment and support services.

Our research findings emphasize several policy recommendations for the province of Alberta, which aim to address the unique challenges faced by individuals with developmental disorders during and after the COVID-19 pandemic. We suggest that enhancing the accessibility and availability of virtual support services is of value in providing continuity of care during times of social isolation or remote learning, ensuring consistent support and treatment for this vulnerable population (Lakes et al., 2022). Moreover, implementing a comprehensive reintegration plan could facilitate a transition back into in-person learning environments and community-based support services (Preyde et al., 2017). Such a plan could contribute to the continued academic and personal growth of individuals with developmental disorders by addressing their specific needs adequately. This also could help mitigate the longer term costs associated with support for those that don’t successfully reintegrate into the community. It is important to acknowledge that when the burden of support primarily falls on the families themselves, it often detracts from the province's available pool of labour and talent.

Finally, we recommend the allocation of additional funding to support research and the development of targeted interventions (Feldman et al., 2022). This investment could foster a deeper understanding of the factors contributing to the observed trends and enable the mitigation of the long-term impacts of the pandemic on this at-risk population. Further research is important to comprehend the factors contributing to these trends and to develop targeted interventions addressing developmental disorders within these specific groups. By adopting any of these policy recommendations, we hope to better support individuals with developmental disorders and their families in the province of Alberta.

4.6 Limitations

While our study provides valuable insights into the impact of the COVID-19 pandemic on developmental disorder utilization rates in Alberta, it is important to acknowledge some limitations. Firstly, our study is limited to Alberta's healthcare system and may not be generalizable to other regions or countries. Secondly, while we used administrative healthcare data to identify mental health-related diagnoses, our study is reliant on the accuracy and completeness of the coding of these diagnoses in the administrative databases. Thirdly, our study does not account for potential confounding factors, such as changes in diagnostic criteria or changes in access to healthcare services. Fourthly, the data used in this study only goes up to 2022, and the impact of the pandemic may continue to affect developmental disorder utilization rates in the future. Lastly, while our study highlights the increased utilization rates of developmental disorders, it is not able to provide a detailed understanding of the factors contributing to the observed trends. Further research is necessary to understand the underlying causes of the increased utilization rates and to identify appropriate interventions to address these trends.

4.7 Conclusion

Overall, this study provides important insights into the impact of the COVID-19 pandemic on individuals with developmental disorders and highlights the need for continued research and action to address this issue. By improving our understanding of the impact of the pandemic on developmental disorders, we can work to ensure that individuals with these conditions receive the support and treatment they need to lead fulfilling and productive lives.

4.8 From Developmental Disorders to Neurocognitive Disorders Amidst the Pandemic

As we delved into the utilization rates for developmental disorders within Alberta's healthcare system, our focus shifts towards another critical area exacerbated by the COVID-19 pandemic: neurocognitive disorders (NCDs). The next chapter's transition underscores the pandemic's broad impact on mental health, extending beyond developmental disorders to include significant increases in NCDs, such as Alzheimer's disease and dementia. Both sections collectively emphasize the importance of a multidimensional approach to healthcare, advocating for enhanced surveillance, support systems, and policy adjustments to address the complex needs of individuals affected by developmental and neurocognitive disorders. By examining the rise in NCDs alongside developmental disorders, we highlight the necessity for comprehensive public health strategies that can navigate the intricacies of healthcare challenges during crisis. This progression from developmental to neurocognitive disorders enriches our understanding of how integrated solutions using ML and AHRs can help guide resources and policy to areas in need, which are not always the area of focus.

4.9 Supplement material

Table S4.1. ICD-9 and ICD-10 code diagnosis codes of mental disorders.

Condition Name	Sub-Condition Name	ICD9 Codes	ICD10 Codes
Developmental Disorders	All	Include: '299','307','312','313','314','315','317','318','319' Exclude: '307.1','307.4','307.5','307.6','307.7','307.8'	Include: 'F70','F71','F72','F73','F78','F79','F80','F81','F82','F83','F84','F88','F89','F90','F91','F92','F94','F95','F98' Exclude: 'F980','F981','F982','F983'
	ADHD	Include: '314'	Include: 'F90'
	Autism	Include: '299.0','299.1'	Include: 'F84'
	Intellectual Disab.	Include: '317','318','319'	Include: 'F70','F71','F72','F73','F78','F79'
	Communication Disorder	Include: '307.0','307.9','315.3','313.23','315.81'	Include: 'F80','F940','F985'
	Special Learning Disorder	Include: '315.0','315.1','315.2','315.9'	Include: 'F81'
	Coordination Disorder	Include: '307.3','315.4','315.5'	Include: 'F82','F83','F984'
	Tics	Include: '307.2'	Include: 'F95'
	Conduct Disorder	Include: '312'	Include: 'F91','F92'

Table S4.2. Number of patients with developmental disorders per 1000 people.

Developmental Disorders n per 1000 heatmap

Year	Quarter	All	CLM	ED	INP
2018	1	23.3	24.1	1.06	2.73
2018	2	24.6	25.5	1.06	2.6
2018	3	22.1	23	0.859	2.49

2018	4	25.4	26.4	1.09	2.52
2019	1	26.3	27.2	1.11	2.47
2019	2	26.7	27.7	0.987	2.32
2019	3	24.7	25.6	0.989	2.21
2019	4	28.1	29.2	1.13	2.41
2020	1	29.4	30.4	1.1	2.77
2020	2	33.9	35.1	1	2.3
2020	3	30.8	31.9	1.05	2.47
2020	4	36.9	37.9	1.29	2.78
2021	1	40.5	41.6	1.37	2.8
2021	2	39.9	41.2	1.24	2.56
2021	3	36.1	37.6	0.947	2.18
2021	4	42.7	44.3	1.21	3.06
2022	1	44.9	46.6	1.09	2.74
2022	2	44.1	46	1.03	2.31
2022	3	41.6	43.5	0.84	2.03

Note: Red represents a larger increase in the number of patients with developmental disorders per 1000 people, while green indicates a lower rate per 1000 relative to the other rates in the data. CLM = Physician's office claims; ED = emergency department visits; INP = inpatients.

References

Boulet, S. L., Boyle, C. A., & Schieve, L. A. (2009). Health care use and health and functional impact of developmental disabilities among US children, 1997-2005. *Archives of Pediatrics & Adolescent Medicine*, *163*(1), 19–26. https://doi.org/10.1001/ARCHPEDIATRICS.2008.506

Breaux, R., Dvorsky, M. R., Marsh, N. P., Green, C. D., Cash, A. R., Shroff, D. M., Buchen, N., Langberg, J. M., & Becker, S. P. (2021). Prospective impact of COVID-19 on mental health functioning in adolescents with and without ADHD: protective role of emotion regulation abilities. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *62*(9), 1132–1139. https://doi.org/10.1111/JCPP.13382

Cielo, F., Ulberg, R., Environmental, D. D. G.-I. J. of, & 2021, undefined. (2021). Psychological impact of the COVID-19 outbreak on mental health outcomes among youth: A rapid narrative review. *Mdpi.ComF Cielo, R Ulberg, D Di GiacomoInternational Journal of Environmental Research and Public Health, 2021•mdpi.Com*, *18*. https://doi.org/10.3390/ijerph18116067

Clark, C. A., Turner, K., Kuntz, J., Perri, A., Deegan, A., Marriott, B., Graham, S., Rahman, A., & McMorris, C. A. (2023). COVID-19 and Neurodevelopmental Disabilities: Examining the Impact of the First 2 Years of the Pandemic on the Demand for Pediatric Inpatient Care. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/S10803-023-06136-X

Currie, G., Finlay, B., Seth, A., Roth, C., Elsabbagh, M., Hudon, A., Hunt, M., Jodoin, S., Lach, L., Lencucha, R., Nicholas, D. B., Shakako, K., & Zwicker, J. (2022). Mental health challenges during COVID-19: perspectives from parents with children with neurodevelopmental disabilities. *Taylor & FrancisG Currie, B Finlay, A Seth, C Roth, M Elsabbagh, A Hudon, M Hunt, S Jodoin, L LachInternational Journal of Qualitative Studies on Health and Well-Being, 2022•Taylor & Francis*, *17*(1). https://doi.org/10.1080/17482631.2022.2136090

Feldman, M. A., Azzano, A., Vause, T. C., & Ward, R. A. (2022). *Early Intervention for Children At-Risk for Autism Spectrum Disorder*. 727–757. https://doi.org/10.1007/978-3-030-88538-0_31

Huang, Y., research, N. Z.-P., & 2020, undefined. (2020). Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0165178120306077

Lakes, K. D., Cibrian, F. L., Schuck, S. E. B., Nelson, M., & Hayes, G. R. (2022). Digital health interventions for youth with ADHD: A mapping review. *Computers in Human Behavior Reports*, *6*, 100174. https://doi.org/10.1016/J.CHBR.2022.100174

Loades, M. E., Chatburn, E., Higson-Sweeney, N., Reynolds, S., Shafran, R., Brigden, A., Linney, C., McManus, M. N., Borwick, C., & Crawley, E. (2020). Rapid Systematic Review: The Impact of Social Isolation and Loneliness on the Mental Health of Children and Adolescents in the Context of COVID-19. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(11), 1218-1239.e3. https://doi.org/10.1016/J.JAAC.2020.05.009

Mahmud, S., Mohsin, M., Dewan, M., Psychology, A. M.-T. in, & 2023, undefined. (2022). The global prevalence of depression, anxiety, stress, and insomnia among general population during COVID-

19 pandemic: A systematic review and meta. *Springer*. https://link.springer.com/article/10.1007/s43076-021-00116-9

Penninx, B. W. J. H., Benros, M. E., Klein, R. S., & Vinkers, C. H. (2022). How COVID-19 shaped mental health: from infection to pandemic effects. *Nature Medicine*, *28*(10), 2027–2037. https://doi.org/10.1038/S41591-022-02028-2

Preyde, M., Parekh, S., Warne, A., & Heintzman, J. (2017). School Reintegration and Perceived Needs: The Perspectives of Child and Adolescent Patients During Psychiatric Hospitalization. *Child and Adolescent Social Work Journal*, *34*(6), 517–526. https://doi.org/10.1007/S10560-017-0490-8/TABLES/3

Rahmati, M., Udeh, R., Keon Yon, D., Won Lee, S., Dolja-Gore, X., McEVoy, M., Kenna, T., Jacob, L., López Sánchez, G. F., Koyanagi, A., Il Shin, J., Smith, L., Unit, D., & Sanitari Sant Joan de Déu, P. (2023). A systematic review and meta-analysis of long-term sequelae of COVID-19 2-year after SARS-CoV-2 infection: A call to action for neurological, physical, and. *Wiley Online LibraryM Rahmati, R Udeh, DK Yon, SW Lee, X Dolja-Gore, M McEVoy, T Kenna, L JacobJournal of Medical Virology, 2023•Wiley Online Library*, *95*(6). https://doi.org/10.1002/jmv.28852

Sibley, M., Ortiz, M., Gaias, L., Reyes, R., … M. J.-J. of psychiatric, & 2021, undefined. (2021). Top problems of adolescents and young adults with ADHD during the COVID-19 pandemic. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0022395621000765

Webb, L., Psychiatry, C. C.-I. J. of G., & 2022, undefined. (2022). The COVID-19 pandemic's impact on older adults' mental health: Contributing factors, coping strategies, and opportunities for improvement. *Ncbi.Nlm.Nih.GovLM Webb, CY ChenInternational Journal of Geriatric Psychiatry, 2022•ncbi.Nlm.Nih.Gov*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8646312/

Xiong, J., Lipsitz, O., Nasri, F., Lui, L., … H. G.-J. of affective, & 2020, undefined. (2020). Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0165032720325891

Xu, E., Xie, Y., & Al-Aly, Z. (2022). Long-term neurologic outcomes of COVID-19. *Nature Medicine 2022 28:11*, *28*(11), 2406–2415. https://doi.org/10.1038/s41591-022-02001-z

Zhang, S., Hao, Y., Feng, Y., environmental, N. L.-I. journal of, & 2022, undefined. (2022). COVID-19 pandemic impacts on children with developmental disabilities: Service disruption, transition to telehealth, and child wellbeing. *Mdpi.ComS Zhang, Y Hao, Y Feng, NY LeeInternational Journal of Environmental Research and Public Health, 2022•mdpi.Com*. https://www.mdpi.com/1660-4601/19/6/3259

Chapter 5
Impact of the Pandemic on Neurocognitive Disorders

This section discusses the impact of the COVID-19 pandemic which has been pervasive, affecting various domains of public health, including mental health disorders. A noteworthy focus within the mental health spectrum is the increase in the treated prevalence of neurocognitive disorders (NCDs) like Alzheimer's disease and dementia, particularly among older populations. Despite the considerable attention that mental health has garnered during the pandemic, little research has delved into the specific effects of the pandemic on NCDs. This paper aims to help fill some of this gap by examining the treated prevalence and healthcare utilization trends related to NCDs before and after the pandemic lockdowns in the Alberta Health system. This research also hopes to show again the value of AHRs in this type of research.

Critique of the Paper

While the paper offers valuable insights into a relatively underexplored area, there are some limitations that must be acknowledged. Firstly, the study's scope is confined to the Alberta healthcare system, limiting the generalizability of the findings. Secondly, the paper relies heavily on administrative data, which may not accurately capture the full spectrum of NCD treated prevalence. The accuracy of diagnosis codes, incomplete documentation, and the possibility of misdiagnosis are factors that could affect the quality of the data. Thirdly, while the paper does a good job at highlighting the observed trends, it falls short of providing a causal explanation. The study is observational in nature and does not consider potential confounding factors such as changes in healthcare policies or population demographics during this period. Finally, the study does not explore how the pandemic's impact may vary across different subgroups, which is crucial for targeted intervention strategies.

Future Directions

Given these limitations, several avenues for future research are apparent. Expanding the study to include multiple healthcare systems would lend more weight to the findings. Prospective studies employing a longitudinal design could better elucidate causal relationships between the pandemic and NCDs. There is also a need for more granular analyses that take into account demographic variables such as age, gender, and socioeconomic status. Such analyses would be invaluable for public health agencies in tailoring their interventions. Research could also focus on the qualitative experiences of individuals with NCDs and their caregivers, to better understand the nuanced challenges faced during the pandemic.

As with the other work done within the Ministry of health this manuscript is in the process of submission to academic journals and awaits final confirmations from all contributing authors.

5.0 Neurocognitive Disorders Before and After the Pandemic

5.1 Abstract

The COVID-19 pandemic has had a significant impact on mental health globally, with a notable rise in neurocognitive disorders (NCDs) such as Alzheimer's disease and dementia. The pandemic has exacerbated modifiable risk factors and limited access to support and treatment services for individuals with NCDs. This study investigates the changes in the treated prevalence and healthcare utilization of the more chronic types of NCDs in Alberta, Canada, during the pandemic, utilizing administrative healthcare databases between 2016 and 2022. Our cohort included 106,489 unique patients, with a median age of 77 years. The results reveal a significant increase in NCD treated prevalence in 2020, followed by a sustained overall increase in NCD in the population. While fluctuations in treated prevalence varied across different care settings, the most significant variations were observed in practitioner claims data. This study demonstrates the need to address the long-term ramifications of increased care for NCDs as a result of the pandemic.

5.2 Introduction

Since the coronavirus disease 2019 (COVID-19) pandemic lock downs we have seen increased rates of various mental health disorders globally (Wang et al., 2022). Certain populations struggle with chronic and ongoing mental health issues (Daly et al., 2022). One particular area of concern is the rise in neurocognitive disorders (NCD) such as Alzheimer's disease and dementia (Golzari-Sorkheh et al., 2023; Wang et al., 2022; Xia et al., 2021).

