BEING FOUND NOT CRIMINALLY RESPONSIBLE ON ACCOUNT OF MENTAL DISORDER IN ALBERTA: POPULATION GENDER DIFFERENCES, AND VIOLENCE PREDICTION WITH THE VRAG-R

A Dissertation Submitted to the College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

In the Department of Psychology

University of Saskatchewan

Saskatoon, SK

By

Robi Lynn Wirove

PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my dissertation work or, in their absence, by the Head of the Department or the Dean of the College of Graduate and Postdoctoral Studies in which my thesis work was done. It is understood that any copying or publication or use of this dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my dissertation.

DISCLAIMER

Reference in this dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this dissertation in whole or part should be addressed to:

Head of the Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A5
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

Objective: This study had two main goals. The first was to examine those individuals who had been found Not Criminally Responsible on Account of Mental Disorder (NCRMD) in Alberta, Canada for gender differences. The second was to examine the predictive accuracy of the Violence Risk Appraisal Guide Revised (VRAG-R) within this population, specifically the discrimination and calibration properties. **Method**: The study was archival and retrospective in nature. There were 574 individuals identified via *The Alberta NCR Project* database for inclusion. Information was taken from various sources including hospital and disposition records. A VRAG-R was scored on every file that contained appropriate information. **Results**: Gender differences were identified that indicated a unique sociodemographic, clinical, and criminological profile for both genders. The VRAG-R demonstrated strong discrimination properties for both total score and bin number for both general and violent recidivism over 5-year, 10-year and global follow up. The calibration properties however indicated that the VRAG-R substantially over estimated risk and that there was poor agreement between expected and observed recidivism rates for the overall population. When examined by gender, these issues remained but to a lesser degree. Examination of both discrimination and calibration for females was not possible. **Discussion**: Results supported previous research indicating that unique sociodemographic, clinical, and criminological profiles exist for the genders. Results indicated strong discrimination and poor calibration properties of the VRAG-R. Though caution should be noted in interpreting these results for a variety of reasons. Overall, the results support the use of the VRAG-R within a population of persons found NCRMD when embedded within a comprehensive risk assessment battery of tools.

ACKNOWLEDGEMENTS

I would like to start by sincerely thanking three incredible humans who have shaped not only the course of my career but also my life. First, Dr. Andrew Haag, thank you for inspiring in me the faith and courage necessary to chase my dream of becoming a psychologist. Also, for all those great existential chats during data collection. Second, Dr. Stephen Wormith, thank you for taking a chance on me and opening the door to my dreams. I'll miss sitting in your office listening to your stories. Third, Dr. Mark Olver, thank you for showing me the way, your guidance and support, your unwavering belief in me (even through all my stats anxiety) and also for the hilarious role plays in first year.

I would also like to extend my sincere thanks to my committee members, Dr. Brian Chartier, Dr. Keira Stockdale and Professor Glen Luther. Your wisdom and guidance were invaluable to this project. Furthermore, my sincere thanks to my external examiner, Dr. Daryl Kroner, your thoughtful feedback and insight was exactly what was needed in order to complete this project.

To my cohort, thanks for being the incredible women that you are. I learned something from each of you. To Gillian and Kelsey especially, thank you for the "debriefs", the exhausted hilarity and all the Tim Horton/Starbucks runs. You made grad school fun and living so far away from home easy.

I have the most incredibly supportive family. Thank you to each of you. You made it possible for me to chase my dreams by giving me a strong foundation and deep roots. To the Momma W., without you, none of this would have been possible. From day one you have been there to help make my dreams a reality. To Justin and Taryn, and Mat and Shauna, thanks for keeping me humble, picking me up off the floor when things went poorly and for all the D&D

nights. To my dad, I know you're watching and I know you're proud. I've felt you walk this path with me the entire time.

To the two loves of my life, Charles (my cat) and Brendan (my husband). Charles, thanks for always being there to keep me grounded, reminding me that the world doesn't revolve around grad school (it revolves around your feeding schedule) and for being something soft for when the world got hard. To Brendan, my one and only, my soulmate, and my best friend, thanks for believing in me, for all the "you got this babe"' pep talks and never letting me give up. You trusted in me and believed in my dream. Just remember, I love you more.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

TABLE OF ABBREVIATIONS

Abbreviation	Explanation
AD	Absolute Discharge
ARB	Alberta Review Board
ASPD	Antisocial Personality Disorder
AUC	Area Under the Curve
CD	Conditional Discharge
DA	Differential Association
DSM - 5	Diagnostic and Statistical Manual of Mental Disorders – 5th edition
FACS	Forensic Assessment Community Services
HRA	High Risk Accused
GPCSL	General Personality and Cognitive Social Learning
NCRMD or NCR	Not Criminally Responsible on Account of Mental Disorder
NGRI	Not Guilty by Reason of Insanity
NTP	National Trajectory Project
PCC	Psychology of Criminal Conduct
RA	Research Assistant
ROC	Receiver Operator Characteristics
SJP	Structured Professional Judgement
SUD	Substance Use Disorder
RB	Review Board
VRAG	Violence Risk Appraisal Guide
VRAG-R	Violence Risk Appraisal Guide - Revised
UST	Unfit to Stand Trial

Note: All abbreviations listed also have explanations included in text as they appear. The list includes alphabetized abbreviations that occur more than once in the text

CHAPTER 1: LITERATURE REVIEW

1.1 General Overview

The following is a review of the relevant research exploring persons found Not Criminally Responsible on Account of Mental Disorder (NCRMD) in Canada, and their intersections with gender, and violence risk assessment using the Violence Risk Appraisal Guide Revised (VRAG-R; Rice, Grant, Harris, Lang, 2013). This review is organized into various sections. It begins by examining perspectives on criminal responsibility, mental illness, and what it means to be found NCRMD in Canada and around the world. It follows with relevant history and rates of NCRMD from across Canada. Next, a clarification between the legal status of Unfit to Stand Trial (UST) and being found NCRMD is made. After that, in attempt to highlight the complex and often costly needs of those persons found NCRMD, what is known about gender, violence, and recidivism within this population is explored. At times throughout this review, comparison with non-mentally disordered offenders is used to highlight their differences and the unique needs of this population. Special considerations such as gendered pathways to crime, and forensic risk assessment theory and history are also explored. The review ends by exploring what is known about the VRAG-R. The NCRMD population is one that has received relatively little attention in research until very recently and what is known is highly dependent upon the jurisdiction that is studied. Similarly, the VRAG-R has only recently been published and requires independent validation. It is hoped that the following study will add knowledge to these areas.