NCDs are defined by the International Classification of Diseases-Tenth Revision (ICD-10) as an impairment in one or more cognitive domains (i.e., learning, memory, attention, executive functioning, perceptual motor functioning, and social cognition) that cannot be attributed to normal aging (WHO, 2019). In Canada, NCDs were the most commonly cited comorbidity with COVID-19 on death certificates of people aged 65 or older (Statistics Canada, 2023). These disorders are chronic in nature and have a profound impact on an individual's quality of life, as well as on the lives of their caregivers and families who are at a higher risk of experiencing sleep problems, depression, and anxiety (Dauphinot et al., 2016). NCDs also place a heavy burden on society as a whole. In 2015, the estimated global cost of dementia was 818 billion US dollars. As the population continues to age and the treated prevalence of dementia climbs, the estimated cost of dementia is projected to rise to two trillion US dollars by 2030 (Prince et al., 2015). A 2020 report to the Lancet Commission discussed 12 modifiable risk factors that if addressed could potentially prevent or delay the onset of 40% of late-onset dementia. The 12 modifiable risk factors are: depression, physical inactivity, low social contact, less education, hypertension, smoking, obesity, excessive alcohol consumption, traumatic brain injury, hearing impairment, diabetes, and air pollution (Livingston et al., 2020). The COVID-19 pandemic adds a new layer of complexity to the situation by proliferating the presence of modifiable risk factors (i.e., depression, physical inactivity, low social contact) and limiting access to support and treatment services (Hellis & Mukaetova-Ladinska, 2022). In Canada, lockdowns and isolation were most severe in long-term care homes, where many residents struggle with NCDs (Stall et al., 2020).

Despite the significant attention given to the impact of the COVID-19 pandemic on mental health, to our knowledge, no research has been done to examine the impact of the pandemic on NCDs. NCDs, which are chronic and mostly affect older people, could have been exacerbated by pandemic-related stress and interruptions in routine care (Mukaetova-Ladinska et al., 2021). Therefore, it is valuable to understand the changes in NCDs before and after the pandemic and to consider the factors contributing to this trend.

Moreover, the pandemic has also affected the delivery of care and support services for individuals with NCDs (Zorzi, 2021). The reduction of auxiliary support staff and limited access to in-person support and treatment services has made it more difficult for individuals with these disorders to access care (Giebel et al., 2021), which may contribute to increased utilization for these disorders. More investigation is needed to inform future policy decisions and to improve access to support and treatment for individuals with NCDs.

We aim to examine the changes in NCDs before and after the COVID-19 pandemic and discuss how this trend is an outlier to what was observed pre-COVID. The findings from our investigation underscore the importance of addressing NCDs in the context of the ongoing pandemic and signals the need for a more comprehensive discussion into the long-term ramifications of this increased utilization of the healthcare system.

5.3 Methods

5.3.1 Data Source

This study used the administrative healthcare databases of practitioner claims, emergency department visits, and inpatient services in the Ministry of Health in Alberta.

5.3.2 Cohort

The study cohort included all residents of Alberta with active health insurance coverage for public healthcare services. This cohort was dynamic in nature, as individuals may enter or exit the cohort annually.

5.3.3 Outcome

Neurocognitive-related diagnoses were selected based on either the International Classification of Diseases, 9th or 10th revision (ICD-9 and ICD-10), depending on the source. These were recorded from physician's office visits, emergency department visits, and inpatient services. Neurocognitive-related service was identified from records with a primary diagnosis code of the following neurocognitive and organic disorders: Dementia, Alzheimer, and Delirium.

5.3.4 Data processing

We extracted the number of unique patients and events related to neurocognitive disorders between 2016 and 2022. To ensure fairness in comparisons across years, we adjusted these numbers for each year's total healthcare numbers using population counts by those with active Alberta Health insurance. The results are presented as the yearly count of unique patients and events per 1000 insured individuals., i.e., the number of subjects with an active Alberta Health care insurance. The result is the number of unique patients and events related to neurocognitive disorders per 1000 patients for each year.

To understand the impact of the COVID-19 pandemic on the health services utilization, we estimated the expected utilization numbers and the associated standard deviations for 2020, 2021, and 2022 assuming a linear increase of the utilization from 2016 to 2019. We also explored the temporal trend of utilization by scaling the observed numbers with population size for each year as a comparison.

For the monthly temporal trend, we extracted the number of unique patients and events with neurocognitive-related services for each month from January 2016 to December 2022. Similar to the yearly data, these numbers were scaled by the healthcare utilization size and by the population size in the corresponding month and year.

As of the date of this study (January 2023), the data for certain months in 2022 may not be entirely complete or representative. Consequently, the findings from this study should be interpreted with caution, particularly when examining trends and patterns for the latter part of 2022.

5.4 Results

This study included 106,489 unique patients that utilized the Alberta Health system between 2016 and 2022. The median age of these patients was 77 years old at the time of this research (Table 1). A majority of these patients were male (58%).

Table 5.1. Cohort characteristics.

Cohort	Sample size	Mean age	Median age	Female (%)
GP	5,241,631	38	37	50
Neurocognitive	106,489	70	77	58

Table 5.1. The cohort size reported in this table is derived from the sum of unique records in the system spanning from 2016 to 2022. "Neurocognitive" represents the subjects with recorded utilization of Neurocognitive disorder information during this period. "GP" refers to the general population, encompassing all other individuals with an active healthcare number without Neurocognitive disorders.

As noted in Figure 5.1, below, between 2020 and 2022, there were notable changes in the utilization rates of events related to NCD per 1000 patients. Specifically, in claims data (CLM), the utilization rates spiked from 17 per 1000 patients in 2019 to nearly 24 per 1000 patients in 2020. While the rates came back down in 2021 and 2022 they remain above what was expected prior to the pandemic. This represents a significant decrease in the utilization rate of events related to NCDs between 2020 and 2022. Unexpectedly, utilization rates in the emergency department (ED) and inpatient care (INP) have fallen well below trends in the four years prior to the pandemic.

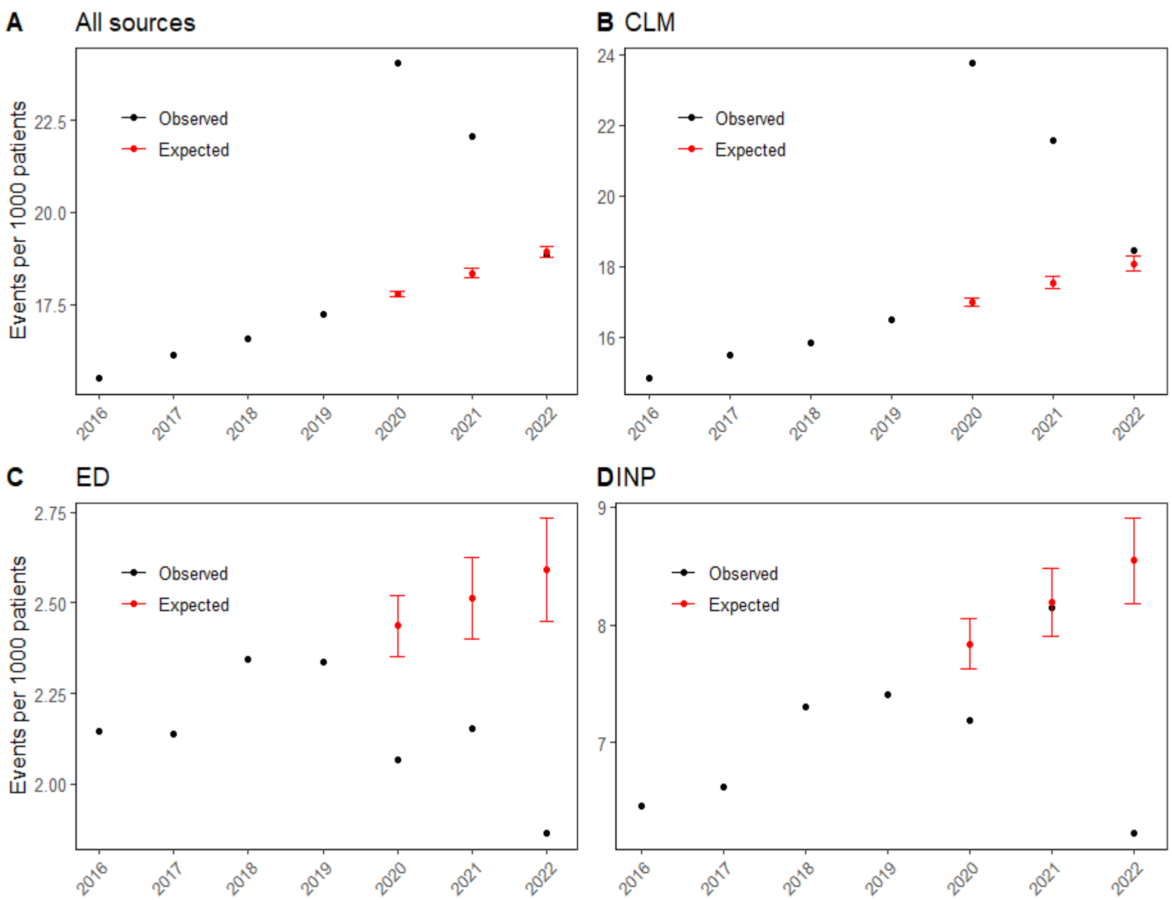

Figure 5.1: Trend of the observed and expected number of patients with neurocognitive disorders adjusted per 1000 patients over different years. 'All sources' is the combination of physician's office visits (CLM), emergency department (ED) visits, and inpatient (INP) data.

The data shows fluctuations in the treated prevalence of NCDs across the years, with a noticeable increase in 2020 and subsequent decline in 2021 and 2022. The results reveal variations across quarters and groups (CLM, ED, and INP), with some groups experiencing higher treated prevalence rates during specific periods (Figure 5.2). The substantial increase in 2020, for example, suggests that there could be external influences or events that impact the trend. Specifically, there was a substantial increase in 2020 with the most notable pattern in the "all" and "CLM" groups in 2020, particularly in the second quarter (Figure 5.2). However, there is a sustained recovery after 2020 where the treated prevalence of neurocognitive disorders in the "all" and "CLM" groups started to decrease and continued to decrease slightly in 2022. This pattern indicates a recovery from the peak observed in 2020. The exception to this trend seems to be the ED group as there was no strong pattern in this group, the treated prevalence of NCDs remained relatively stable and low throughout the years. The values fluctuate within a narrow range (1.33 to 1.94), suggesting minimal variation in the treated prevalence of NCDs in this group. Whereas the "INP" group had a noticeable drop in treated prevalence in the third quarter of 2022, reaching the lowest value (4.86) in the entire dataset.

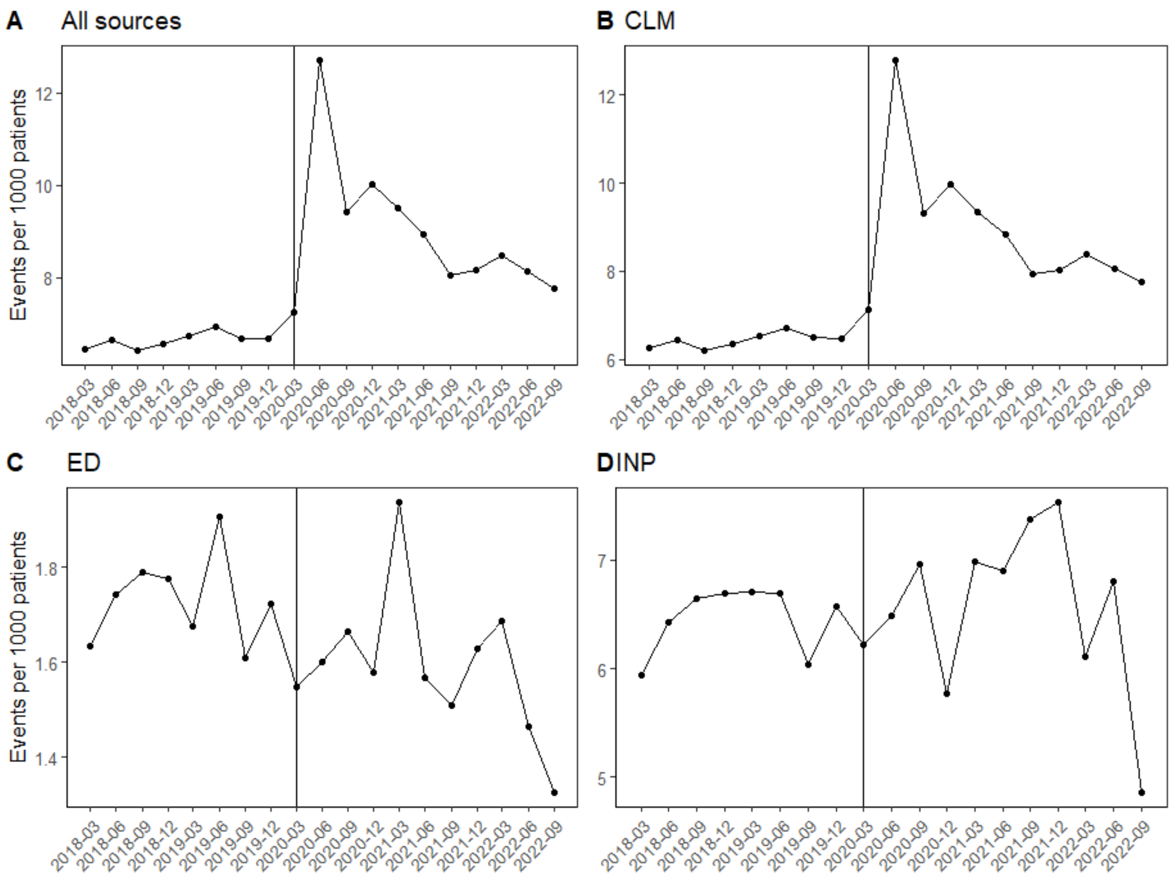

Figure 5.2: The number of patients with neurocognitive disorders scaled for 1000 patients along different quarters. 'All sources' is the combination of physician's office visits (CLM), emergency department (ED) visits, and inpatient (INP) data.

5.5 Discussion

In our analysis of 106,489 patients within the Alberta Health system from 2016 to 2022, we observed significant fluctuations in the treated prevalence of NCDs, particularly during the pandemic years, with a notable and sustained increase in NCDs since the pandemic onset. The rise in NCDs amid the COVID-19 pandemic represents a chronic and growing issue for healthcare systems.

One possible explanation is that older members of the population are accessing the healthcare system at a higher rate for COVID-related reasons, which in turn presents more opportunities for NCDs to be discovered in patients. Another potential reason is that severe disease in older patients visiting emergency departments is often associated with delirium, which may be confused with NCD or increase the risk for NCD (Hellis & Mukaetova-Ladinska, 2022). The pandemic has led to a reduction of protective factors, such as social activities that are mentally stimulating and promote movement. This has been accompanied by an increase in risk factors, including isolation, loneliness, less stimulating activities, and more immobility. Additionally, COVID-19 related risk factors may also play a role in the increased visibility of NCD during the pandemic. These factors include increased depression, as well as potential central nervous system (CNS) damage resulting from hypoxia, blood circulation issues, and disruptions to neuronal pathways (Hellis & Mukaetova-Ladinska, 2022).

Another potential reason for the decreased utilization of the emergency and inpatient care services in the province is changes in patient behaviors after the pandemic. Specifically, fears of infection from COVID-19 may have led to more people seeking care closer to home for themselves or family members, even for non-COVID related issues. This could have contributed to an increase in the number of NCD cases detected during the pandemic in practitioner claims but not in hospitals.

Given these observations, we propose the following policy recommendations for the province to address this growing concern. Enhance early detection and intervention programs (Krolak-Salmon et al., 2019). This can be achieved by training healthcare professionals to identify early signs of NCDs and implementing standardized screening tools in primary care settings. Promote mental health and social engagement, as the pandemic has reduced protective factors and increased risk factors for NCDs (Chyu et al., 2022). By encouraging mental health support services and facilitating social activities, especially for older adults, the province can help mitigate the negative effects of isolation and loneliness exacerbated during the pandemic.

Overall, the COVID-19 pandemic has highlighted the importance of surveillance but also the potential impact of infectious disease outbreaks on neurological health. More research is needed to fully understand the relationship between COVID-19 and NCD, as well as to support strategies for mitigating the negative effects of the pandemic on neurological health. Our analysis demonstrates that AHRs could be valuable in identifying the prevalence of neurocognitive disorders during significant events. The data shows a significant increase in conditions like Alzheimer's disease and dementia, especially among older adults as well as a disruption to the overall patterns for utilization. The integration of AHRs into our study has allowed for a detailed understanding of these trends, emphasizing the need for enhanced virtual care options and community-based support systems. These data-driven insights support existing recommendations and advocate for the development of targeted interventions. For instance, the increased use of telemedicine, and remote monitoring through wearables could mitigate the effects of reduced in-person care. Further, our findings highlight the necessity for policy adjustments to ensure better resource allocation and support for mental health services during and after the pandemic. This approach not only addresses the immediate challenges but also lays the groundwork for resilient healthcare systems capable of managing future crises effectively.

5.6 Limitation

While this study provides valuable insights into the treated prevalence of NCDs in Alberta, there are several limitations that should be considered when interpreting the results. First, the study only includes data from the Alberta healthcare system, which may not be generalizable to other healthcare systems or populations. Second, the study relies on administrative data from physician's offices, emergency department visits, and inpatient services. This data may not accurately capture the full extent of NCDs in the population, as some individuals may not seek medical attention for their symptoms. Additionally, the accuracy of the diagnosis codes used to identify NCDs may be limited by factors such as misdiagnosis or incomplete documentation. Third, the study only examines changes in NCDs before and after the COVID-19 pandemic, without considering other potential confounding factors such as changes in population demographics or healthcare policies. Fourth, the study does not explore potential differences in the impact of the pandemic on different subgroups of the population, such as those based on age, gender, or socioeconomic status. Finally, the study does not provide a causal explanation for the observed changes in NCDs during the pandemic, as it is based on observational data. Future studies may be needed to establish causal relationships between the pandemic and NCDs, as well as to identify potential interventions to mitigate the negative effects of the pandemic on neurological health.

5.7 Conclusion

Our study highlights the significant impact of the COVID-19 pandemic on NCDs, particularly in the utilization rates of events related to NCDs per 1000 patients. While fluctuations in the treated prevalence of NCDs were observed in the years prior to the pandemic, a sustained increase in NCDs was observed since the pandemic onset. The pandemic appears to have led to a reduction of protective factors and an increase in risk factors, which may contribute to the increased visibility of NCD during the pandemic. Furthermore, changes in patient behaviors may have contributed to an increase in the number of NCD cases detected during the pandemic. We believe these findings to be significant and emphasize the need for continued surveillance and research on the impact of COVID-19 on neurological health, as well as the development of strategies to mitigate the negative effects of the pandemic on individuals with NCD and their families. The long-term ramifications of this increased utilization of the healthcare system for NCDs should be further investigated to inform future policy decisions and improve access to support and treatment for individuals with NCDs in the future.

To fully contextualize our recommendations, it would be valuable to map out how these suggestions are supported by other datasets in the ministry. The AHRs have highlighted changes in patterns in accessing healthcare resources due to the rise in neurocognitive disorders, which should lead to a re-evaluation of current healthcare policies and support systems. Moreover, our findings advocate for increased funding in mental health research, focusing on the long-term effects of pandemic-induced stress and isolation on neurocognitive health. By integrating these data-driven insights, we can advance existing recommendations and develop more robust, evidence-based strategies to address the challenges posed by the pandemic on mental health. This approach not only supports immediate awareness but also lends evidence to support the healthcare system in future public health emergencies.

5.8 From Neurocognitive Disorders to Anticipatory Mental Health Care

Transitioning from the analysis of the NCD surge during the pandemic, we pivot to the promising role of AHRs in enhancing mental health preparedness for disasters. This pivot underscores the value in a forward-thinking approach in healthcare, moving from identifying challenges to implementing predictive solutions. Both discussions stress the importance of leveraging data to inform and improve mental health interventions in times of crisis. By connecting the increase in neurocognitive disorders with the potential of AHRs for early mental health crisis intervention, we highlight the evolution towards a proactive healthcare model. This approach seeks not only to address immediate health concerns but also to anticipate and mitigate future mental health impacts resulting from disasters with data-informed healthcare solutions.

5.9 Supplement Material

Table S5.1. ICD-9 and ICD-10 codes of the neurocognitive disorders.

Condition Name	Sub-Condition Name	ICD9 Codes	ICD10 Codes
NeuroCognitive and Organic Disorders (NCO)	All	Include: '290','293','294','331','341'	Include: 'F00','F01','F02','F03','F04','F05','F06','F07','F09','G30'
	Dementia	Include: '290','294.1','331.1','331.2'	Include: 'F00','F01','F02','F03','G30'
	Alzheimers	Include: '290.0','290.1','290.2','290.3','331.0'	Include: 'F00','G30'
	Organic	Include: '331','341','293.8','293.9','294.0','294.8','294.9'	Include: 'F06','F09','F070','F078','F079'

Table S5.2. Number of patients with neurocognitive disorders scaled for 1000 patients for all sources between 2018 and 2022.

Healthcare Settings Comparison: Quarterly Neurocognitive Disorders per 1000

Year	Quarter	All	CLM	ED	INP
2018	1	6.44	6.25	1.63	5.95
2018	2	6.66	6.45	1.74	6.43
2018	3	6.42	6.2	1.79	6.64
2018	4	6.56	6.34	1.78	6.7
2019	1	6.73	6.52	1.68	6.71
2019	2	6.94	6.71	1.91	6.69
2019	3	6.66	6.5	1.61	6.04

2019	4	6.67	6.47	1.72	6.57
2020	1	7.25	7.12	1.55	6.23
2020	2	12.7	12.8	1.6	6.49
2020	3	9.41	9.32	1.66	6.97
2020	4	10	9.97	1.58	5.77
2021	1	9.52	9.35	1.94	6.99
2021	2	8.94	8.83	1.57	6.9
2021	3	8.05	7.95	1.51	7.37
2021	4	8.16	8.03	1.63	7.54
2022	1	8.48	8.38	1.69	6.11
2022	2	8.12	8.05	1.46	6.8
2022	3	7.75	7.76	1.33	4.86

Note: Red represents a larger increase in the number of patients with neurocognitive disorders per 1000 people, while green indicates a lower rate per 1000 relative to the other rates in the data. CLM = practitioner claims; ED = emergency department visits; INP = inpatients.

References

Canadian Institute for Health Information. (2023). *Canadian Institute for Health Information*. Retrieved from Codes and Classifications: https://www.cihi.ca/en/submit-data-and-view-standards/codes-and-classifications

Cartus, A. R., Samuels, E. A., Cerdá, M., & Marshall, B. D. (2023). Outcome class imbalance and rare events: An underappreciated complication for overdose risk prediction modeling. *Addiction* , https://doi.org/10.1111/add.16133.

Dauphinot, V., Ravier, A., Novais, T., Delphin-Combe, F., Moutet, C., Xie, J., . . . Krolak-Salmon, P. (2016). Relationship Between Comorbidities in Patients With Cognitive Complaint and Caregiver Burden: A Cross-Sectional Study. *Journal of the American Medical Directors Association*, https://doi.org/10.1016/j.jamda.2015.10.011.

Ellis, R. J., Wang, Z., Genes, N., & Ma'ayan, A. (2019). Predicting opioid dependence from electronic health records with machine learning. *BioData Mining*, https://doi.org/10.1186/s13040-019-0193-0.

Hellis, E., & Mukaetova-Ladinska, E. B. (2022). Informal Caregiving and Alzheimer's Disease: The Psychological Effect. *Medicina* , https://doi.org/10.3390/medicina59010048.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., . . . Kales, H. (2020). Dementia prevention, intervention, and care. *The Lancet Commission*, https://doi.org/10.1016/s0140-6736(20)30367-6.

Loades, M. E., Chatburn, E., Higson-Sweeney, N., Reynolds, S., Shafran, R., Brigden, A., . . . Crawley, E. (2020). Rapid Systematic Review: The Impact of Social Isolation and Loneliness on the Mental Health of Children and Adolescents in the Context of COVID-19. *Journal of the American Academy of Child & Adolescent Psychiatry*, https://doi.org/10.1016/j.jaac.2020.05.009.

Murphy, G. H., Beadle-Brown, J., Wing, L., Gould, J., Shah, A., & Holmes, N. (2005). Chronicity of Challenging Behaviours in People with Severe. *Journal of Autism and Developmental Disorders*, https://doi.org/10.1007/s10803-005-5030-2.

Nathan M. Stall, A. J. (2020). For-profit long-term care homes and the risk of COVID-19 outbreaks and resident deaths. *Canadian Medical Association Journal*, https://doi.org/10.1503/cmaj.201197.

National Institute on Drug Abuse. (2022). *Drug Overdose Death Rates*. Retrieved from National Institutes of Health: https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates

Prince, M. J., Wu, F., Guo, Y., Robledo, L. M., O'Donnell, M., Sullivan, R., & Yusuf, S. (2015). The burden of disease in older people and implications for health policy and practice. *The Lancet*, https://doi.org/10.1016/s0140-6736(14)61347-7.

Public Health Agency of Canada. (2023). *Opioid- and Stimulant-related Harms in Canada. Ottawa: Public Health Agency of Canada*. Retrieved from Federal, provincial, and territorial Special Advisory Committee on the Epidemic of Opioid Overdoses: https://health-infobase.canada.ca/substance-related-harms/opioids-stimulants/

SAS. (2019). *Exploring SAS Viya.* Retrieved from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://support.sas.com/content/dam/SAS/support/en/books/free-books/exploring-sas-viya-data-mining-machine-learning.pdf

Schell, R. C., Allen, B., Goedel, W. C., Hallowell, B. D., Scagos, R., Li, Y., . . . Ahern, J. (2021). Identifying Predictors of Opioid Overdose Death at a Neighborhood Level With Machine Learning. *American Journal of Epidemiology*, https://doi.org/10.1093/aje/kwab279.

Statistics Canada. (2021). *Deaths and age-specific mortality rates, by selected grouped causes*. Retrieved from https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310039201

Tseregounis, I. E., & Henry, S. G. (2021). Assessing opioid overdose risk: a review of clinical prediction models utilizing patient-level data. *Translational Research*, https://doi.org/10.1016/J.TRSL.2021.03.012.

World Health Organization. (2019). *International Statistical Classification of Diseases and Related Health Problems (ICD)*. Retrieved from Classication of Diseases: https://www.who.int/standards/classifications/classification-of-diseases

Chapter 6
AHRs in Forecasting Mental Health Outcomes Post-Disaster

This chapter sets out to critically examine the role that AHRs might have in transforming disaster response and mental health care. It focuses on the evolving challenge of mental health crises following disasters and explores how AHRs can transition from static records to dynamic, predictive tools. This transformation is essential for anticipating and mitigating the mental health effects of disasters, leveraging ML to analyze complex AHR data patterns, and forecast needs accurately.

Critique of the Chapter

While this chapter provides a comprehensive overview of the potential for AHRs to revolutionize mental health care in the context of disasters, it would benefit from a more detailed analysis with more case studies where AHRs have been effectively utilized in disaster scenarios, offering tangible evidence of their impact. Additionally, while the ethical considerations of privacy and consent are acknowledged, a deeper exploration into how these challenges have been addressed in real-world applications would strengthen the argument. The chapter also assumes a level of interoperability and data sharing that is not universally available, suggesting an area for further exploration on overcoming these practical challenges.

Future Directions

To advance this chapter towards a comprehensive academic paper, a few changes are necessary. First, integrating specific case studies or pilot programs that demonstrate the successful application of AHRs in predicting and mitigating mental health issues post-disaster would provide empirical support to the theoretical framework presented. This is an area of research currently being explored by the author. This could include detailed examples of ML models that have accurately forecasted mental health needs and the interventions that followed in Alberta. Second, a more rigorous examination of the technological and infrastructural prerequisites for implementing AHR-based predictive models in various healthcare systems worldwide would add depth, addressing the variability in healthcare infrastructure across different regions. Lastly, expanding on the collaboration between governmental, technological, and healthcare sectors could offer insights into a multidisciplinary approach for policy development, highlighting successful strategies for overcoming the ethical and logistical barriers identified. These enhancements would not only solidify the chapter's contribution to academic discourse but also provide a clear roadmap for implementing AHRs in disaster mental health preparedness and response, making it a valuable resource for practitioners and policymakers.

6.0 AHRs in Forecasting Mental Health Outcomes Post-Disaster

This chapter addresses the growing challenge of mental health in the aftermath of disasters and the pivotal role of AHRs in improving response and preparedness. This section examines how AHRs could transform from static data collections into dynamic, predictive tools, aiding in the anticipation and mitigation of post-disaster mental health crises.

The discussion includes the potential for machine learning to interpret complex AHR data patterns to forecast mental health needs accurately. It suggests collaboration with tech and academic sectors to refine these predictive models. Additionally, it looks at how AHRs could identify populations at risk, proposing targeted, pre-emptive mental health interventions in disaster-prone areas.

A future vision for AHRs is proposed, central to an integrated disaster response system that could reshape mental health care. Necessary policy and legislative changes to enable this vision are also considered, with a call to action for a united effort to harness AHRs for disaster mental health preparedness and response.

6.1 Introduction to the Escalating Challenge of Disasters and Mental Health

The onset of the 21st century has witnessed a disturbing increase in the frequency and intensity of disasters, both natural and man-made. The resulting devastation goes beyond the immediate physical and economic toll, casting long shadows over the mental well-being of affected populations. Emerging trends indicate a surge in disaster-related mental health conditions, with an escalating number of individuals presenting symptoms of PTSD, anxiety, and depression post-disaster (Goldmann & Galea, 2014). These conditions can persist, often undiagnosed and untreated, leading to chronic psychological distress (North & Pfefferbaum, 2013).

The complexities of mental health needs in the aftermath of such tragedies are multi-layered, often exacerbated by the displacement of communities, the disintegration of social support structures in the disaster area, and the overarching uncertainty of recovery and rehabilitation. Recent reports by various health organizations underscore the pressing need for mental health services to be considered a primary component of disaster response (United Nations Office for Disaster Risk Reduction, 2022). Integrating AHRs into disaster response strategies could enable continuous monitoring and evaluation of these mental health needs well beyond current monitoring, allowing for timely adjustments to interventions and resource allocation, thus better addressing the evolving challenges faced by affected populations.

Amid this backdrop, AHRs stand poised to revolutionize mental health interventions. AHRs, with their comprehensive datasets, hold the potential to transform reactive mental health care into proactive, predictive, and personalized medicine. For instance, during the COVID-19 pandemic, administrative data was utilized to track mental health trends and allocate resources effectively. By analyzing patterns in this data, health systems were able to predict spikes in mental health issues and prepare accordingly. These methodologies can be applied to post-disaster scenarios, demonstrating the adaptability and utility of these tools in various contexts. By leveraging historical health data, AHRs can be valuable in identifying at-risk populations before the onset of symptoms, through the exploration of historical mental health outcomes in disasters, thus enabling the pre-emptive deployment of mental health resources (Patel, et al., 2018).

The predictive power of AHRs is further enhanced when coupled with disaster psychology research. Studies in this field provide valuable insights into the psychosocial impact of disasters and the resilience factors that may protect against mental health deterioration (Pfefferbaum & North, 2020). When these insights are integrated with AHRs through sophisticated data analytics, a potent tool could emerge for forecasting mental health needs and tailoring interventions accordingly.

Healthcare organizations, anticipating these developments, are beginning to advocate for and develop integrated systems that allow for the rapid assimilation and analysis of AHRs in response to disasters (World Health Organization, 2013). Predictive modeling techniques are continually refined, offering the promise of not only understanding but also anticipating the mental health aftermath of a disaster. In doing so, they lay the groundwork for a responsive mental health care system that is agile, adaptive, and attuned to the emerging challenges of a world increasingly prone to disasters.

This narrative is not without its cautions; the utilization of AHRs in such predictive capacities must navigate the ethical considerations of privacy and consent. Nonetheless, the path forward is being charted, and with AHRs potential in guiding mental health professionals towards a horizon where timely and effective intervention is not just an aspiration but a tangible reality.