1.2 Introduction to Not Criminally Responsible on Account of Mental Disorder (NCRMD)

Determining criminal responsibility for a person who commits a crime while suffering from a mental disorder varies by country (Grossi & Green, 2017; Allnutt, Samuels, & O'driscoll, 2007; Bal & Koenraadt, 2000). For example, one country may view criminal responsibility as a

dichotomous factor (e.g., in some states within the United States one is either criminally responsible or not criminally responsible), while another country may view it as a continuous factor (e.g., in the Netherlands a graded 5-point system is used to describe criminal responsibility), while a third country may not regard it as a factor at all (e.g., in Sweden mentally ill offenders are still, in principle, considered responsible) (Grossi & Green, 2017; Bal & Koenraadt, 2000). In Canada, criminal responsibility is considered a dichotomous factor (i.e., one is either criminally responsible or not) and requires two essential elements: (a) the *actus reus* which is a prohibited act and, (2) the *mens rae* which is a knowledge or intention (Ferguson & Ogloff, 2011). In laymen's terms the *mens rae* is often thought of as being a knowledge or intention that the act was wrong. However, a more nuanced examination of the defense for criminal responsibility and *mens rae* demonstrates that this is not entirely the case, especially when someone commits a crime while under the influence of a mental disorder.

Table 1.1

Section 16 Criteria

Legislative Criteria NCRMD
1. No person is criminally responsible for an act committed or an omission made while suffering from a mental disorder that rendered the person incapable of appreciating the nature and quality of the act or omission or of knowing that it was wrong.
2. Every person is presumed not to suffer from a mental disorder so as to be exempt from criminal responsibility by virtue of subsection (1), until the contrary is proved on the balance of probabilities.
3. The burden of proof that an accused was suffering from a mental disorder so as to be exempt from criminal responsibility is on the party that raises the issue.

Criminal Code, R.S.C 1985, c. C-46, s.16

It is Section 16 of the *Canadian Criminal Code* (*Criminal Code*) that outlines the defense for a person who may have committed a crime while under the influence of a mental disorder

(Table 1; Criminal Code, R.S.C. 1985, c. C-46, s.16). Specifically, the criminal code defines a mental disorder as a "disease of the mind" (Criminal Code, R.S.C. 1985, c. C-46, s. 2). Thus, a person who has been found to have been suffering under the influence of a mental disorder at the time of their crime can be found *Not Criminally Responsible on Account of Mental Disorder* (NCRMD). This does not imply that if one commits a crime and suffers from a mental disorder that she or he is necessarily exempt from criminal responsibility. Instead the NCRMD defense (i.e., Section 16) acknowledges that a mental disorder may have impacted an individual's ability to know or appreciate that the act was wrong. In other words, the defense recognizes that an individual may have had intention (*mens rea*) in their actions but those actions may have been skewed to such an extent by their mental illness that they were incapable of discerning right from wrong and, in some cases, to even know what they were doing, and thus should not be held criminally responsible (Ferguson & Ogloff, 2011; Latimer & Lawrence, 2006).

Latimer and Lawrence (2006) described the finding of NCRMD as unique within the Canadian legal system as it represents neither a guilty plea nor a finding of innocence. Instead, it represents a third option wherein pursuant to section 672.38 of the *Criminal Code* an accused who is found to be NCRMD is diverted to a provincial or territorial review board unless they are immediately granted an absolute discharge by the court (Criminal Code, 1985). Once under the jurisdiction of a review board, an accused found NCRMD may be granted one of three dispositions outlined in Section 672.54: (a) absolute discharge, (b) conditional discharge or (c) detention in custody of a hospital. In the case of an absolute discharge, the accused is no longer under the jurisdiction of the review board and is free to go back into the community with no lingering restrictions upon his or her liberty. When granted a conditional discharge, an accused may be supervised in the community with restrictions placed upon their liberty. Section 672.54

(b) of the *Criminal Code* grants the review board authority to impose any restrictions it deems appropriate (Criminal Code, 1985). The third option, detention in custody of a hospital, is self-explanatory wherein the accused is detained within the confines of a hospital. Once again, the review board has the ability to impose conditions as deemed appropriate. Notably, the court that renders the finding of NCRMD, may also determine a disposition if it is deemed that the court is able to do so and it is believed that a disposition should be made without delay (Criminal Code, 1985; Latimer & Lawrence, 2006). Review boards represent a balancing act between entrenched civil liberties and public safety as it is the goal of the review board to assess a person found NCRMD and decide upon the disposition that most balances the person's right to freedom, and the safety of the public. Recently however, this has changed as new legislation has instructed review boards to place public safety as the paramount consideration in review board hearings (Bill C-14, 2014).

Specific guidelines within the *Criminal Code* dictate the creation of review boards. These guidelines stipulate that review boards are made up of no fewer than five members who have been elected by the Lieutenant Governor in council for that province. Three of the five members have specific requirements that must be met. For example, the review board must be chaired by a judge or someone who holds the equivalent credentials, at least one member must be a psychiatrist who is able to be licensed in the province or territory of that review board and at least one other member of the review board must have experience working in the field of mental health and be able to be licensed to practice medicine or psychology within that province (Criminal Code, 1985). The remaining two members are often professionals or lay persons who have experience within this area (e.g., lawyer).

1.2.1 Criminal Responsibility History and Relevant Legislation

Allnut, Samuels, O'Driscol (2007) noted that societies have long been debating the appropriate course of action for when a mentally ill person commits a crime. For example, evidence of this debate can be seen in Roman Law in the third century, in the 'code of Justinian' in the sixth century, in the 'Wild Beast Test' of the thirteenth century, and in the four formalized forms of insanity set out by the Lord Chief Justice of England, Mathew Hale, in the sixteenth century. Historically though, most attempts at dealing with this issue were haphazard, vague, and ill-defined. It was not until the M'Naghten standard in the nineteenth century that a more systematized manner of dealing with the insanity plea and determining criminal responsibility was established (Allnutt, Samuels, & O'Driscoll, 2007).

The M'Naghten standard of 1843 laid the foundation for the current Section 16 of the *Criminal Code* and formalized how criminal responsibility is thought of today in Canada (Ferguson & Ogloff, 2011; Allnutt, Samuels, & O'Driscoll, 2007). The M'Naghten standard was established as a result of an attempt on the life of then British Prime Minister, Sir Robert Peel by Daniel M'Naghten. M'Naghten believed the Prime Minister and the Tories were conspiring to kill him and the only way to save himself was to take the life of the Prime Minister. Unfortunately, M'Naghten mistook the Prime Minister's Secretary, Edward Drummond, for the Prime Minister and shot him in error. Edward Drummond eventually succumbed to his injuries. At M'Naghten's trial it was evident that he was mentally ill and not able to understand the consequences of his actions. Daniel M'Naghten was eventually acquitted of his charges. After his acquittal, there was a significant uproar regarding the decision and the House of Lords was called upon to set out clear standards for the insanity defence (Ferguson & Ogloff, 2011). The result of this clarification was the M'Naghten Standard and it laid out three elements for the insanity defence. The basic tenets, in laymen's terms, of each of the three elements of the

standard as they were outlined in 1843 are described in Table 2. Ferguson and Ogloff (2011) noted the M'Naghten standard is often considered a "cognitive" test of insanity because there is a focus on the quality of the accused's thought process at the time of the offense. This focus of the M'Naghten standard provided the foundation and general ethos of today's Section 16 parameters.