The prospects for AHRs in enhancing disaster-related mental health care are significant, marked by a potential shift from the historically retrospective analysis of health data to a predictive, preventative, and patient-centered approach. As we delve further into the capabilities and mechanisms of AHRs in Sections 6.2 and 6.3, the role of these records in predicting and mitigating the mental health impacts of disasters becomes increasingly apparent, highlighting an urgent call for innovation and integration in the field of disaster mental health.

6.2 AHRs: From Archival Richness to Predictive Proficiency

The evolution of AHRs from static repositories of medical histories to dynamic engines capable of predictive analytics could herald a new era in disaster-related mental health forecasting. This transformation is driven by integrating AHRs with real-time data collection and sophisticated analytical tools, forming a feedback loop that continuously enhances predictive capabilities and intervention strategies. Further, it is anchored in anticipated advancements in the collection and nuanced integration of biopsychosocial data, which are set to exponentially enhance predictive capabilities within this specialized field.

AHRs, in their traditional form, offer a rich historical snapshot of patient interactions with health services, capturing a wide spectrum of data from diagnostic codes to medication prescriptions. However, the emerging trends extend beyond this retrospective function. The infusion of AHRs with real-time data streams, such as biometric data from wearable technologies and environmental sensor data, is primed to provide a more comprehensive picture of an individual's health (Topol, 2019). Incorporating these real-time data streams allows health systems to rapidly adapt to new information, ensuring mental health interventions are based on the most current and relevant data. The potency of AHRs in disaster mental health forecasting could grow significantly if they are also augmented by broader data sets that encompass environmental exposures, individual behavioral patterns, and potentially even genomic information, which can be particularly salient in understanding susceptibility to post-traumatic stress disorder (PTSD) and other disaster-related mental health conditions (Smoller, 2015). The inclusion of social determinants of health into AHRs—an array of socio-economic, educational, and environmental factors—has the potential to greatly amplify their utility in the context of disaster psychology. Socioeconomic status, for example, has been shown to correlate with the prevalence and severity of mental health outcomes following disasters (Galea, Merchant, & Lurie, 2020). Machine learning algorithms, leveraging this enriched data, have the potential to discern complex patterns indicative of mental health vulnerabilities in the face of disasters—patterns that would otherwise remain undetected through conventional analytical means. This relevance is underscored by case studies discussed earlier, where machine learning was used to predict opioid overdoses and developmental disorders during the pandemic, showcasing the broad applicability of these techniques across various disaster-related mental health scenarios.

Privacy considerations are paramount as AHRs evolve. The burgeoning field of privacy-preserving technologies promises to mitigate the risks associated with the broader use of AHRs for predictive analytics. Innovations such as synthetic data, differential privacy and federated learning are potentially key technologies to protect against the risks of personal data exposure, ensuring that individual identities remain shielded while allowing for the collective benefit of shared knowledge (Rieke, et al., 2020). Furthermore, blockchain technology's potential application to AHRs offers an intriguing possibility of a secure, immutable ledger, enhancing data integrity and exchange while upholding individual autonomy over personal health information (Kuo, Kim, & Ohno-Machado, 2017). This secure exchange mechanism is not just a protective measure; it also facilitates the construction of a comprehensive data ecosystem, essential for accurate disaster mental health forecasting.

Ultimately, the transition of AHRs towards predictive proficiency rests not only on the integration of expansive and deep data but also hinges on the parallel development of technologies that safeguard the privacy of the data subjects. As we navigate this promising yet complex terrain, the simultaneous enhancement of data analytic capabilities and privacy safeguards in AHRs will have to be a cornerstone in the future of mental health intervention strategies amidst disasters, if we want to enable healthcare systems to deliver timely, targeted, and effective mental health interventions in an increasingly unstable environment.

6.3 Machine Learning and Predictive Modeling with AHRs

The application of machine learning algorithms to AHRs represents a convergence of data science and clinical insights, yielding predictive models with the potential to transform mental health care in the aftermath of disasters. The incorporation of machine learning into AHRs enables the extraction of complex patterns and the forecasting of potential mental health crises at individual and population levels, with a precision previously unattainable through traditional statistical methods.

Machine learning algorithms can integrate diverse data within AHRs, from diagnostic codes to pharmaceutical prescriptions, alongside more granular data points such as laboratory results, and even the frequency and timing of healthcare service utilization. These algorithms learn from the data to identify risk factors and signals that precede mental health issues, particularly those exacerbated by disaster-related stressors.

Case studies, such as those examining the aftermath of hurricanes or large-scale fires, reveal the efficacy of machine learning in identifying populations at risk for PTSD, depression, or anxiety disorders. For instance, after Hurricane Katrina, researchers used AHRs to track mental health prescriptions and medical visits to project spikes in PTSD occurrences in affected regions (Kessler, et al., 2008). In these models, predictors included not only prior mental health conditions but also factors like loss of employment, housing instability, and the breakdown of social support structures.

Furthermore, theoretical models have posited the application of unsupervised learning techniques to discern latent patterns within AHRs that might not be immediately obvious to human observers or traditional analyses. Such models could, theoretically, identify clusters of symptoms or conditions that frequently precede mental health crises in post-disaster contexts, thus enabling pre-emptive interventions.

The promise of machine learning in enhancing AHRs for disaster mental health preparedness also invites collaborations between healthcare providers, data scientists, and policymakers. By working together, these stakeholders can ensure the development of models that are not only accurate but also ethically constructed and considerate of the nuances of mental health conditions (Char, Shah, & Magnus, 2018).

To fully realize the potential of machine learning in this domain, several critical steps must be undertaken. These include the standardization of AHR data formats to ensure interoperability, the ethical use of data with due respect for privacy and consent, and the development of robust models that are generalizable across diverse populations and disaster scenarios.

As ML and AHRs continue to evolve, the future points towards a proactive, data-informed approach to disaster mental health, with predictive models serving as both early-warning systems and guides for resource allocation and intervention planning, ensuring that support reaches those who need it most, precisely when they need it.

6.4 Identifying Populations at Risk Using AHRs

Identifying populations at risk for adverse mental health outcomes following disasters is a potential application of AHRs, particularly with the integration of algorithms capable of real-time risk assessment. The evolution of these algorithms is likely now at a point where larger government entities and organizations could harness live data streams, such as social media activity, emergency calls, and environmental monitoring, which could provide immediate insights into areas affected by disasters.

The potential of these algorithms lies in their ability to analyze vast and varied data sources in real-time, thereby offering an unprecedented immediacy in the identification of at-risk populations. For instance, by incorporating live data on power outages, supply chain disruptions, and shelter occupancy rates, algorithms could quickly flag regions where the compounded stressors may lead to heightened mental health risks. Additionally, sentiment analysis of social media platforms can offer a real-time pulse on community morale and stress levels, further refining risk assessments.

I anticipate future research to focus on refining these algorithms for greater accuracy. This refinement process will likely involve the incorporation of advanced ML techniques, such as deep learning and newer large language models, which can handle unstructured data and learn more complex representations. Furthermore, longitudinal studies could aid in the improvement of these models by providing insights into the progression of mental health outcomes over time, allowing for the adjustment of an algorithm's parameters based on long-term data (Rieke, et al., 2020).

Another area of potential advancement is the use of simulation-based approaches to test the effectiveness of different algorithmic strategies in virtual environments that mimic real-world disasters. This could enable researchers to refine their models in a controlled setting before deployment in actual disaster scenarios, thus enhancing the models' predictive accuracy and reliability (Schrittwieser, et al., 2020).

The accuracy of these predictive models also depends on the level of detail of the AHRs. Enriching these records with more detailed information about individual and community resilience factors, such as access to social support networks and mental health services, could significantly improve the ability of algorithms to identify those at greatest risk especially when we consider urban vs rural disasters. Moreover, interdisciplinary research that integrates insights from psychology, sociology, and disaster science could provide a more comprehensive framework for risk assessment models.

To operationalize these advancements, research must also navigate the ethical considerations of real-time monitoring and predictive analytics, ensuring that privacy is preserved and that the benefits of such technology are equitably distributed. As we advance in this direction, the aspiration is that AHRs, augmented by ML, will not only identify at-risk populations but also facilitate the delivery of pre-emptive interventions, thereby mitigating the mental health impact of disasters.

6.5 Targeted Interventions Informed by AHRs

The strategic deployment of mental health resources in disaster-prone areas represents a novel application of AHRs, leveraging their predictive capabilities to inform and guide interventions. The contemplation of a regional, national, or global digital infrastructure, underpinned by AHRs, signifies a forward-thinking approach to disaster mental health preparedness.

This envisioned infrastructure could likely integrate AHRs with geographic information systems (GIS) and disaster risk management databases to create a comprehensive system capable of not only forecasting potential mental health crises but also allocating resources efficiently. For example, by identifying regions with a high density of individuals with pre-existing mental health conditions and overlaying this data with environmental risk factors, the system could direct resources such as mobile counseling units or telepsychiatry services to those areas in anticipation of disaster events.

Future pilot programs are crucial for testing the efficacy of targeted interventions informed by AHR-driven predictive analytics. These pilots could be designed to evaluate both the effectiveness of the interventions and the efficiency of the infrastructure in deploying them. By implementing such programs in diverse settings and disaster scenarios, researchers could gather valuable data on the system's performance, leading to iterative improvements.

For example, a pilot program in a fire-prone region might utilize AHRs to identify communities with elevated levels of PTSD following past events. The program could then pre-deploy mental health resources to these communities during the fire season and assess the intervention's impact on mitigating PTSD symptoms compared to communities that did not receive pre-emptive resources.

The success of these interventions would also depend on their cultural and contextual relevance, necessitating collaboration between data scientists, healthcare professionals, and community stakeholders. By tailoring interventions to the specific needs and characteristics of each community, the effectiveness of pre-emptive deployments can be increased. Allowing for a global level system that can tailor outcomes at the local and individual needs of the community. Moreover, the long-term sustainability and scalability of these interventions would be enhanced by continuously updating the AHRs with post-intervention outcomes. This would not only help refine the predictive models but also provide an evidence base for the allocation of resources, thus informing policy at both national and global levels (Patel, et al., 2018). As such, targeted interventions informed by AHRs hold the promise of transforming mental health disaster response from a reactive to a proactive stance. The establishment of a robust digital infrastructure and the conduct of pilot programs are the next logical steps in operationalizing this vision, with the ultimate goal of enhancing the resilience of disaster-prone communities and mitigating the mental health impact of catastrophic events.

6.6 Monitoring and Adjusting Interventions through AHRs

The dynamic nature of disaster impacts on mental health necessitates a similarly dynamic system of intervention monitoring and adjustment. AHRs could serve as the cornerstone of continuous learning systems, designed to adjust mental health interventions in real-time, responding to feedback from ongoing treatment outcomes.

In such a system, AHRs would not merely serve as static records but as living documents that are continuously updated with patient responses to interventions. For instance, imagine a scenario where AHRs are updated in real-time with data on patients' symptomatic responses to different treatment modalities following a disaster. ML algorithms could analyze this data to identify which interventions are most effective for specific symptoms or demographics. This analysis could then be used to dynamically adjust treatment protocols to enhance their effectiveness. Continuous learning systems would necessitate the development of sophisticated models that can analyze complex, non-linear data in real-time. This concept isn't new as technologies like large language models are starting to pave the way for such ideas. These models could be trained to recognize the signs of improvement or deterioration in mental health status and suggest modifications to the treatment regime accordingly. For example, if a particular medication or therapy is found to be less effective for a subset of patients, the system could recommend alternative interventions.

The global standardization of intervention monitoring could further refine the efficacy and responsiveness of AHR-driven continuous learning systems. Establishing international standards for AHRs, like the increased use and rollout of ICD codes, in the context of disaster mental health response would ensure consistency in data collection, analysis, and response across different healthcare systems and regions. These standards could facilitate cross-border collaborations, enabling a unified response to global mental health challenges posed by disasters. This would also mean that responses could be predictive and tailored to the region’s needs and norms, even though resources are coming from potentially vary different regions and people. Beyond the use of ICD codes, such international standards would need to include protocols for data entry, outcome measurement, privacy protection, and the ethical use of predictive analytics. They could also define benchmarks for treatment efficacy and guidelines for adjusting interventions based on data-driven insights (World Health Organization, 2022). The development of these continuous learning systems and international standards would benefit from multi-sectoral collaboration involving governments, health organizations, technology companies, and patient advocacy groups. Together, these stakeholders could work toward a shared goal of enhancing the precision and adaptability of mental health interventions in the aftermath of disasters.

In sum, the potential of AHRs to revolutionize the monitoring and adjustment of mental health interventions in disaster settings is contingent upon the advancement of continuous learning systems and the establishment of international standards. By embracing these innovations, the global healthcare community can look forward to improved outcomes in disaster mental health responses, tailored to the evolving needs of affected populations.

6.7 AHRs in the Broader Context of Disaster Preparedness and Response

In envisioning the future of disaster preparedness and response, AHRs stand as a potential linchpin in an integrated response framework that could fundamentally transform emergency management and mental health care delivery. The incorporation of AHRs into disaster response strategies heralds a shift towards a more proactive, personalized, and precise approach to mental health interventions in crisis situations. This envisioned framework includes real-time data collection, continuous feedback mechanisms, and advanced analytics, which collectively create a system that not only responds to crises but also anticipates and mitigates them.

An integrated disaster response framework featuring AHRs would be characterized by its anticipatory and adaptive capabilities. It would leverage real-time data from AHRs to forecast mental health needs and tailor interventions accordingly. Developing this framework would involve establishing standardized protocols for data collection and analysis, creating secure data-sharing networks, and forming multidisciplinary teams to oversee the implementation and evaluation of interventions. Such a system would facilitate seamless collaboration between emergency responders, healthcare providers, and mental health professionals, ensuring that mental health care is a central consideration in disaster response efforts.

The actualization of this vision would rely on robust legislative and policy initiatives. Firstly, legislation would need to be enacted to ensure that AHRs can be utilized effectively during disasters. This may involve mandating the inclusion of mental health indicators in AHRs and establishing protocols for their use in emergencies. Additionally, policy changes are required to support the interoperability of health information systems, enabling the seamless exchange of AHRs across different jurisdictions and healthcare providers.

Privacy laws would also need to be revisited and potentially revised to protect sensitive data while allowing for the necessary sharing of information during a disaster. This could involve creating exceptions within existing health privacy regulations such as HIPAA in the

United States and PIPEDA in Canada, to allow for the sharing of health data for the purpose of disaster response without compromising individual privacy.

Furthermore, investment in the technological infrastructure to support the integration of AHRs into disaster preparedness and response would be valuable. This includes funding for the development of advanced analytics tools, training for healthcare providers in their use, and the creation of secure networks for data sharing.

Policies encouraging the development of emergency preparedness plans that specifically address mental health needs, informed by insights from AHRs, could also be useful. These plans could outline strategies for rapid mental health assessments, triage procedures, and the deployment of mental health resources in the immediate aftermath of a disaster.

Lastly, international collaboration on policies related to AHRs in disaster response is needed to address global disparities in mental health care during emergencies. The establishment of an international consortium dedicated to advancing the use of AHRs in disaster mental health could be proposed. This consortium could act as a platform for sharing best practices, standardizing data protocols, and fostering global collaboration. By doing so, it would ensure that the benefits of AHRs are realized globally, promoting a unified approach to disaster mental health preparedness and response.

While these concepts have immense hurdles to overcome, the broader integration of AHRs into disaster preparedness and response demands concerted efforts in policy-making, technological advancement, and international cooperation. By establishing a framework that centralizes the role of AHRs in emergency management, a more effective, efficient, and empathetic response to the mental health impacts of disasters could be achieved.

6.8 Conclusion and Future Directions

The convergence of AHRs and disaster mental health readiness presents a novel opportunity to reshape our approach to health crises. Realizing the added potential of AHRs in disaster mental health readiness and response demands a concerted effort from stakeholders across the spectrum of healthcare, emergency management, technology, industry, and government.

It is important that research in this field adopts a multidisciplinary lens, drawing from the insights of psychiatry, epidemiology, data science, and disaster response. Only through a synthesis of these diverse perspectives can we develop comprehensive models that accurately forecast mental health needs and interventions in the aftermath of disasters.