Section 16 of the *Criminal Code* has mirrored the M'Naghten Standard since its inception into the *Criminal Code* in 1892 and has changed very little over time. Notable instances where changes have been suggested or made have included the following: (a) *the Report of the Royal Commission on the Law of Insanity as a Defence in Criminal Cases (*McRuer Report; Privy Council Office, 1956); (b) *Bill C-30: An Act to Amend the Criminal Code and to Amend the National Defence Act and the Young Offenders Act in Consequence Thereof* (1992), which was a result of the Supreme Court of Canada decision *R. v. Swain* (1991); (c) *Winko v. British Columbia (Forensic Psychiatric Institute)* (1999); (d) *Bill C-10: An Act to Amend the Criminal Code (mental disorder) and to Make Consequential Amendments to Other Acts (2005);* and (e) *Bill C-14, The Not Criminally Responsible Reform Act (2014).*

Table 1.2

M'Naghten Criteria

1843 M'Naghten Standard Criteria
1. It must be determined that the accused was suffering from 'a defect of reason, disease of the mind'
2. Evidence must support the fact that the accused did not know what they were doing was wrong as a result of their mental disease.
3. An inquiry into whether the accused understood that what they were doing was wrong must be made

Ferguson, and Ogloff, 2011

The Royal Commission on the Law of Insanity as a Defense in Criminal Cases (1956) or as it was also known, The McRuer Report, was commissioned to inquire into criminal law in Canada regarding what was then called the "insanity defence" or *Not Guilty by Reason of Insanity* (NGRI). The commission held public meetings in all capital cities across Canada, as well as in Vancouver, Ottawa, and Montreal. The resulting report made 16 recommendations. These recommendations focused on the wording and terminology of Section 16 but also commented on such topics as whether to adopt the law of diminished capacity, making changes to laws regarding provocation, and a suggestion to the Supreme Court regarding the creation of a uniform explanation to jurors on the proper interpretation of Subsection 2 of Section 16 which deals with presumption. Overall few changes were made to the existing legislation as the existing laws were determined to be satisfactory.

In 1991, the first major change would be set in motion as the Supreme Court of Canada decision *R v. Swain* concluded that the indeterminate detention of an accused found NGRI was unconstitutional. At that time, those found NGRI remained under strict custody in a psychiatric facility at the 'pleasure of the lieutenant governor' of that province (Grant,1997). Rather than immediately striking down a portion of the Criminal Code that could result in the release of many individuals held on warrants across the country, the Supreme Court suspended its declaration for half a year to enable parliament to pass legislative reform. In February of 1992 most of Bill C-30*: An Act to Amend the Criminal Code and to Amend the National Defence Act and the Young Offenders Act in Consequence Thereof* (1992), was proclaimed into law.

The decision of the Supreme Court in *R. v. Swain* and the following Bill C-30 came at the end of a lengthy national discussion on the status of the insanity defence in Canada (Pilon, 2002). The Law Reform Commission of Canada first published a working paper on the topic in

1975 citing the need for reform and made several recommendations. This was followed in 1982 by the Department of Justice initiating the Mental Disorder Project which released a final report in 1985 ("Review of the Mental Disorder Provisions", 2002). Together these reports contained suggestions which eventually helped to inform Bill C-30. As a result of the long debate and recognition of the need for reform, Bill C-30 made sweeping changes to the insanity defence. The most significant change being that the insanity defence was no longer recognized as *Not Guilty by Reason of Insanity* but instead became *Not Criminally Responsible on Account of Mental Disorder (NCRMD;* Viljoen, Roesch, Ogloff & Zapf, 2003). The bill also made changes to the process of assessment orders, the determination of fitness to stand trial, and clarified the creation, role, and powers of provincial and territorial review boards. Bill C-30 was never fully proclaimed into law (i.e., certain sections of the bill were not proclaimed) and eventually was updated by Bill C-10 in 2005. The bill remains important in the history of NCRMD legislation as the impact of its changes are still debated today. For example, many argue that the changes made by Bill C-30 (1992) have increased the number of persons being found NCRMD as the lack of indeterminate status makes it a much more appealing course of action (Brodsky, Defence, & Information, 2017).

In 1999, the Supreme Court of Canada determined in *Winko v. British Columbia (Forensic Psychiatric Institute,* 1999; *Winko*) that if a person found NCRMD did not pose a significant threat to the safety of the public, then a review board must order an absolute discharge. The *Winko* decision also outlined that NCRMD status does not create a presumption of dangerousness, in that "the past offense committed while the NCRMD accused suffered from a mental illness is not, by itself, evidence that the NCRMD accused continues to pose a significant risk to the safety of the public" (Winko, 1999, para. 62). Lacroix, O'Shaughnessy,

McNiel and Binder (2017) noted that the data on the NCRMD population since this decision appears to support this stance.

The *Winko* decision also clarified that the indeterminate status of NCRMD until an absolute discharge was given was not unconstitutional as the NCRMD status was not meant to be punitive. In paragraph 93 of the decision it states:

> …it has been determined that the NCR offender is not morally responsible for his or her criminal act. Punishment is morally inappropriate and ineffective in such a case because the NCR accused was incapable of making a meaningful choice upon which the punishment model is premised. Because the NCR accused's liberty is not restricted for the purpose of punishment, there is no corresponding reason for finitude. The purposes of restriction on his liberty are to protect society and to allow the NCR accused to seek treatment. This requires a flexible approach that treats the length of the restriction as a function of these dual aims and renders a mechanistic comparison of the duration of confinement inappropriate. (Winko v. British Columbia, para. 93)

Thus, with the *Winko* decision, the principle of proportionality which is often applied in the sentencing of offenders was deemed to be inappropriate in the case of those found to be NCRMD (Latimer & Lawrence, 2006).

In 2002, a parliamentary review was conducted as required by Bill C-30. The review made 19 recommendations to address both outstanding parts of Bill C-30 that had not been fully proclaimed and new issues that had arisen since the bill's induction (Standing Committee on Justice and Human Rights, 2002). New legislation in the form of Bill C-10 was introduced to implement those recommendations and in May 2005, Bill C-10 received Royal Assent (Raaflaub,

2005). Changes resulting from the bill included repealing previously unproclaimed provisions, expanding the authority of review boards, items related to victim participation in hearings, and permitting courts to make certain decisions regarding cases involving people found Unfit to Stand Trial.