Investment in technology is critical—particularly in developing advanced analytical tools that can process and interpret the vast amounts of data contained within AHRs. Significant funding would also need to be allocated to the creation of secure platforms for data sharing, which would be essential for the rapid sharing of information during emergencies.

Policy reform is another cornerstone for progress. Laws and regulations need to evolve to support the utilization of AHRs in emergency situations, balancing the imperatives of privacy with the need for swift and decisive mental health interventions when disasters strike.

To galvanize international efforts and ensure a collective movement forward, the proposal of an international consortium dedicated to advancing the use of AHRs in disaster mental health could be a potential solution, this kind of work could be done under the umbrella of the WHO, which already works on several aspects of this problem. This consortium could serve as a central location for innovation, dialogue, and resource sharing, fostering a global partnership in pursuit of resilient healthcare systems that can withstand the mental health impacts of disasters.

Such a consortium could also act as a steward for the development of global standards and best practices, driving consistency in how AHRs are employed across different countries and healthcare contexts. It could facilitate cross-border research collaborations, develop funding mechanisms for pilot projects, and advocate for policies that support the integration of AHRs into disaster preparedness plans.

In summary, this chapter emphasizes the importance of AHRs in enhancing our ability to predict, prepare for, and respond to the mental health consequences of disasters. As we look to the future, it is incumbent upon us to lay the groundwork for a new era of disaster mental health readiness—one that is underpinned by data-driven insights, supported by policy, and realized through global cooperation. The journey ahead is complex and challenging, but the destination—a world better equipped to protect and promote mental health in the face of disasters—is a compelling and worthy pursuit.

6.9 From Disaster Mental Health Forecasting to Wearable Integration in Care

Moving from the exploration of AHRs in forecasting mental health outcomes post-disaster, we transition to the potential for innovative integration of AHRs with wearable device data. The next chapter emphasizes the progression towards a more anticipatory and personalized mental health care framework, driven by the combination of AHRs and real-time health data. Both chapters underscore the potential benefits of harnessing data analytics to refine and elevate mental health interventions in an era marked by increasing disaster-related challenges. By drawing a connection between the predictive capabilities of AHRs in the aftermath of disasters and the potential of wearable technologies to enhance these predictions, we underscore a strategic pivot towards a healthcare system that has the capacity to not just be reactive but also pre-emptive.

References

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp1714229.

Galea, S., Merchant, R. M., & Lurie, N. (2020). The Mental Health Consequences of COVID-19 and Physical Distancing. *JAMA Internal Medicine*, https://doi.org/10.1001/jamainternmed.2020.1562.

Goldmann, E., & Galea, S. (2014). Mental Health Consequences of Disasters. *Annual Review of Public Health*, https://doi.org/10.1146/annurev-publhealth-032013-182435.

Kessler, R. C., Galea, S., Gruber, M. J., Sampson, N. A., Ursano, R. J., & Wessely, S. (2008). Trends in mental illness and suicidality after Hurricane Katrina. *Molecular Psychiatry*, https://doi.org/10.1038/sj.mp.4002119.

Kuo, T.-T., Kim, H.-E., & Ohno-Machado, L. (2017). Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association*, https://doi.org/10.1093/jamia/ocx068.

North, C. S., & Pfefferbaum, B. (2013). Mental health response to community disasters: a systematic review. *JAMA*, https://doi.org/10.1001/jama.2013.107799.

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., . . . Jordans, M. J. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet*, https://doi.org/10.1016/S0140-6736(18)31612-X.

Pfefferbaum, B., & North, C. S. (2020). Mental Health and the Covid-19 Pandemic. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp2008017.

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H., Albarqouni, S., . . . Cardoso, M. J. (2020). The future of digital health with federated learning. *npj Digital Medicine*, https://doi.org/10.1038/s41746-020-00323-1.

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., . . . Silver, D. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* , https://doi.org/10.1038/s41586-020-03051-4.

Smoller, J. W. (2015). The Genetics of Stress-Related Disorders: PTSD, Depression, and Anxiety Disorders. *Neuropsychopharmacology* , https://doi.org/10.1038/npp.2015.266.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

United Nations Office for Disaster Risk Reduction. (2022). *Global Assessment Report on Disaster Risk Reduction.* United Nations.

World Health Organization. (2013). *Building back better: sustainable mental health care after emergencies.* World Health Organization.

World Health Organization. (2022). *Mental health in emergencies*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/mental-health-in-emergencies

Chapter 7
Integrating AHRs with Wearable Device Data

This chapter explores the transformative potential of combining AHRs with real-time data from wearable devices in mental health care. It assesses how this integration could improve patient engagement, enable predictive analytics, and foster a proactive approach to treatment. Highlighting both opportunities and challenges, including ethical and privacy concerns, the chapter draws on case studies and pilot programs to illustrate practical applications. It helps readers to envision a future where mental health care is personalized and empowered by data-driven insights, emphasizing the importance of navigating technological, regulatory, and economic considerations to realize the benefits of this innovative integration.

Critiques of the Chapter

This chapter presents an aspirational perspective on enhancing mental health care through technology. However, it could be improved by addressing a few key areas. Firstly, while the potential of integrating AHRs with wearable device data is well-articulated, the chapter could benefit from a deeper dive into the technical challenges and solutions related to data integration, interoperability, and the real-time processing of wearable device data. A more detailed exploration of these aspects would provide a fuller understanding of the practicalities involved in achieving the envisioned integration. Additionally, while the chapter touches on ethical considerations, a more comprehensive analysis of the ethical dilemmas, particularly around consent in the context of continuous monitoring, would enrich the discussion. Additionally, incorporating perspectives on patient autonomy and the potential for over-surveillance would offer a more balanced view of the implications of this technological integration. Finally, the chapter could extend its discussion on the economic impacts and healthcare cost implications by including more concrete examples or models that demonstrate cost savings or return on investment from similar integrations in other fields or pilot studies.

Future Directions for the Chapter as a Paper

To evolve this chapter into a comprehensive academic paper, several enhancements are potentially necessary. The paper could start by expanding on the technical specifics of integrating AHRs with wearable device data, including the use of emerging technologies and standards for data interoperability, privacy-preserving data analytics, and the management of large-scale health data ecosystems. A deeper analysis of ethical considerations, drawing on case studies or theoretical models, would provide good insights into managing the balance between innovation and individual rights. The paper could also benefit from a more detailed examination of the economic implications, including cost-benefit analyses, potential for reducing healthcare expenditures, and the economic barriers to widespread adoption. These additions would not only extend the academic contribution of the chapter but also offer practical guidance for researchers, practitioners, and policymakers interested in the intersection of healthcare technology and mental health care.

7.0 Integrating AHRs with Wearable Device Data

The integration of AHRs with data from wearable devices represents a significant step forward in the field of mental health care. This section explores the transformative potential of combining real-time, granular data from wearables with the comprehensive background provided by AHRs. By analyzing current practices, technological advancements, and the capacity for predictive analytics, we examine the implications of this integration across various domains, including the enhancement of patient engagement, the evolution of regulatory frameworks, and the economic impacts on the healthcare system. Moreover, integrating AHRs with wearable device data can significantly enhance the management of various mental health conditions, such as opioid use disorder, neurocognitive disorders, and developmental disorders. For instance, real-time data from wearables can help identify early warning signs of opioid misuse, track cognitive decline in neurocognitive disorders, and monitor developmental progress in children, providing critical insights that inform timely and personalized interventions.

We navigate through the challenges and opportunities presented by this integration, considering the ethical implications and the importance of data privacy in the burgeoning age of digital health. Case studies and pilot programs are discussed to provide a practical perspective on the application of these technologies in real-world settings. This chapter culminates in a forward-looking discussion about the future of mental health care—envisioning a system that is proactive, personalized, and empowered by data-driven insights.

7.1 Real-time, Granular Insights into Mental Health Status

Technological advancements in wearable devices have substantially augmented the potential for real-time health monitoring, providing continuous streams of data that were previously inaccessible outside of clinical settings. Wearables now extend beyond mere step counters and now encompass sophisticated sensors capable of tracking a wide array of physiological and behavioral metrics.

Devices such as smartwatches and fitness trackers can monitor heart rate variability (HRV), which is increasingly recognized as an important psychophysiological marker linked to the autonomic nervous system (ANS) (Shaffer & Ginsberg, 2017). The ANS responses can be indicative of psychological states such as stress or relaxation. Moreover, accelerometers and gyroscopes in wearables detect movement patterns, which can be informative about activity levels, sleep patterns, and circadian rhythms—all relevant to mental health (Faurholt-Jepsen, et al., 2014).

Sleep is a particularly critical aspect of mental health, with irregularities often serving as a precursor or symptom of psychiatric disorders (Wulff, Gatti, Wettstein, & Foster, 2010). Wearables can track sleep duration, disturbances, and quality by measuring movement and heart rate throughout the night. This data provides granular insights into sleep patterns that could, for instance, signal the onset of a depressive episode before it fully emerges (Beattie, Kyle, Espie, & Biello, 2015).

Furthermore, skin conductance sensors measure electrodermal activity (EDA), which is linked to arousal and can be a proxy for emotional states and stress (Boucsein, 2013). Combined with self-reported mood logs through companion smartphone apps, this physiological data can enrich our understanding of an individual's mental health landscape in situ.

Research suggests that when these real-time data points are aggregated and analyzed using ML algorithms, they can predict mental health states with considerable accuracy. Torous et al. (2020) demonstrated the utility of smartphone data in predicting episodes of mania and depression in bipolar disorder, highlighting the transformative potential of integrating such granular data with traditional healthcare records.

Incorporating real-time data from wearables into daily mental health monitoring has the potential to revolutionize patient care (Torous, Staples, & Onnela, 2015). By continuously updating healthcare providers on a patient's physiological and psychological status, interventions can be precisely timed and personalized, fostering a more dynamic and responsive care approach. For example, significant changes in HRV or EDA might prompt a mental health professional to check in with a patient, potentially averting a crisis. Such real-time data could be invaluable in managing opioid use disorders by detecting signs of relapse or overdose risk, monitoring cognitive function in neurocognitive disorders, and tracking developmental milestones in children, ensuring that interventions are timely and tailored to individual needs.

In general, it seems that even with a few basic sensors that already exist in some of the most popular smartwatches that the integration of real-time data from wearable devices with AHRs could substantially transform mental health monitoring. By providing a detailed and dynamic picture of an individual's physiological and behavioral status, healthcare professionals can anticipate changes in mental health, offering timely interventions that could pre-empt the progression of mental health issues. This capability is particularly beneficial for proactive mental health strategies across various conditions, such as providing immediate support during crises, optimizing treatment, managing chronic disorders, and supporting developmental health.

7.2 Transforming Predictions: From Reactionary to Proactive

Historically, AHRs have been indispensable in chronicling patient histories, offering a retrospective glance at an individual's interactions with health services. However, their utility in predicting mental health issues is inherently constrained by their static nature. AHRs often lag behind the current status of a patient, reflecting past diagnoses, treatments, and health care encounters, sometimes years out of date, rather than presenting an up-to-the-minute picture (Monteith, Glenn, Geddes, Whybrow, & Bauer, 2016). Moreover, the retrospective analysis of AHRs typically identifies risk factors and patterns after a mental health condition has manifested, thus situating AHRs predominantly within a reactionary framework of mental health care.

The infusion of real-time data from wearable devices into mental health assessments could significantly enhance this challenge. Real-time monitoring introduces the possibility of identifying nuanced patterns and trends that may predict episodes or changes in mental health status. This information can be enhanced by the patient's historical information contained in AHRs. For instance, sudden alterations in sleep patterns or activity levels, when analyzed alongside historical health records, could signal impending manic or depressive episodes in patients with bipolar disorder (Faurholt-Jepsen, et al., 2014). Similarly, subtle changes in HRV could signal an oncoming anxiety attack, allowing for timely intervention even through prompts given to the patient via their smart device. The real-time data thus holds immense potential to shift the system from reactive to one that is proactive. By providing ongoing updates to healthcare providers and the patient directly, this data integration enables the development of predictive models that are significantly more nuanced and immediate than those relying solely on AHRs. These models could, in effect, identify warning signs and trigger interventions before the patient is even aware of a potential issue, thus transitioning mental health care into a realm where prevention is as attainable as treatment.

Envisioning the future of mental health care, one can anticipate a discipline profoundly influenced by the proactive utilization of integrated data. ML models trained on combined datasets of real-time and historical health information could anticipate patient needs and tailor care pathways accordingly (Torous & Baker, 2016). This could not only improve outcomes but also empower patients by offering them insights into their mental health, fostering a partnership model of care between patients and providers.

The concept of digital phenotyping, which involves using data from smartphones and other wearables to detect the onset of mental health issues, illustrates this potential (Insel, 2017). Such technology can enable continuous assessment and the application of just-in-time adaptive interventions (JITAIs), which are treatments that are adaptive to an individual's context and administered as problems are detected, often in real-time (Nahum-Shani, et al., 2017).

In sum, the proactive future of mental health care powered by data integration promises a transformative impact on both the prediction and prevention of mental health issues. By leveraging the depth and immediacy of wearable data in conjunction with the breadth of AHRs, a new horizon in computational psychiatry and personalized mental health care is on the cusp of realization.

7.3 The Role of Machine Learning and Advanced Analytics

ML algorithms represent a central component in the synthesis of AHRs with data derived from wearable technology for the purpose of advancing mental health analytics. The ability of ML to handle large, complex, and non-linear datasets makes it an ideal set of tools for discerning patterns and predictive signals within the intricate web of factors influencing mental health. ML algorithms can integrate the continuous stream of biometric and behavioral data from wearables with the extensive, multi-dimensional datasets found in AHRs, thus enabling a level of analysis that is both nuanced and predictive.

For instance, deep learning models, a subset of ML, have shown significant promise in detecting depressive symptoms by analyzing speech patterns and facial expressions captured through smartphones (Alhanai, Ghassemi, & Glass, 2018). When combined with clinical history from AHRs, such models could provide a more comprehensive understanding of an individual's mental health trajectory and potential triggers for intervention. In opioid use disorder, similar models could analyze behavioral cues and physiological data to predict relapse. For neurocognitive disorders, these models could monitor speech and movement patterns to detect early signs of cognitive decline. In developmental disorders, they could track speech and motor development to identify delays and recommend early interventions.

Looking towards future developments, it could be anticipated that ML models will become increasingly sophisticated, perhaps advancing to the point of unsupervised learning, where they can detect new patterns and correlations without explicit programming. Moreover, the evolution of ML could lead to the development of personal health navigators — intelligent systems that not only monitor health indicators but also interact with users to guide them towards healthier behavioral patterns, adapting recommendations in real-time based on the user's data (Luxton, 2014).

However, with these advancements comes a paramount concern for ethics and privacy. The management and analysis of sensitive health data necessitate stringent protocols to ensure confidentiality and consent. Data de-identification, robust encryption, and transparent data governance policies are essential safeguards to protect individual privacy. Ethical ML in health care requires algorithms to be free from bias and their development process to be transparent, allowing scrutiny and assurance that these tools serve all segments of the population equitably (Vayena, Blasimme, & Cohen, 2018).

Moreover, the intersection of wearable technology and mental health analytics must navigate the delicate balance between beneficial oversight and intrusive surveillance. Ethical considerations must include the right to informational self-determination and the need for informed consent processes that genuinely reflect the ongoing, dynamic nature of data collection in this field (Martinez-Martin, Insel, Dagum, Greely, & Cho, 2018).

As ML tools continue to penetrate the domain of mental health analytics, the imperative to harmonize these technological advancements with ethical needs will become increasingly pronounced. Ensuring that privacy and ethical considerations keep pace with technological innovation will be essential to maintaining public trust and realizing the full potential of integrating AHRs with wearable device data for mental health care.

7.4 Case Studies and Pilot Programs

Exploring the emerging integration of AHRs with wearable device data involves a close examination of specific instances where such endeavors have been undertaken. The case studies and pilot programs described herein provide critical insights into the practical applications and implications of this intersection, offering valuable lessons for future initiatives.