Table 1.3

Timeline of Notable NCRMD History

1843	M'Naghten Standard established
1892	Inception of insanity defence within the Canadian Criminal Code
1955	McRuer Report
1975	Law Reform Commission of Canada, *The Criminal Process and Mental Disorder*
1976	Law Reform Commission of Canada, *Mental Disorder in the Criminal Process*
1985	Department of Justice, *Mental Disorder Project Criminal Law Review*
1991	*R. v. Swain* Concluded that the indeterminate detention of those found NGRI was unconstitutional Decision delayed to enable parliament to pass legislative reform
1992	Bill C-30 Replaced NGRI with NCRMD Creation of provincial review boards Not fully proclaimed to law
1999	*Winko v. British Columbia* (*Forensic Psychiatric Institute*) Clarification of indeterminant status of NCRMD Clarification of presumption of dangerousness Requirement for RBs to grant AD if no longer a danger to public safety
2002	Parliamentary review required by Bill C-30 completed
2005	Bill C-10 Implemented recommendations from parliamentary review Addressed provisions not proclaimed into law from Bill C-30
2014	Bill C-14 Four key amendments including creation of High Risk Accused designation

Most recently, in July of 2014, Bill C-14*, The Not Criminally Responsible Reform Act* officially came into effect and made four significant amendments to the criminal code proceedings pertaining to those who have a mental disorder. The bill was a response to increasing public pressure to ensure public safety and mental health. The four main amendments included within the bill were: (a) increasing participation and/or consideration of the victims involved, (b) public safety is placed as the paramount consideration for review board decisions, (c) "significant threat to the safety of the public" was given a new statutory definition and (d) the creation of the designation of a "high-risk accused" (HRA; Bill C-14, 2014). The potential impacts of these changes are still debated and there is skepticism over whether there is evidence to support the creation of the "high risk accused" designation (Grantham, 2014; Lacroix, O'Shaughnessy, McNiel, & Binder, 2017; Goossens, Nicholls, Charette, Wilson, Seto, & Crocker, 2019). Charette et al. (2015) and Grantham (2014) both noted that the designation of HRA may suggest that those persons labeled as such have a higher probability of reoffending and as of yet, there is still no conclusive evidence to support this interpretation. Goossens, Nicholls, Charette, Wilson, Seto, and Crocker, (2019) further supported this idea wherein they argued that the HRA designation is more closely linked to index severity and as a result is largely missing the mark when it comes to protecting public safety. Charette et al. (2015) also noted that those who had been found NCRMD for severe index offences were less likely to reoffend than those who had not committed severe crimes, or those who had not committed crimes against the person. Thus considering all the issues related to the HRA designation, many have argued that it may be vulnerable to multiple challenges based on the Canadian Charter of Rights and Freedoms (Canadian Charter, 1982; Lacroix, O'Shaughnessy, McNiel, & Binder, 2017).

1.2.2 NCRMD Rates

Grossi and Green (2017) consider researchers, policy makers, and mental health professionals as having an obligation to understand the cultural and legal context in which criminal responsibility research is conducted. This is because research done in one jurisdiction may impact and inform legislation in another. Indeed, knowledge about the NCRMD population often depends on the country of interest and jurisdiction within that country. As such, it is not easy to capture a national or global image of how many people are found not criminally responsible and to determine realistic rates of prevalence and incidence (Grossi & Green, 2017). One reason for this, aside from how criminal responsibility may be conceptualized, is that in some countries, criminal responsibility legislation and administration is the responsibility of each individual state or province. In this way, each state or province may be left to define criminal responsibility as it sees fit and to administer the related legislation. In some cases, this may include even choosing not to have any version of an insanity defence. The United States exemplifies such a complicated situation as each state is left to define criminal responsibility, to administer legislation, and even in the case of four states, to decide not to have any version of an insanity defence available (Appelbaum, 2013). Thus, the heterogeneity that can exist between and within countries can make it difficult to capture a national or global image of those who are found not criminally responsible or equivalent (Grossi & Green, 2017).

According to the available data from Statistics Canada (2014) from 2005 - 2012 in which data from ten provinces or territories was available, the Canadian NCRMD population had an incident rate of approximately 7.5 to 9.1 per 10,000, and represented about 1% of criminal court cases on an annual basis. The National Trajectory Project (NTP), a leading national research endeavor on individuals found NCRMD in Canada, examined the trajectories of 1800 men and

women found NCRMD in the nation's most populated provinces (i.e., British Columbia, Ontario and Quebec) between April 2005 and December 2008 (Crocker, Nicholls, Seto, & Côté, 2015; Crocker, Nicholls, Seto, Charette, et al., 2015). These provinces also had the highest numbers of persons found NCRMD within Canada (Crocker, Nicholls, Seto, Côté, et al., 2015). Several notable interprovincial differences in rates of person's found NCRMD were identified and the authors noted that this may be due to the fact that in Canada NCRMD legislation is federally defined but provincially administered (Crocker, Nicholls, Seto, Charette, et al., 2015; Crocker, Nicholls, Seto, Côté, et al., 2015). Baillie (2015) noted that given the notable provincial differences identified by the NTP, caution should be exercised when using national statistics to make assumptions regarding local circumstances. The authors of the NTP also noted that different trends over time were observed regarding the number of persons entering the provincial review board systems. This information added fuel to the debate on whether the number of persons being found NCRMD had been increasing since the 1992 legislative changes (Crocker, Nicholls, Seto Côté et al, 2015). Demarais, Hucker, Brink and De Freitas (2008) argued that the apparent increase in NCRMD rates was not as extreme or intense as it was often thought to be and that more interprovincial differences may exist than differences created by changes in any legislation. Demarais et al. (2008) also noted that given this, research on those found NCRMD might benefit from multi-site sampling or a focus on individual provinces. Penney et al. (2018) who studied the forensic population in Ontario cited a pragmatic reason for knowing the trends within each province: the potential economic consequences. Forensic services are not inexpensive and an increasing population could mean increased costs for provinces. Thus, some of the most prominent NCRMD research in Canada echo's the sentiments made by Grossi and

Green (2017) and highlights the importance of looking at each individual province to acquire the most accurate information regarding NCRMD rates in Canada.

Prior to the commencement of *The Alberta NCR Project*, a research endeavor captained by Dr. Andrew Haag at Alberta Hospital Edmonton, very little was known about the NCRMD population within Alberta. It is now known given the work done by *The Alberta NCR Project* that as of October 2018, there have been 574 cases to come under the Alberta Review Board's (ARB) jurisdiction. The first known case dated back to 1941 and as of 2016 there were 209 active cases under the ARB's jurisdiction (Haag, Cheng, Wirove, 2016). *The Alberta NCR Project* has multiple research projects currently underway (this study included) in the hopes of continuing to increase what is known regarding the NCRMD population in Alberta.