In a notable pilot program that focused on patients with Major Depressive Disorder (MDD), wearable devices were employed to track various physiological markers such as physical activity, sleep patterns, and HRV. The subsequent integration of these data with AHRs—which contained comprehensive medical histories, medication regimens, and prior episode details—afforded researchers a multi-dimensional perspective on the predictive factors of depressive episodes. Notably, the combined data streams demonstrated a superior predictive capacity compared to the use of AHRs in isolation, suggesting a marked improvement in the timeliness and appropriateness of therapeutic interventions (Torous, Staples, & Onnela, 2015).

However, this integration was not without challenges. The considerable volume of data generated necessitated robust solutions for data storage and management. Additionally, disparities in data standards between wearable technologies and AHRs presented significant interoperability challenges. Privacy considerations also emerged as a critical concern, necessitating stringent data governance policies to ensure the confidentiality of sensitive health information.

Another case study of interest took place during the COVID-19 pandemic and involved monitoring stress and burnout among healthcare workers. Wearables were utilized to measure biometric stress markers, which were then correlated with professional parameters such as work schedules and patient interactions, as recorded in AHRs. Early findings from this study indicated a tangible link between work-related factors and stress levels, offering a data-driven basis for potential scheduling and workload adjustments (Clingan, et al., 2021). Yet, maintaining consistent use of the wearable devices by the healthcare workers proved to be a significant hurdle, affecting the completeness and reliability of the data collected. The need for sophisticated analytical models to meaningfully interpret biometric data became evident, as did the importance of integrating these insights into workforce management practices.

In the context of chronic disease management, a review that centered on patients with diabetes showcased how continuous glucose monitoring (CGM) data, gathered through wearable devices, could be synthesized with AHRs to predict and better manage glycemic episodes. This study highlighted the utility of predictive analytics in enhancing glucose control and underscored the need for real-time intervention protocols. It also emphasized the importance of patient engagement, demonstrating that providing patients with access to their integrated data could significantly boost self-management (Rhee, Kim, Shin, & Stein, 2020). Here too, concerns around algorithmic bias and the ethical dimensions of continuous monitoring were present, emphasizing the need for careful consideration of these issues as this field progresses. The lessons from chronic disease management can be applied to mental health, particularly in managing chronic disorders, where continuous monitoring and real-time interventions may significantly improve outcomes.

Collectively, these case studies underscore the transformative impact that the integration of AHRs with wearable device data can have on healthcare. They point to the necessity of scalable data infrastructure, standardized data exchange protocols, robust privacy frameworks, and the active engagement of both patients and healthcare providers. The ethical and privacy-related challenges that arise from such integrations cannot be overemphasized and require diligent attention to ensure that the advancements in healthcare technology translate into equitable and responsible improvements in patient care.

7.5 Implementation Challenges and Considerations

The successful integration of AHRs with data from wearable devices is contingent upon overcoming several significant barriers. One such barrier is the issue of interoperability and the lack of standardized protocols that can seamlessly integrate the high-frequency, heterogeneous data from wearables with the structured, often siloed data of AHRs. A potential solution to this challenge is the proactive development and adoption of healthcare data standards that enable cross-platform communication and data harmonization.

Moreover, the imperative of maintaining the privacy and security of health data adds layers of complexity to this integration. The expanded data ecosystem presents a larger surface for potential breaches, necessitating stringent cybersecurity measures and compliance with privacy regulations. Building secure data-sharing platforms with end-to-end encryption, and adopting frameworks such as the Zero Trust model, can substantially mitigate these risks. The Zero Trust model is a security concept centered on the belief that organizations should not automatically trust anything inside or outside their perimeters and instead must verify anything and everything trying to connect to their systems before granting access (Microsoft, 2023). Patient consent and engagement also emerge as pivotal factors. To surmount this challenge, it is important to foster transparency with patients regarding data usage and actively involve them in the benefits of technology through education and outreach efforts.

Technical and resource constraints present another challenge, as the assimilation and analysis of vast amounts of data require substantial computational resources and specialized personnel. Investment in cloud computing infrastructures and the hiring and training of skilled data scientists are viable strategies to address these limitations and something already seen in the larger tech companies. Additionally, algorithmic transparency and bias are concerns that must be addressed to avoid perpetuating health disparities. Implementing rigorous validation processes and adopting inclusive algorithm development practices are important steps toward creating equitable and effective analytical tools. Finally, navigating the regulatory and ethical landscape is essential for the responsible advancement of this integration. Ensuring that research and implementation strategies undergo ethical review and adhere to regulatory standards is necessary for maintaining public trust and safeguarding participants' welfare.

In integrating AHRs with wearable device data, the path forward lies in a multidisciplinary approach that considers technical feasibility, ethical integrity, and above all, patient-centeredness. Strategic investments, collaborations across sectors, and ongoing dialogue with stakeholders are necessary to navigate these challenges and harness the full potential of these technologies to revolutionize mental health care.

7.6 Regulatory and Policy Framework Evolution

The current regulatory and policy landscape for integrating AHRs with data from wearable devices is at a critical juncture. The dynamic nature of digital health data collection presents unique challenges to existing privacy laws and regulatory frameworks that were not originally designed to address the nuances of continuous, passive data monitoring and the complexities of data analytics in mental health care.

For most provinces in Canada, PIPEDA sets the groundwork for the protection of personal data in the private sector, including health information in certain contexts. This is complemented by provincial legislations such as Ontario's Personal Health Information Protection Act (PHIPA), which regulates the use of personal health information within the province (Office of the Privacy Commissioner of Canada, 2020).

As these technologies progress, there is a clear need for regulatory evolution to keep pace with advancements. The existing frameworks will require updates to ensure they encapsulate the scope of data generated by wearable devices, as well as the analytical capabilities of machine learning in health care. A number of researchers are already championing the development of regulatory policies that specifically address the collection, use, and sharing of digital health data (Vayena, Blasimme, & Cohen, 2018).

Future policy developments are likely to see an emphasis on national and international standards for data interoperability, which will be essential for the seamless integration of AHRs and wearable device data across different jurisdictions and health care systems (Topol, 2019). Moreover, as precision medicine advances, there will be a push towards regulations that facilitate the sharing and analysis of data while ensuring robust protections for patient privacy (National Academies of Sciences, Engineering, and Medicine, 2019).

One of the most significant anticipated developments in policy will revolve around consent. With wearable technologies enabling continuous data collection, traditional consent models are often inadequate. Innovative consent frameworks, such as dynamic consent, could become more prevalent, allowing individuals to have greater control over their data (Kaye, et al., 2014). Another important area will be the formulation of policies to govern the ethical use of ML and AI in health care. Ensuring algorithmic transparency and addressing potential biases will be challenging but likely necessary to maintain trust and to avoid perpetuating disparities in health care (Char, Shah, & Magnus, 2018). Finally, regulatory bodies may consider the concept of data stewardship, where data is managed by trustees who oversee the ethical use of health data for research and health care improvement, balancing the need for innovation with privacy concerns (Aitken, Jorre, Pagliari, Jepson, & Cunningham-Burley, 2016).

In the end, the regulatory and policy framework for the integration of wearable device data with AHRs is poised for transformation. To effectively manage this integration while upholding high standards of patient privacy and data security, there is a need for an adaptive, forward-looking approach that ensures ethical governance and fosters trust among all stakeholders in the health care ecosystem.

7.7 Economic Impact and Healthcare Cost Implications

The integration of AHRs with data from wearable devices is anticipated to have a substantial economic impact on the healthcare system. By merging the longitudinal health data with the real-time physiological and behavioral data from wearables, there is a potential significant shift in the approach to healthcare delivery—especially in the domain of mental health, where early intervention and continuous monitoring could significantly alter patient outcomes and system efficiencies. Cost savings in healthcare could be substantial with the introduction of predictive analytics using integrated data. For instance, by identifying early signs of mental health deterioration, interventions can be applied before conditions worsen and require more intensive, and often expensive, treatment options. Research has indicated that mental health conditions, when left untreated, can escalate to serious levels, resulting in increased emergency room visits, hospitalizations, and a need for more intensive services—all of which contribute to higher healthcare costs (Stephens & Joubert, 2001). Moreover, the proactive healthcare model facilitated by the combination of AHRs and wearable data may lead to better resource allocation. By preventing mental health crises, healthcare systems can avoid the economic strain of acute care episodes and reduce the long-term costs associated with chronic mental health conditions. This shift towards preventive care could be a key driver in controlling healthcare expenses (Wang, et al., 2005).

Cost-benefit considerations also include the investment in technology and infrastructure needed to integrate and analyze the data from wearable devices. While the initial expenditure may be significant, the return on investment can be realized through the improved health outcomes and efficiencies in care provision. The data gathered from wearables can enable healthcare providers to deliver more personalized care, which is often more cost-effective than generalized treatment approaches (Kang & Exworthy, 2022).

Another consideration is the potential revenue generation through the development and commercialization of new wearable technologies and data analytics platforms. As healthcare continues to evolve, there is a growing market for innovative solutions that can deliver on the promise of improved care at reduced costs. Companies and healthcare systems that can leverage integrated data effectively may be able to develop new business models and revenue streams. However, it is necessary to address the disparities in access to wearable technologies to ensure that the economic benefits of integrated data are not limited to certain populations. There is a risk of exacerbating health inequities if only individuals with higher socioeconomic status can afford wearable devices and thus benefit from data-driven preventive care (Kumar, et al., 2013).

In the end, the economic impact of integrating AHRs with wearable device data is multifaceted. While it offers a strong potential for significant cost savings and efficiencies in healthcare delivery, there is a need for strategic investments and equitable access to technology. Policymakers and healthcare leaders must consider the full spectrum of economic implications as they navigate the integration of these advanced data analytics into mental health care.

7.8 Patient Engagement and Empowerment

The intersection of real-time data from wearables and AHRs presents an unprecedented opportunity for patient empowerment in mental health management. Access to granular, personalized data can transform patients from passive recipients of healthcare into active participants in their treatment and wellness journey. Real-time data equips patients with immediate feedback on their physiological and psychological state, potentially highlighting correlations between lifestyle choices and mental health symptoms. For example, variations in sleep patterns, physical activity, and heart rate variability can be monitored to assess their impact on mood and anxiety levels, offering patients concrete data to support self-care decisions (Faurholt-Jepsen, et al., 2014). Such empowerment through self-monitoring can lead to a greater sense of control over one's mental health, fostering engagement and adherence to treatment plans.

The future of patient engagement is likely to be bolstered by a new generation of tools and platforms that provide integrated health insights. Innovations in app development, for example, are starting to deliver personalized, algorithm-driven advice for managing stress or mood dysregulation. These tools could incorporate cognitive-behavioral therapy techniques, mindfulness practices, and tailored recommendations for activity or social engagement, all informed by the individual's own data (Torous & Keshavan, 2016). Additionally, gamification elements can be embedded into these platforms to incentivize healthy behaviors and adherence to treatment protocols. By offering rewards or achievements for consistency in self-care practices or for reaching personal health milestones, these applications can make the management of mental health both engaging and rewarding, similar to how they are being used now in the physical fitness domains (Fitzpatrick, Darcy, & Vierhile, 2017). However, it's not just about individual data points; it's the integration of these into a cohesive narrative about one's health that can empower patients. Future platforms could utilize ML algorithms to not only present data but also interpret it in an easily understandable manner, providing suggestions and motivational support tailored to the user's specific context and progress (Insel, 2017). This could be achieved by having tailored large language models trained on mental health data that could then evaluate and recommend healthy behaviours based on the data from the wearables.

Telehealth platforms may also evolve, integrating real-time data feeds from wearables to enable clinicians to provide more dynamic and responsive care. This could further strengthen the therapeutic alliance as clinicians and patients would engage in a more collaborative dialogue about the patient's mental health, informed by objective data.

As we look to the future, the emphasis will be on creating a seamless, user-friendly experience that bridges the gap between data collection and actionable insights. Privacy and data security will remain paramount as these tools develop, ensuring that patient empowerment does not come at the cost of data vulnerability (Martinez-Martin, Insel, Dagum, Greely, & Cho, 2018).

7.9 Ethical Considerations and Data Privacy with Wearables

The integration of AHRs with data from wearable technologies in mental health care raises a complex array of ethical considerations, likely more complex than those already mentioned. The real-time, continuous nature of data collection from wearables creates a detailed digital record of an individual's physiological and potentially psychological state.

This rich data set, while invaluable for treatment and monitoring, also increases the risk of privacy breaches and misuse of information, especially when it is not processed locally (Martinez-Martin, Insel, Dagum, Greely, & Cho, 2018).

Concerns center on the potential for unauthorized access to sensitive data, as well as the possible ramifications of data sharing without explicit patient consent. Data may be exploited for purposes that extend beyond patient care, including commercialization or employment discrimination. Moreover, the predictive power of integrated data might lead to stigmatization or other harms if not carefully managed and regulated.

To mitigate these risks, strategies for ensuring data privacy must be embedded at every level of technology design and implementation. This can include employing state-of-the-art encryption methods, secure data storage solutions, and stringent access controls like Zero trust implementation mentioned in the previous section. The principle of data minimization, which posits that only the data necessary for specific care purposes should be collected, can further help protect patient privacy (Layman, 2020).

Healthcare providers and technology developers must work in tandem to ensure compliance with all applicable data protection laws, such as HIPAA, PIPEDA, and the General Data Protection Regulation (GDPR) in Europe. Additionally, organizations should adhere to the guidelines and ethical standards set by professional bodies, such as the American Psychological Association's guidelines for the ethical practice in telepsychology (American Psychological Association, 2023).

The principle of informed consent will also be paramount in the ethical use of wearables in mental health care. Patients should be provided with clear, accessible information about how their data will be used, who will have access to it, and how it will be protected. Moreover, they should have the autonomy to opt-in or opt-out of data sharing with ease, without any impact on their access to care.

The design of wearable technologies and digital health platforms should incorporate 'privacy by design' principles, ensuring that privacy measures are not just an afterthought but a fundamental component of product development. This could include giving users granular control over their data, such as the ability to delete records or control the specific types of data and at when times that is being recorded and shared (Cavoukian, 2018).

Finally, there is a pressing need for interdisciplinary research that examines the ethical implications of wearable technology in mental health, exploring the balance between potential benefits and risks to patient privacy. Such research should guide the development of robust ethical frameworks and regulatory policies to safeguard the well-being and rights of individuals in the context of digital mental health care.

7.10 Future Directions and Research Opportunities

The integration of AHRs with data derived from wearable devices could significantly impact the landscape of mental health research. This integration is expected to provide researchers with an unprecedented volume of longitudinal and real-time data, which can offer novel insights into the precursors and progressions of mental health conditions. As mentioned in previous sections, these insights have the potential to transform mental health care from a reactive to a proactive discipline, with a strong emphasis on early detection and preventive care.

One of the long-term implications of this data integration is the potential to uncover subtle patterns and predictors of mental health issues that are currently obscured in the more episodic and coarse data of AHRs alone. For instance, continuous monitoring of physiological markers, like heart rate variability and sleep patterns, may reveal nuances in the onset of mood disorders or anxiety, potentially leading to early intervention strategies (Torous, Staples, & Onnela, 2015). However, gaps remain in the current research, particularly in the understanding of how best to synthesize and interpret the vast amount of data generated by wearables. There is also a need for more robust, longitudinal studies that link wearable data with clinical outcomes to validate the predictive power of these tools in diverse populations. Another gap lies in the development of standardized protocols for data integration, analysis, and interpretation to ensure consistency and reliability across studies (Gomes, Pato, Lourenço, & Datia, 2023). The potential areas of study to support the growth of this interdisciplinary field are vast. Research is needed to explore the best practices for engaging patients in the use of wearable devices in a way that respects their privacy and autonomy. Additionally, studies on the implications of wearable data on clinical decision-making processes and health outcomes will be critical. Further, there is a need to explore how these technologies can be scaled and made accessible across different socioeconomic groups to avoid exacerbating existing health disparities. In the end, the integration of AHRs with data from wearable devices offers a transformative potential for mental health research and care. By providing a rich, nuanced view of an individual's health in real-time, this intersection of data could facilitate a major shift toward personalized, pre-emptive mental health care strategies. Researchers must continue to advance our understanding of how to effectively analyze and apply wearable data, while healthcare providers should prepare to integrate these insights into clinical practice. Policymakers must develop regulations that protect patients' privacy while enabling innovation. Only through concerted and collaborative efforts can the full vision of a data-driven, patient-centered approach to mental health care be realized.