1.2.3 Unfit to Stand Trial

Related to NCRMD is the legal concept of Unfit to Stand Trial (UST). The two concepts are closely related but differ in significant ways. Most notably, NCRMD is focused on the mental state of the accused at the time of the offense while, UST is focused on the mental state of the accused at the time of court proceedings (Viljoen, Roesch, Ogloff, & Zapf, 2003). The *Criminal Code* defines the criteria for being found UST in section two (Table 1.4) and further expands upon it in section 672.22 (Criminal Code, R.S.C 1985). Like individuals found NCRMD, an accused found UST will come under the jurisdiction of a provincial or territorial review board. However unlike individuals found NCRMD, a court or review board does not have the authority to order an absolute discharge. Instead a person found UST will stay under the jurisdiction of a review board until they are deemed fit or their charges have been stayed or withdrawn (Latimer & Lawrence, 2006).

Supreme Court of Canada decision *R. v. Demers* (2004) determined that detaining an accused who is "permanently" unfit and who does not pose a significant threat to the safety of the public infringes upon civil liberties given in the *Charter of Rights and Freedoms* (Canadian Charter, 1982). Bill C-10 (2005) addressed these issues. It granted the court the authority to stay proceedings if three specific conditions were met: (a) the accused is unlikely to become fit, (b) the accused does not pose a significant threat to the safety of the public, and (c) a stay of proceedings is in the interest of the proper administration of justice (Latimer & Lawrence, 2006).

Table 1.4

Legislative Criteria for Unfit to Stand Trial

Unfit to stand trial means:
Unable on account of mental disorder to conduct a defence at any stage of the proceedings before a verdict is rendered or to instruct counsel to do so, and in particular, unable on account of mental disorder to: a) Understand the nature or object of the proceedings b) Understand the possible consequences of the proceedings, or c) Communicate with counsel

Criminal Code, R.S.C 1985, c. C-46, s. 2

1.3 Gender and Individuals Found NCRMD

According to the available data from Statistics Canada (2014) from 2005-2012 in which ten provinces or territories reported, the Canadian NCRMD population was mostly male (87%). This converges with existing literature in that an approximate gender ratio of five to one, males to females, is often found in the Canadian NCRMD population (Crocker, Nicholls, Seto, Charette, et al., 2015; Nicholls et al., 2015; Livingston, Wilson, Tien, & Bond, 2003; Haag, Cheng, Wirove, 2016). This gender ratio also converged with an international sample from Japan, where 2,094 mentally ill offenders who were found either partially or fully not criminally

responsible on account of mental disorder were studied between the years of 1980 and 1994 (Xie, 2000). In this sample, 87% were also male. However, Norko et al. (2016) in the United States studying the Connecticut Psychiatric Security Review Board (i.e., the approximate equivalent of provincial or territorial review boards in Canada) found a gender ratio of 10:1 for males to females in this population of insanity acquittees.

Nicholls et al. (2015) in conjunction with the NTP, noted that females found NCRMD in Canada had psychosocial, clinical, and criminogenic profiles that were distinct from the males within their sample. For example, females were more likely to have been diagnosed with mood or personality disorders while males were more likely to have been diagnosed with schizophrenia spectrum disorders or substance abuse disorders. Index offenses did not differ between the genders, apart from women who were more likely than men to have committed or attempted murder. These authors also noted that not only did the genders have distinct psychosocial, clinical, and criminogenic profiles but also specific victim profiles. Females were more likely to have offended within private relationships (e.g., significant others, family members, children) than males who were more likely to have offended against acquaintances or strangers. Penney, et al. (2018) found similar results regarding the forensic population in Ontario between 1987 and 2012.

Distinct gender profiles were also found in an international sample from Japan (Xie, 2000). For example, females were more likely to be older at time of index offence, had received more education, were more likely to be married, had less recorded criminal involvement, were more likely to have attacked a family member and were also more likely to have committed a single lethal crime than the males within the same sample. These characteristics closely resemble those found within the NTP. Logan and Blackburn (2009) in the United States examined

mentally disordered offenders in the general prison population (i.e., not found NGRI/NCRMD) and found a similar victim profile for males and females as noted above. Results such as this could indicate that distinct victim profiles by sex are not limited to those found NGRI/NCRMD.

This type of specific gender information can be vital to review boards in formulating dispositions as it may impact factors such as treatment, and public safety. While gendered differences were found in profiles, Nicholls et al. (2015) noted that these differences were not necessarily inconsistent with traditional models of offender assessment and treatment (i.e., Risk-Needs-Responsivity; RNR; Andrews, Bonta, & Hoge, 1990) and if anything, further supported the importance of individual assessment and client-centered services (i.e., gender as a responsivity issue).

1.3.1 Gender Differences in General Offender Population

In Canada, general offenders are managed much differently than those who have been found NCRMD. In fact, the federal, provincial, and territorial governments share the responsibility of administering correctional services. The federal system is responsible for overseeing those who are 18 and older, and have been sentenced for longer than two years and those who are on conditional release in the community (i.e., parole or statutory release). Provincial or territorial governments are responsible for overseeing those who will serve custodial sentences of two years or less, on remand (i.e., awaiting trial) and those who are serving community sentences (i.e., probation). On an average day in Canada there are 40,147 adults in custody, this includes provincial, territorial, and federal custodial settings; 25,405 of these are in provincial or territorial custody and 14,742 are in federal custody (Reitano, 2017). Of the 14,742 in federal custody, 680 are women (i.e., 4.6% of the population; Sapers, 2016). This can be compared with the fact that in the United States, the Bureau of Federal Prisons (2017)

indicated that of all federally sentenced offenders only 7% are female. Despite the variability in gender make-up, the general trend across all offenders, criminally responsible or otherwise, is that female offenders represent a significantly smaller portion of the population.

It is noteworthy that in Canada in the last decade the number of federally sentenced women has increased by 30% (Zinger, 2018). It is currently unknown whether the rate of females found NCRMD is rising in Canada. This represents another area of study within the NCRMD population that has not been thoroughly examined. Sapers (2016) also noted that compared to their male counter parts women in federal custody had higher rates of mental health needs. This was reiterated by Nicholls, Cruise, Greig, and Hinz (2015), who noted this also stands true for females in American prisons. Derkzen, Booth, Taylor and McConnell (2013) noted that a pattern exists in the research that female offenders often have higher rates of substance abuse, major depressive disorder, and Posttraumatic Stress Disorder (PTSD). A more thorough investigation of general female offenders is outside the scope of this study, however for more information the reader can refer to Nicholls, Cruise, Greig, and Hinz (2015) for an in-depth overview.

1.3.2 Gendered Pathways to Crime

What drives such notable gender differences within offenders is a topic of much debate. This debate frequently focuses on how similar males and females are in how they come to be involved in crime, how they progress through their criminal careers, and what criminological factors most pertain to their treatment and assessment of risk for reoffending. This debate is further complicated among those found NCRMD in that mental illness adds another layer of examination. For now, just looking at the debate of males versus females and crime, multiple important points can be explored.

To begin, Blanchette (2004) noted that the literature often refers to how people become involved in crime as 'pathways to crime' and that it is generally understood, that most of what is known about pathways to crime, offending trajectories, and recidivism is based on samples of white, male offenders. Thus, many researchers, Blanchette included, have called for more work to determine if the current assumptions about crime can be applied to both males and females equally. The Risk-Needs-Responsivity (RNR; Andrews, Bonta, & Hoge, 1990) model would argue that there are more differences within genders than between and thus the risks and needs associated with crime would apply more or less equally to both genders (RNR is explained further below). A host of critical feminist researchers have argued that this is untrue and that females have unique pathways to crime, criminal careers, and risks for recidivism. Indeed, they have argued that females require specific and distinct consideration from their male counterparts and that a common approach is markedly inappropriate. For example: Daly (1992, 1994) a trailblazer of the gender specific pathways to crime discourse, suggested five unique pathways to crime for women (e.g., harmed and harming women, battered women, street women, drug-connected women, and other). Block, Blokland, van der Werff, van Os and Nieuwbeerta (2010) also argued that the assumptions currently held about criminal career trajectories may not apply in the same way for women. They pointed to a trend in their research of an older onset of offending, as many women in their study had a criminal career that began after 45 years of age. Such a late onset would be unexpected in a more traditional RNR framework. Delsis (2002) also argued that drug dependence and crime are closely connected for females and may be an important consideration when examining career paths and trajectories. Referring to the RNR model, Blanchette, (2004) questioned whether the traditional risks involved in offending (i.e., the Central Eight) applied equally as well to a female population, arguing that there is research for

additional risks that may need to be considered for women or even risks that may better fit their distinct context. Regarding the central eight, Wilson, Crocker, Nicholls, Charette, and Seto (2015) noted that there may not be so much of a "big four" for women but a "big five" when considering the central eight criminological factors. The authors suggested that the evidence could support substance abuse being added to the current group of major risk factors for women.

Continuing to explore the feminist critique of more traditional models in a correctional setting, Blanchette (2004) argued that in order to have effective correctional services for women, women need to be considered distinct from their male counterparts. Blanchette reasoned that women are a heterogeneous group and they have many unique considerations. An example of such a unique consideration was argued by Chesney-Lind (2000) who stated the importance of considering women's victimization histories when working towards a holistic classification within an effective correctional intervention plan. Blanchette (2004) acknowledged that the risk classification of women in corrections is a precarious endeavor as issues such as base rates and lack of research can impede accurate assessment, an argument that more traditional models such as RNR have long cited as prominent barriers to fully understanding female offending. It is an interesting point that despite the heated debate that often rages in the literature over this topic, both sides are invested in more research being done and believe that correctional policy and practice should be based on empirical data and research (Taylor and Blanchette, 2009)

Mental illness further complicates this debate and raises the question of whether mental illness impacts the factors involved in males and females' pathways to crime, criminal trajectories or risks for recidivism. Bonta, Law, and Hanson (1998) argued that from their meta-analysis regarding general and violent recidivism, the same factors that predict for non-mentally ill offenders were also valid for mentally ill offenders. They found that clinical symptoms did not

predict over and above the factors normally considered within the principles of RNR. The idea that mental illness itself is not a predictor of violence is one that is often found and supported in the research.

A solution regarding mental illness, gender, pathways to crime, and criminal trajectories is outside the scope of this study and, at the moment few conclusions can be drawn due to the divergent literature that currently exists (Taylor and Blanchette, 2009). Though at the very least it could and should be argued that gender is an important responsivity issue and needs to be attended to in seriousness. This includes more research being done on the potential gender differences that might exists within various offender populations (i.e., general, mentally disordered, forensic). As more gender informed tools are created, validated, and show incremental validity over gender neutral tools, they should be increasingly used within the female population. This would demonstrate a dedication to evidence driving correctional and forensic practice.

1.4 Criminological Aspects of Those Who have Been Found NCRMD

In the following subsections, topics regarding criminological characteristics such as violence, conditional discharge, and recidivism are examined in relation to both individuals who are found NCRMD and gender. As has already been done, this information is juxtaposed with the general offender population as a way to highlight the many unique aspects of the NCRMD population.

1.4.1 Violence and the NCRMD Population

According to the available data from Statistics Canada (2014) almost two-thirds (63%) of the NCRMD cases brought forward between 2005 and 2012 included crimes against the person; major assault represented 20% of those cases. The National Trajectory Project (NTP) found similar results with 64.9% of the index offenses within their sample being crimes against the

person, these authors also noted that females were as likely as males to have an index offense that included a crime against the person (Crocker, Nicholls, Seto, Charette, Côté & Caulet, 2015; Nicholls, Crocker, Seto, Charette, & Côté, 2015). Also, noted by these researchers was that females found NCRMD had substantially more charges that involved death or attempting to cause death (Nicholls, Crocker, Seto, Charette, & Côté, 2015). Interestingly, this finding was also found in the research done by Xie (2000) in a Japanese sample. Livingston, Wilson, Tien, and Bond (2003) examined a NCRMD cohort within British Columbia shortly after the passing of Bill C-30 in 1992 and found that the most serious offense was assault for almost half (45.5%) of the cohort; assault-type offenses made up 36.4% of the total cohort index offense. In Alberta, direct violence, coded as all violence excluding sexual offences, homicide and attempted homicide, accounted for 46.9% of the population indexes, with homicide representing 18.5% and attempted homicide 10.6% (Haag, Cheng, Wirove, 2016). Clearly, violence as related to the index offense is a consistent theme within this population. Though caution should be taken when interpreting this as the *Winko v. British Columbia* (Forensic Psychiatric Institute, 1999) decision reinforces that index offense severity is not to be taken as evidence that a person found NCRMD will continue to pose a danger to public safety. Indeed, the purpose of the NCRMD defense is to highlight that the individual was suffering under a defect of reason at the time of their offense - no matter how severe.

1.4.2 Violence and The General Offender Population

Information regarding violence within the NCRMD population can be juxtaposed with information about violence in the general offender population and in civil psychiatric patients. For example, 69% of all the federal offenders in Canada in 2016 were serving time for a violent offense. Broken down by gender, 54.3% of all women and 69.9% of all men serving federal

sentences were doing so for at least one violent offence (Public Safety, 2016). CANSIM (2017) data from Statistics Canada, indicated that crimes against the person accounted for approximately 25% of all guilty pleas in Canadian criminal courts not including traffic violations. Women in the general offender population committed violent crimes or any crime at a lower rate than did men. This trend is not unique to Canada and has also been found in the United States of America, and in Australia (Tye and Mullen, 2006; Nicholls, Cruise, Greig, & Hinz, 2015. This trend is notable in the context of considering how the low base rate of female offending in general and even lower base rate of violent female offending can impact our ability to predict violence or the reoccurrence of violence, NCRMD or otherwise.