References

Aitken, M., Jorre, J. d., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Medical Ethics*, https://doi.org/10.1186/s12910-016-0153-x.

Alhanai, T., Ghassemi, M., & Glass, J. (2018). Detecting depression with audio/text sequence modeling of interviews. *Electrical Engineering*, https://doi.org/10.21437/interspeech.2018-2522.

American Psychological Association. (2023). *Guidelines for the Practice of Telepsychology*. Retrieved from https://www.apa.org/practice/guidelines/telepsychology

Beattie, L., Kyle, S. D., Espie, C. A., & Biello, S. M. (2015). Social interactions, emotion and sleep: A systematic review and research agenda. *Sleep Medicine Reviews*, https://doi.org/10.1016/j.smrv.2014.12.005.

Boucsein, W. (2013). *Electrodermal Activity: Second Edition.* Springer.

Cavoukian, A. (2018). *Privacy by design: The 7 foundational principles.* Retrieved from https://privacy.ucsc.edu/resources/privacy-by-design---foundational-principles.pdf

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp1714229.

Clingan, C. A., Dittakavi, M., Rozwadowski, M., Gilley, K. N., Cislo, C., Barabas, J., . . . Tewari, M. (2021). Monitoring Health Care Workers at Risk for COVID-19 Using Wearable Sensors and Smartphone Technology: Protocol for an Observational mHealth Study. *JMIR Publications*, https://preprints.jmir.org/preprint/29562.

Faurholt-Jepsen, M., Frost, M., Vinberg, M., Christensen, E. M., Bardram, J. E., & Kessing, L. V. (2014). Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research*, https://doi.org/10.1016/j.psychres.2014.03.009.

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, https://doi.org/10.2196/mental.7785.

Gomes, N., Pato, M., Lourenço, A. R., & Datia, N. (2023). A Survey on Wearable Sensors for Mental Health Monitoring. *Sensors* , https://doi.org/10.3390/s23031330.

Insel, T. R. (2017). Digital Phenotyping: Technology for a New Science of Behavior. *JAMA*, https://doi.org/10.1001/jama.2017.11295.

Kang, H. S., & Exworthy, M. (2022). Wearing the Future—Wearables to Empower Users to Take Greater Responsibility for Their Health and Care: Scoping Review. *JMIR mHealth and uHealth*, https://doi.org/10.2196/35684.

Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2014). Dynamic consent: a patient interface for twenty-first century research networks. *European Journal of Human Genetics*, https://doi.org/10.1038/ejhg.2014.71.

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., . . . Swendeman. (2013). Mobile health technology evaluation: the mHealth evidence workshop. *American Journal of Preventive Medicine*, https://doi.org/10.1016/j.amepre.2013.03.017.

Layman, E. J. (2020). Ethical Issues and the Electronic Health Record. *The Health Care Manager*, https://doi.org/10.1097/hcm.0000000000000302.

Luxton, D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice*, https://doi.org/10.1037/a0034559.

Martinez-Martin, N., Insel, T. R., Dagum, P., Greely, H. T., & Cho, M. K. (2018). Data mining for health: staking out the ethical territory of digital phenotyping. *npj Digital Medicine*, https://doi.org/10.1038/s41746-018-0075-8.

Microsoft. (2023). *Embrace proactive security with Zero Trust*. Retrieved from Microsoft Security: https://www.microsoft.com/en-ca/security/business/zero-trust#:~:text=What%20is%20Zero%20Trust%3F,originates%20from%20an%20open%20network.

Monteith, S., Glenn, T., Geddes, J., Whybrow, P. C., & Bauer, M. (2016). Big data for bipolar disorder. *Int J Bipolar Disord*, https://doi.org/10.1186/s40345-016-0051-7.

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2017). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, https://doi.org/10.1007/s12160-016-9830-8.

National Academies of Sciences, Engineering, and Medicine. (2019). Integrating Health Care and Social Services for People with Serious Illness: Proceedings of a Workshop. *National Academies Press*, https://doi.org/10.17226/25350.

Office of the Privacy Commissioner of Canada. (2020). *Office of the Privacy Commissioner of Canada*. Retrieved from https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/pos_ai_202001

Rhee, S. Y., Kim, C., Shin, D. W., & Stein, S. R. (2020). Present and Future of Digital Health in Diabetes and Metabolic Disease. *Diabetes & Metabolism Journal*, https://doi.org/10.4093/dmj.2020.0088.

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers Public Health*, https://doi.org/10.3389/fpubh.2017.00258.

Stephens, T., & Joubert, N. (2001). The Economic Burden of Mental Health Problems in Canada. *Chronic Diseases in Canada*, 18-23.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

Torous, J., & Baker, J. T. (2016). Why Psychiatry Needs Data Science and Data Science Needs Psychiatry. *JAMA Psychiatry*, https://doi.org/10.1001/jamapsychiatry.2015.2622.

Torous, J., & Keshavan, M. (2016). The role of social media in schizophrenia: evaluating risks, benefits, and potential. *Current Opinion in Psychiatry*, https://doi.org/10.1097/yco.0000000000000246.

Torous, J., Staples, P., & Onnela, J.-P. (2015). Realizing the potential of mobile mental health: new methods for new data in psychiatry. *Current Psychiatry Reports*, https://doi.org/10.1007/s11920-015-0602-0.

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, https://doi.org/10.1371/journal.pmed.1002689.

Wang, P. S., Berglund, P., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, https://doi.org/10.1001/archpsyc.62.6.603.

Wulff, K., Gatti, S., Wettstein, J. G., & Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nature Reviews Neuroscience*, https://doi.org/10.1038/nrn2868.

Chapter 8
Conclusion

This thesis embarked on a comprehensive exploration of the applications and untapped potential of AHRs in augmenting the landscape of mental health care. Beginning with a contextual look at the historical challenges facing psychiatry and public health, we scrutinized the role and limitations of diagnostic mechanisms like the DSM and ICD in mental health. Given the current grim landscape of morbidity, mortality, and economic impact associated with mental health, the research identified an immediate need for innovative approaches.

ML emerged as a new frontier in psychiatry, opening the door for a convergence of public health data, precision health, and computational methods. This synthesis has paved the way for more complex, data-driven strategies in mental health care, particularly important amid the increasing issues triggered by various disasters. Early chapters navigated the complex terrain of privacy concerns and emphasized the need for ethical and secure handling of sensitive data. Collaborative research endeavors were highlighted, emphasizing the collaboration between governmental institutions, clinicians, and academia in the ethical utilization of AHRs.

The core of this work featured a series of case studies illustrating the transformative potential of AHRs in current mental health research. From crafting predictive models for individual-level opioid overdoses to uncovering shifts in developmental disorder treated prevalence and illuminating healthcare utilization patterns for neurocognitive disorders, the power of AHRs was examined. The impact was particularly noticeable in the unprecedented healthcare challenges induced by the COVID-19 pandemic.

Towards the end, the research ventured into the realm of future applications. AHRs were positioned as pivotal tools for forecasting mental health outcomes in post-disaster scenarios. Here, the intricate relationships between disaster-induced stress, trauma, and observable mental health patterns were dissected. The thesis also envisioned a future where

AHRs are seamlessly integrated with wearable technology. Such integration could provide real-time, granular insights into mental health, fundamentally transforming our predictive capabilities from being reactionary—based on reported symptoms or delayed administrative data—to proactive and preventive.

To sum up, this thesis not only provided a detailed account of the current state and inherent challenges of mental health care but also laid out a roadmap for future, transformative practices. By combining cutting-edge technology with traditional research methods, this research stands as a small milestone toward a new paradigm in mental health care—one that is proactive, data-driven, and deeply rooted in interdisciplinary collaboration. The coupling of AHRs with other real-time data sources like wearables, fortified by ML algorithms, offers a promising avenue for early intervention and more individualized mental health care strategies. Through these multi-faceted approaches and integrative strategies, the dream of a more informed, proactive, and humane approach to mental health is not only plausible but within reach.

Bibliography

Aitken, M., Jorre, J. d., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Medical Ethics*, https://doi.org/10.1186/s12910-016-0153-x.

Alhanai, T., Ghassemi, M., & Glass, J. (2018). Detecting depression with audio/text sequence modeling of interviews. *Electrical Engineering*, https://doi.org/10.21437/interspeech.2018-2522.

Alpert, S. A. (2003). Protecting Medical Privacy: Challenges in the Age of Genetic Information. *Journal of Social Issues*, https://doi.org/10.1111/1540-4560.00066.

American Psychiatric Association . (1952). *Diagnostic and Statistical Manual: Mental Disorders.*

American Psychiatric Association . (2017). *DSM 5 Diagnostic And Statistical Manual Of Mental Disorders 5 Ed.* Washington, DC: American Psychiatric Publishing.

American Psychological Association. (2023). *Guidelines for the Practice of Telepsychology*. Retrieved from https://www.apa.org/practice/guidelines/telepsychology

Ashley, E. A. (2016). Towards precision medicine. *Nature Reviews Genetics*, 17, 507–522.

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*, https://doi.org/10.1377/hlthaff.2014.0041.

Beattie, L., Kyle, S. D., Espie, C. A., & Biello, S. M. (2015). Social interactions, emotion and sleep: A systematic review and research agenda. *Sleep Medicine Reviews*, https://doi.org/10.1016/j.smrv.2014.12.005.

Benitez, K., & Malin, B. (2010). Evaluating re-identification risks with respect to the HIPAA privacy rule. *Journal of the American Medical Informatics Association*, https://doi.org/10.1136/jamia.2009.000026.

Boucsein, W. (2013). *Electrodermal Activity: Second Edition.* Springer.

Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Cognitive Neuroscience and Neuroimaging*, https://doi.org/10.1016/j.bpsc.2017.11.007.

Calhoun, V. D., Sui, J., Kiehl, K., Turner, J., Allen, E., & Pearlson, G. (2012). Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder. *Frontiers of Psychiatry*, https://doi.org/10.3389/fpsyt.2011.00075.

Canadian Institute for Health Information. (2023). *Canadian Institute for Health Information*. Retrieved from Codes and Classifications: https://www.cihi.ca/en/submit-data-and-view-standards/codes-and-classifications

CanLII. (2018). *Facebook agrees to stop using non-users' personal information in users' address books*. Retrieved from https://canlii.ca/t/hvs0v

Cartus, A. R., Samuels, E. A., Cerdá, M., & Marshall, B. D. (2023). Outcome class imbalance and rare events: An underappreciated complication for overdose risk prediction modeling. *Addiction* , https://doi.org/10.1111/add.16133.

Cavoukian, A. (2018). *Privacy by design: The 7 foundational principles.* Retrieved from https://privacy.ucsc.edu/resources/privacy-by-design---foundational-principles.pdf

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. Retrieved from https://www.semanticscholar.org/paper/CRISP-DM-1.0%3A-Step-by-step-data-mining-guide-Chapman-Clinton/54bad20bbc7938991bf34f86dde0babfbd2d5a72

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp1714229.

Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., . . . Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(15)00471-x.

Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, https://doi.org/10.1038/s41551-021-00751-8.

Chevrier, R., Foufi, V., Gaudet-Blavignac, C., Robert, A., & Lovis, C. (2019). Use and Understanding of Anonymization and De-Identification in the Biomedical Literature: Scoping Review. *Journal of Medical Internet Research*, https://doi.org/10.2196/13484.

Choi, E., Biswal, S., Malin, B., Duke, J., Stewart, W. F., & Sun, J. (2018). Generating Multi-label Discrete Patient Records using Generative Adversarial Networks (Version 3). *arXiv*, https://doi.org/10.48550/ARXIV.1703.06490.

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three Approaches to Understanding and Classifying Mental Disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest*, 72-145. https://doi.org/10.1177/1529100617727266.

Clingan, C. A., Dittakavi, M., Rozwadowski, M., Gilley, K. N., Cislo, C., Barabas, J., . . . Tewari, M. (2021). Monitoring Health Care Workers at Risk for COVID-19 Using Wearable Sensors and Smartphone Technology: Protocol for an Observational mHealth Study. *JMIR Publications*, https://preprints.jmir.org/preprint/29562.

Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying Mental Health Signals in Twitter. *Workshop on Computational Linguistics and Clinical Psychology*, https://doi.org/10.3115/v1/w14-3207.

Dauphinot, V., Ravier, A., Novais, T., Delphin-Combe, F., Moutet, C., Xie, J., . . . Krolak-Salmon, P. (2016). Relationship Between Comorbidities in Patients With Cognitive Complaint and Caregiver Burden: A Cross-Sectional Study. *Journal of the American Medical Directors Association*, https://doi.org/10.1016/j.jamda.2015.10.011.

David W Bates 1, S. S.-M. (2014). Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*, https://doi.org/10.1377/hlthaff.2014.0041.

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, https://doi.org/10.1146/annurev-clinpsy-032816-045037.

El Emam, K., & Dankar, F. K. (2008). Protecting Privacy Using k-Anonymity. *Journal of the American Medical Informatics Association*, https://doi.org/10.1197/jamia.m2716.

Ellis, R. J., Wang, Z., Genes, N., & Ma'ayan, A. (2019). Predicting opioid dependence from electronic health records with machine learning. *BioData Mining*, https://doi.org/10.1186/s13040-019-0193-0.

Emam, K. E., Arbuckle, L., Koru, G., Eze, B., Gaudette, L., Neri, E., . . . Gluck, J. (2012). De-identification Methods for Open Health Data: The Case of the Heritage Health Prize Claims Dataset. *Journal of Medical Internet Research*, https://doi.org/10.2196/jmir.2001.

Emam, K. E., Moura, P., Locton, V., Jonker, E., & Kardash, A. (2021). Practices for Generating Non-identifiable Data: FINAL REPORT . *OFFICE OF THE PRIVACY COMMISSIONER OF CANADA CONTRIBUTIONS PROGRAM* , https://www.priv.gc.ca/en/opc-actions-and-decisions/research/funding-for-privacy-research-and-knowledge-translation/completed-contributions-program-projects/2020-2021/p_2020-21_10/.

Faurholt-Jepsen, M., Frost, M., Vinberg, M., Christensen, E. M., Bardram, J. E., & Kessing, L. V. (2014). Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research*, https://doi.org/10.1016/j.psychres.2014.03.009.

Fazel, M., Hoagwood, K., Stephan, S., & Ford, T. (2014). Mental health interventions in schools 1: Mental health interventions in schools in high-income countries. *Lancet Psychiatry*, 377-387. https://doi.org/10.1016/s2215-0366(14)70312-8.

Felix, E., Afifi, T., Kia-Keating, M., Brown, L., Afifi, W., & Reyes, G. (2015). Family functioning and posttraumatic growth among parents and youth following wildfire disasters. *American Journal of Orthopsychiatry*, https://doi.org/10.1037/ort0000054.

Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F., & Berk, M. (2017). The new field of 'precision psychiatry'. *BMC Medicine*, https://doi.org/10.1186/s12916-017-0849-x.

First, M. B., Reed, G. M., Hyman, S. E., & Saxena, S. (2015). The development of the ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. *World Psychiatry*, https://doi.org/10.1002/wps.20189.

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, https://doi.org/10.2196/mental.7785.

Friston, K. J., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(14)70275-5.

Fussell, E., & Lowe, S. R. (2014). The impact of housing displacement on the mental health of low-income parents after Hurricane Katrina. *Social Science & Medicine*, https://doi.org/10.1016/j.socscimed.2014.05.025.

Galea, S., Merchant, R. M., & Lurie, N. (2020). The Mental Health Consequences of COVID-19 and Physical Distancing. *JAMA Internal Medicine*, https://doi.org/10.1001/jamainternmed.2020.1562.

Garies, S., Ng, P., Dickinson, J. A., McDonald, T., O'Beirne, M., McBrien, K. A., . . . Quan, H. (2022). Leaving the Walkman and ICD-9 Behind: Modernizing the Disease Classification System Used by Canadian Physicians. *Health Policy*, https://doi.org/10.12927/hcpol.2022.26907.