The McArthur Studies, which followed over 1,000 civil inpatients in three states in the United States attempted to determine risk factors most pertinent to risk prediction for this population, indicated that violence risk attributed to people with mental disorders vastly exceeded the risk that those with mental disorders actually had (Monahan et al. 2001). The authors suggested that for people who do not abuse drugs or alcohol, there is no reason to anticipate that their risk for violence would be any greater than the average person (Monahan, 2002). This conclusion was congruent with Hiday (1997) who argued for the importance of social factors to be considered in provoking violence in people with mental health issues.

1.4.3 Recidivism

Bonta, Ruge, and Dauvergne (2003) noted that there is no universally agreed upon measure for recidivism, although in general it is commonly defined as a return to crime. These authors argued that recidivism data may be one of the most important indicators of impact of criminal justice interventions and that recidivism data provides information that has value for multiple areas of the criminal justice system (e.g., crime prevention, the police, the courts,

corrections). Recidivism data are also very important for evaluation, treatment, and rehabilitation efforts. Bonta et al. (2003) further stated that many variables may impact the measurement of recidivism and specifically identified four major factors: (a) the definition of crime, (b) the nature or type of crime, (c) the length of follow up time, and (d) the difference in sample of offenders. Lacroix, O'Shaughnessy, McNiel and Binder (2017) noted that historically it has been difficult to determine precise recidivism rates for those persons found NCRMD for a variety of reasons. Some of those reasons are related to the issues mentioned by Bonta et al. and some are related to previously raised issues (e.g., how criminal responsibility is defined, administered, and reported) but others have to do with the individual person found NCRMD and the specific context of the individual's situation or more administrative factors dependent upon institutional policy. For example, factors possibly vary depending upon what stage of the proceedings an individual is currently in or what disposition an individual is currently under or whether institutions (e.g., hospitals) have different policies regarding institutional breaches of conditions.

1.4.4 Recidivism and the NCRMD Population

Charette et al. (2015) in conjunction with the National Trajectory Project (NTP), examined recidivism rates of persons found NCRMD and identified several key findings. First, at a three year follow up after being granted an absolute discharge, the rates of recidivism for those found NCRMD was 17% following the index offence, 22% following conditional discharge and 22% following absolute discharge. The authors noted that these rates were lower than those found in the general offender population (34%) and much lower for those general offenders who had been treated for a mental disorder (70%) (Johnson, & Grant, 2000; Villeneuve,& Quinsey, 1995). Charette et al. (2015) further suggested that this finding indicated

that the NCRMD population is adequately managed under a review board. Second, the authors found that a history of criminal convictions and NCRMD findings was a good predictor of future offenses. Third, even after controlling for a variety of factors (e.g., number of prior offences, seriousness of index offence, diagnostic category, review board supervision), there were provincial differences in rates of recidivism for those individual found NCRMD. The findings that individuals who are found NCRMD have lower rates of recidivism comparatively to other groups of offenders, that they are well managed by review boards, that number of previous convictions is a good predictor of future offenses, and that there are provincial differences in rates of recidivism are all supported by other research across Canada and the globe (Richer, Cheng, & Haag, 2018; Goossens, Nicholls, Charette, Wilson, Seto, & Crocker, 2019; Friendship, McClintock, Rutter, & Maden, 1999; Grann, Danesh, & Fazel, 2008; Hayes, Kemp, Large, & Nielssen, 2014; Fazel, Fimińska, Cocks, & Coid, 2016; Norko et al. 2016; Simpson, Chatterjee, Duchcherer, Ray, Prosser, & Penney, 2018; Tabita, de Santi, & Kjellin, 2012).

At the intersection of individuals who are found NCRMD, recidivism, and gender, limited amounts of literature exists. What is known about female general offenders is thought to extend to those females who have been found NCRMD (e.g., lower rates of reoffending). Though, more research needs to be conducted on recidivism and potential gender differences within the NCRMD population to know for sure.

1.4.5 Recidivism and the General Offender Population

Different offender samples produce different recidivism rates and different measures of recidivism make it difficult to compare across samples and populations. That being said, for a Canadian comparison to the NCRMD population, research by Bonta, Rugge, and Dauvergne (2003) can be examined. These authors studied recidivism in three cohorts of Canadian federal

offenders (i.e., samples include all those released from federal penitentiaries, except those on temporary passes, in the fiscal years of 1994/95, 1995/96, 1996/97) and used as their operational definition of recidivism as a new conviction within two years of release. What they found was recidivism rates of 44%, 42.8% and 40.6% for each of the cohorts respectively, with confidence interval overlap for the first two groups but not for the first and third. A similar trend was found for non-violent reconviction as the third group had a lower reconviction rate than the first group. There was no change in violent reconviction rate over the three groups. Females made up a small portion of this group (2.9%) and the bulk of the most serious offenses counted for females and males were crimes against the person. Across all categories of reconviction examined (i.e., any, non-violent, violent) men showed higher reconviction rates than women. Due to the small sample size of women the authors called for extreme caution regarding the female results. While comparisons across populations and samples need to be examined with prudence, it is notable that for those found NCRMD the rates of recidivism are often quite lower than in comparison to the general offender population.

1.4.6 Conditional Discharge and the NCRMD Population

A conditional discharge is given to a person found NCRMD when it is believed that they may be successfully supervised within the community with limitations placed upon their liberty (Latimer & Lawrence, 2006). A conditional discharge may be revoked at any time. For example, if any of the conditions are broken, if a new charge is acquired or if it is believed that the risk involved in the person remaining in the community is rising. Crocker et al. (2015) in conjunction with the NTP, examined outcomes of review board decisions and conditional discharge data. They noted that there were significant differences in the trajectories of a NCRMD person by province even after holding the number of past offenses, index severity and

diagnoses constant. What did not vary by province was the fact that a high number of past offences reduced the chance of a person under the review board of receiving an absolute discharge or a conditional discharge. The authors noted that there was some evidence to support the idea that the severity of the index offence was related to the duration of detention within a forensic setting and total duration under the review board. Within this study, outcomes by gender were not examined and there was not an analysis of review board decisions and conditional discharge information by gender.

Vitacco et al. (2008) studied a sample of persons found Not Guilty By Reason of Insanity (NGRI) in Wisconsin who were in the community on a conditional discharge for five years (2000-2004). What the authors found was a high revocation rate but that the actual rate of new crimes was low as most revocations were a result of conditional release violations (e.g., not taking medication). Vitacco et al. (2011) evaluated the conditional release status of 76 American females found NGRI in the state of Wisconsin and found that over a seven-year period 68% (n = 52) maintained their conditional release while 32% (n = 24) had their conditional release revoked due to either a rule violation or a new criminal charge. Notably, none of the new charges acquired were for violent crimes. The findings of this study were in line with their previous research. Vitacco, Vauter, Erickson, Ragatz (2014) conducted a prospective follow up study of 127 individuals found NGRI in Virginia who were released on a conditional discharge. In this study, three-quarters of the sample (75.6%) maintained their conditional release. The authors noted that demographic information was not a predictor of revocation (e.g., gender, age, education status, ethnicity). The authors also noted three factors that were significantly related to conditional discharge revocations: (a) substance-abuse diagnosis, (b) higher supervision level, (c) mental health issues requiring short-term hospitalization in a non-forensic facility. It bears

repeating that the results of such studies would indicate that those found NCRMD can be successfully maintained within the community.