Giuffrè, M., & Shung, D. L. (2023). Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. *npj Digital Medicine* , https://doi.org/10.1038/s41746-023-00927-3.

Global Newswire. (2021). *Alberta Project Explores Use of Synthetic Data as Alternative to Health Data Utilization*.

Goldmann, E., & Galea, S. (2014). Mental Health Consequences of Disasters. *Annual Review of Public Health*, https://doi.org/10.1146/annurev-publhealth-032013-182435.

Gomes, N., Pato, M., Lourenço, A. R., & Datia, N. (2023). A Survey on Wearable Sensors for Mental Health Monitoring. *Sensors* , https://doi.org/10.3390/s23031330.

Gonzales, A., Guruswamy, G., & Smith, S. R. (2023). Synthetic data in health care: A narrative review. *PLOS Digital Health*, https://doi.org/10.1371/journal.pdig.0000082.

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *Journal of Clinical Psychiatry*, https://doi.org/10.4088/jcp.14m09298.

Hayes, J., Melis, L., Danezis, G., & De Cristofaro, E. (2017). LOGAN: Membership Inference Attacks Against Generative Models. *arXiv*, https://doi.org/10.48550/ARXIV.1705.07663.

Health Canada. (2019). *Guidance document on Public Release of Clinical Information*. Retrieved from https://www.canada.ca/en/health-canada/services/drug-health-product-review-approval/profile-public-release-clinical-information-guidance.html

Hellis, E., & Mukaetova-Ladinska, E. B. (2022). Informal Caregiving and Alzheimer's Disease: The Psychological Effect. *Medicina* , https://doi.org/10.3390/medicina59010048.

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., . . . Madan, I. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(20)30168-1.

Horgan, D., Paradiso, A., McVie, G., Banks, I., Wal, T. V., Brand, A., & Lawler, M. (2015). Is precision medicine the route to a healthy world? *The Lancet*, https://doi.org/10.1016/s0140-6736(15)61404-0.

Information and Privacy Commissioner of Ontario. (2016). *De-identification Guidelines for Structured Data*. Retrieved from https://www.ipc.on.ca/resource/de-identification-guidelines-for-structured-data/

Insel, T. R. (2017). Digital Phenotyping: Technology for a New Science of Behavior. *JAMA*, https://doi.org/10.1001/jama.2017.11295.

International Organization for Standardization. (2023). *Health informatics – Digital health – Standards for global use*.

J Torous, J.-P. O. (2017). New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. *Translational Psychiatry*, https://doi.org/10.1038/tp.2017.25.

Jr, R. C., & Clayton, P. D. (1996). Privacy, confidentiality, and electronic medical records. *Journal of the American Medical Informatics Association*, https://doi.org/10.1136/jamia.1996.96236282.

Kang, H. S., & Exworthy, M. (2022). Wearing the Future—Wearables to Empower Users to Take Greater Responsibility for Their Health and Care: Scoping Review. *JMIR mHealth and uHealth*, https://doi.org/10.2196/35684.

Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2014). Dynamic consent: a patient interface for twenty-first century research networks. *European Journal of Human Genetics*, https://doi.org/10.1038/ejhg.2014.71.

Kendler, K. S. (2016). The Phenomenology of Major Depression and the Representativeness and Nature of DSM Criteria. *The American Journal of Psychiatry*, 771-780. https://doi.org/10.1176/appi.ajp.2016.15121509.

Kessler, R. C., Galea, S., Gruber, M. J., Sampson, N. A., Ursano, R. J., & Wessely, S. (2008). Trends in mental illness and suicidality after Hurricane Katrina. *Molecular Psychiatry*, https://doi.org/10.1038/sj.mp.4002119.

Kisely, S., Lin, E., Lesage, A., Gilbert, C., Smith, M., Campbell, L. A., & Vasiliadis, H.-M. (2009). Use of administrative data for the surveillance of mental disorders in 5 provinces. *Canadian Journal of Psychiatry*, https://doi.org/10.1177/070674370905400810.

Knapp, M., McDaid, D., & Parsonage, M. (2011). *Mentalhealthpromotion andmentalillness prevention: The economic case.* London: Department of Health.

Koutsouleris, N., Kahn, R. S., Chekroud, A. M., Leucht, S., Falkai, P., Wobrock, T., . . . Hasan, A. (2016). Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry*, https://doi.org/10.1016/s2215-0366(16)30171-7.

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., . . . Swendeman. (2013). Mobile health technology evaluation: the mHealth evidence workshop. *American Journal of Preventive Medicine*, https://doi.org/10.1016/j.amepre.2013.03.017.

Kuo, T.-T., Kim, H.-E., & Ohno-Machado, L. (2017). Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association*, https://doi.org/10.1093/jamia/ocx068.

Layman, E. J. (2020). Ethical Issues and the Electronic Health Record. *The Health Care Manager*, https://doi.org/10.1097/hcm.0000000000000302.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., . . . Kales, H. (2020). Dementia prevention, intervention, and care. *The Lancet Commission*, https://doi.org/10.1016/s0140-6736(20)30367-6.

Loades, M. E., Chatburn, E., Higson-Sweeney, N., Reynolds, S., Shafran, R., Brigden, A., . . . Crawley, E. (2020). Rapid Systematic Review: The Impact of Social Isolation and Loneliness on the Mental Health of Children and Adolescents in the Context of COVID-19. *Journal of the American Academy of Child & Adolescent Psychiatry*, https://doi.org/10.1016/j.jaac.2020.05.009.

Luxton, D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice*, https://doi.org/10.1037/a0034559.

MacEachern, S. J., & Forkert, N. D. (2021). Machinelearningforprecisionmedicine. *Genome* , https://doi.org/10.1139/gen-2020-0131.

Martinez-Martin, N., Insel, T. R., Dagum, P., Greely, H. T., & Cho, M. K. (2018). Data mining for health: staking out the ethical territory of digital phenotyping. *npj Digital Medicine*, https://doi.org/10.1038/s41746-018-0075-8.

Mathers, C. D., & Loncar, D. (2006). Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLOS Medicine*, https://doi.org/10.1371/journal.pmed.0030442.

Microsoft. (2023). *Embrace proactive security with Zero Trust*. Retrieved from Microsoft Security: https://www.microsoft.com/en-ca/security/business/zero-trust#:~:text=What%20is%20Zero%20Trust%3F,originates%20from%20an%20open%20network.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Science*, https://doi.org/10.1016/j.tics.2011.11.018.

Monteith, S., Glenn, T., Geddes, J., Whybrow, P. C., & Bauer, M. (2016). Big data for bipolar disorder. *Int J Bipolar Disord*, https://doi.org/10.1186/s40345-016-0051-7.

Mosquera, L., Emam, K. E., Ding, L., Sharma, V., Zhang, X. H., Kababji, S. E., . . . Eurich, D. T. (2023). A method for generating synthetic longitudinal health data. *BMC Medical Research Methodology*, https://doi.org/10.1186/s12874-023-01869-w.

Murphy, G. H., Beadle-Brown, J., Wing, L., Gould, J., Shah, A., & Holmes, N. (2005). Chronicity of Challenging Behaviours in People with Severe. *Journal of Autism and Developmental Disorders*, https://doi.org/10.1007/s10803-005-5030-2.

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2017). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, https://doi.org/10.1007/s12160-016-9830-8.

Nathan M. Stall, A. J. (2020). For-profit long-term care homes and the risk of COVID-19 outbreaks and resident deaths. *Canadian Medical Association Journal*, https://doi.org/10.1503/cmaj.201197.

National Academies of Sciences, Engineering, and Medicine. (2019). Integrating Health Care and Social Services for People with Serious Illness: Proceedings of a Workshop. *National Academies Press*, https://doi.org/10.17226/25350.

National Institute on Drug Abuse. (2022). *Drug Overdose Death Rates*. Retrieved from National Institutes of Health: https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates

Neria, Y., Nandi, A., & Galea, S. (2007). Post-traumatic stress disorder following disasters: a systematic review. *Psychological Medicine*, https://doi.org/10.1017/s0033291707001353.

Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry: Interpersonal and Biological Processes*, https://doi.org/10.1521/psyc.65.3.207.20173.

North, C. S., & Pfefferbaum, B. (2013). Mental health response to community disasters: a systematic review. *JAMA*, https://doi.org/10.1001/jama.2013.107799.

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp1606181.

Office for Civil Rights. (2022). *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. Retrieved from U.S. Department of Health and Human Services: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Office of the Privacy Commissioner of Canada. (2020). *Consultation on the OPC's Proposals for ensuring appropriate regulation of artificial intelligence*. Retrieved from https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/pos_ai_202001/ (accessed 10 May 2022).

Office of the Privacy Commissioner of Canada. (2020). *Office of the Privacy Commissioner of Canada*. Retrieved from https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/pos_ai_202001

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., . . . Jordans, M. J. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet*, https://doi.org/10.1016/S0140-6736(18)31612-X.

Paton, D., & Johnston, D. (2001). Disasters and communities: vulnerability, resilience and preparedness. *Disaster Prevention and Management: An International Journal*, https://doi.org/10.1108/eum0000000005930.

Pfefferbaum, B., & North, C. S. (2020). Mental Health and the Covid-19 Pandemic. *New England Journal of Medicine*, https://doi.org/10.1056/nejmp2008017.

Price, N. (2021). Problematic Interactions Between AI and Health Privacy Problematic Interactions Between AI and Health Privacy. *Utah Law Review Utah Law Review* , https://ssrn.com/abstract=3797161.

Prince, M. J., Wu, F., Guo, Y., Robledo, L. M., O'Donnell, M., Sullivan, R., & Yusuf, S. (2015). The burden of disease in older people and implications for health policy and practice. *The Lancet*, https://doi.org/10.1016/s0140-6736(14)61347-7.

Public Health Agency of Canada. (2023). *Opioid- and Stimulant-related Harms in Canada. Ottawa: Public Health Agency of Canada*. Retrieved from Federal, provincial, and territorial Special Advisory Committee on the Epidemic of Opioid Overdoses: https://health-infobase.canada.ca/substance-related-harms/opioids-stimulants/

Purnell, A. (2020). *How synthetic data will improve Veteran health and care*. Retrieved from VA News: https://news.va.gov/81908/synthetic-data-improve-veteran-care/

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, https://doi.org/10.1056/nejmra1814259.

Reifels, L., Pietrantoni, L., Prati, G., Kim, Y., Kilpatrick, D. G., Dyb, G., . . . O'Donnell, M. (2013). Lessons learned about psychosocial responses to disaster and mass trauma: an international perspective. *European Journal of Psychotraumatology*, https://doi.org/10.3402/ejpt.v4i0.22897.

Rhee, S. Y., Kim, C., Shin, D. W., & Stein, S. R. (2020). Present and Future of Digital Health in Diabetes and Metabolic Disease. *Diabetes & Metabolism Journal*, https://doi.org/10.4093/dmj.2020.0088.

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H., Albarqouni, S., . . . Cardoso, M. J. (2020). The future of digital health with federated learning. *npj Digital Medicine*, https://doi.org/10.1038/s41746-020-00323-1.

SAS. (2019). *Exploring SAS Viya.* Retrieved from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://support.sas.com/content/dam/SAS/support/en/books/free-books/exploring-sas-viya-data-mining-machine-learning.pdf

Schell, R. C., Allen, B., Goedel, W. C., Hallowell, B. D., Scagos, R., Li, Y., . . . Ahern, J. (2021). Identifying Predictors of Opioid Overdose Death at a Neighborhood Level With Machine Learning. *American Journal of Epidemiology*, https://doi.org/10.1093/aje/kwab279.

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., . . . Silver, D. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* , https://doi.org/10.1038/s41586-020-03051-4.

Service, C. R. (2023). *The Office of the National Coordinator for Health Information Technology (ONC)*. Retrieved from https://crsreports.congress.gov

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers Public Health*, https://doi.org/10.3389/fpubh.2017.00258.

Shorter, E. (1997). *A history of psychiatry: From the era of the asylum to the age of Prozac.* John Wiley & Sons.

Shultz, J. M., Thoresen, S., Flynn, B. W., Muschertl, G. W., Shaw, J. A., Espinel, Z., . . . Cohen, A. M. (2014). Multiple vantage points on the mental health effects of mass shootings. *Current Psychiatry Reports*, https://doi.org/10.1007/s11920-014-0469-5.

Smoller, J. W. (2015). The Genetics of Stress-Related Disorders: PTSD, Depression, and Anxiety Disorders. *Neuropsychopharmacology* , https://doi.org/10.1038/npp.2015.266.

Statistics Canada. (2021). *Deaths and age-specific mortality rates, by selected grouped causes*. Retrieved from https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310039201

Stephens, T., & Joubert, N. (2001). The Economic Burden of Mental Health Problems in Canada. *Chronic Diseases in Canada*, 18-23.

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., . . . Bunting, B. (2017). Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry*, https://doi.org/10.1192/bjp.bp.116.188078.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

Torous, J., & Baker, J. T. (2016). Why Psychiatry Needs Data Science and Data Science Needs Psychiatry. *JAMA Psychiatry*, https://doi.org/10.1001/jamapsychiatry.2015.2622.

Torous, J., & Keshavan, M. (2016). The role of social media in schizophrenia: evaluating risks, benefits, and potential. *Current Opinion in Psychiatry*, https://doi.org/10.1097/yco.0000000000000246.

Torous, J., Staples, P., & Onnela, J.-P. (2015). Realizing the potential of mobile mental health: new methods for new data in psychiatry. *Current Psychiatry Reports*, https://doi.org/10.1007/s11920-015-0602-0.

Treasury Board of Canada. (2020). *Privacy Implementation Notice 2020-03: Protecting privacy when releasing information about a small number of individuals*. Retrieved from https://www.canada.ca/en/treasury-board-secretariat/services/access-information-privacy/access-information-privacy-notices/2020-03-protecting-privacy-releasing-information-about-small-number-individuals.html

Tseregounis, I. E., & Henry, S. G. (2021). Assessing opioid overdose risk: a review of clinical prediction models utilizing patient-level data. *Translational Research*, https://doi.org/10.1016/J.TRSL.2021.03.012.

United Nations Office for Disaster Risk Reduction. (2022). *Global Assessment Report on Disaster Risk Reduction.* United Nations.

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, https://doi.org/10.1371/journal.pmed.1002689.

Wang, P. S., Berglund, P., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, https://doi.org/10.1001/archpsyc.62.6.603.

Wang, W., & Krishnan, E. (2014). Big data and clinicians: a review on the state of the science. *JMIR Medical Informatics*, https://doi.org/10.2196/medinform.2913.

Wilkinson, K., Green, C., Nowicki, D., & Von Schindler, C. (2020). Less than five is less than ideal: replacing the "less than 5 cell size" rule with a risk-based data disclosure protocol in a public health setting. *Canadian Journal of Public Health*, https://doi.org/10.17269/s41997-020-00303-8.

Wind, T. R., Rijkeboer, M., Andersson, G., & Riper, H. (2020). The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. *Internet Interventions*, https://doi.org/10.1016/j.invent.2020.100317.

World Health Organization. (1948). *International Classification of Diseases.*

World Health Organization. (2013). *Building back better: sustainable mental health care after emergencies.* World Health Organization.

World Health Organization. (2018). *International Classification of Diseases (11th ed.).* Geneva, Switzerland: WHO.

World Health Organization. (2019). *International Statistical Classification of Diseases and Related Health Problems (ICD)*. Retrieved from Classication of Diseases: https://www.who.int/standards/classifications/classification-of-diseases

World Health Organization. (2020). *Mental health and psychosocial considerations during the COVID-19 outbreak*. Retrieved from World Health Organization: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf

World Health Organization. (2022). *Mental health in emergencies*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/mental-health-in-emergencies

Wulff, K., Gatti, S., Wettstein, J. G., & Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nature Reviews Neuroscience*, https://doi.org/10.1038/nrn2868.

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. Annual Review of Clinical Psychology. *Annual Review of Clinical Psychology*, 49-71. https://doi.org/10.1146/annurev-clinpsy-032816-045020.

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensice Psychiatry*, https://doi.org/10.1016/j.comppsych.2014.09.007.