1.4.7 Conditional Discharge and the General Offender Population

According to Statistics Canada ("*Parole, Pardons and Clemency" 2016*) the number of federal offenders on conditional release was 9,189. Of these, a high majority of both day parole (91%) and full parole (88%) successfully completed their conditional release. Full parole in this case included only those offenders serving determinate sentences, those serving indeterminate sentences were not included in these numbers. This can be compared with those offenders who were out on statutory release; here a lower percentage (63%) of offenders were able to complete their releases. Statistics Canada also noted that offenders on statutory release were more likely to have their release revoked due to a breach of condition or a new offense than federal offenders on day parole or full parole. Statistics Canada did not break this information down by gender so it is currently unknown if there is a gender difference that exists in the federal offender population who are out on conditional release. Once again, care must be taken in making comparisons across samples

1.5 NCRMD and Indigenous Issues

While it is outside of the scope of this study, it is important to note that while Indigenous people make up 4.9% of the Canadian population they make up 25% of the federally sentenced prison population and of the women who are federally sentenced 36% of them are Indigenous (Statistics Canada, 2017; Sapers, 2016). These high numbers within the federally sentenced prison population are in stark contrast to the number of Indigenous people found NCRMD. Nicholls, Wilson, Charette, Crocker, Seto (2015, March) noted that 2.9% of NCRMD persons nationally were Indigenous with only 7.7% of the NCRMD population in British Columbia being

Indigenous. In Alberta, persons of Indigenous background made up 8.8% of the NCRMD population (A. Haag, personal communication, November 2017). Thus the glaring difference in rates between the general non-offender population, the general offender population, and those found NCRMD with Indigenous background constitutes a research topic worthy of its own study.

1.6 Forensic Risk Assessment

A vital task for review boards is the ability to accurately assess a person's risk for reoffending. One way to accomplish this task is through the use of structured risk assessment measures. The extensive use of risk assessment measures, not only by review boards but by psychologists and other mental health professionals who are increasingly being called upon to determine risk for reoffending, is indicated by the exponential proliferation of risk assessment measures, literature, and research done on the topic (Mills, 2017; Viljoen, McLachlan, & Vincent, 2010; Ferguson & Ogloff, 2011). In fact, Viljoen, McLachlan and Vincent (2010) noted that some risk assessment measures are so well validated that the omission of their use in the determination of risk may be questioned in court.

1.6.1 Generations of Risk Assessment

Risk assessment measures though are not a monolithic enterprise and often vary vastly in their approach and conceptualization (Heilbrun, Fairfax-Columbo, Wagage, & Brogan, 2017). The differences in risk assessment measures over time has been historically conceptualized as evolving through generations (Andrews & Bonta, 2015; Andrews, Bonta, & Wormith, 2006; Mills, 2017). The first generation of risk assessment was defined by the use of pure clinical judgment. During this time, professional wisdom and the 'gut instincts' of the assessor were relied upon to make determinations of risk. First generation risk assessment or clinical wisdom is often criticized for its vulnerability to cognitive biases. Brook (2017) noted that the

fundamental attribution error, illusory correlation, and the tendency to assign undue weight to certain factors because of recency or salience were just a few of the examples of cognitive biases that may impede the accuracy of unstructured professional judgement. Indeed, regarding violence, it is well established within the literature that unstructured clinical judgment cannot be relied upon to make accurate predictions (Monahan, 1988).

Second generation risk assessment measures disavow pure clinical intuition and instead are statistically constructed and focus primarily on static items. Static items are historic facts or pieces of information that are not likely to change (e.g., age at first offence, sex). Second generation risk assessment tools are often called actuarial measures because of their reliance on statistics in their creation and scoring. The items in these tools are selected based on whether they are predictive of recidivism and in general are atheoretical in nature. These items are then given a weighting and all together summed to arrive at a final risk score. In a meta-analysis done by Ægisdóttir et al. (2016) it was found that actuarial measures tended to be more accurate than pure clinical judgment. Criticisms of second generation risk assessment measures generally focus on the idea that they are unable to capture change, and are limited in their ability to inform treatment. Examples of second generation measures are the VRAG (Harris, Rice, & Quinsey, 1993), the VRAG-R (Rice, Harris, & Lang, 2013) and the Static-99R (Hanson, and Thornton, 2000; Helmus, Thornton, Hanson, & Babchishin, 2012).

Third generation risk assessment measures were developed from theory and research and contain both static (i.e., items that are resistant to change) and dynamic (i.e., items that are expected to vary and change with time) factors. Dynamic items often reflect criminogenic needs (e.g., employment, education, antisocial associates, antisocial attitudes; Andrews and Bonta, 2015) and these dynamic items are used to guide intervention. The use of both static and

dynamic items reflects the idea that risk can fluctuate and by looking at both static and dynamic factors a more holistic view can be taken. These are also frequently actuarial in nature, given the that items are summed to generate a numeric score, which in turn, is used to generate a risk rating, and often also a recidivism estimate over a defined follow-up period. An example of a third generation risk assessment measure is the Level of Service Inventory - Revised (LSI-R; Andrews & Bonta, 1995). Limitations of the third generation measures generally focus on the fact that they do not capture change over various timepoints in treatment.

Fourth generation risk assessment measures only recently arrived in the field of forensic risk assessment. These are also often, but not exclusively, actuarial in nature, and were created to not only assess risk but also to guide treatment and to evaluate change as a result of treatment or other extraneous factors. Previous risk measures were only used at single time points however fourth generation measures are meant to be used over the duration of a case. For example, they can be used from entry into the justice system, to treatment, to release, to community case management, to case closure. Examples of fourth generation risk assessment measures include the Level of Service/Case Management Inventory (LS/CMI; Andrews, Bonta, & Wormith, 2004) and the Violence Risk Scale (VRS; Wong and Gordon, 1999).

A contemporary offshoot or hybrid of the generations is labelled Structured Professional Judgement (SPJ). Within SPJ, professional judgement is focused on specified items without any systemic mechanism linking scores to decisions (Andrews, Bonta, & Wormith, 2006). This is done in the hopes that both the professional wisdom of the assessor and the context of the offender may be considered while reducing the influence of cognitive biases. The intended strength of SPJ lies in the assessor's ability to appraise risk taking into account factors such as situational or special circumstances and other factors that might not be mentioned in the

measure. SPJ's strength lies in its flexibility, and its accuracy is most often comparable to actuarial measures (Heilbrun, Fairfax-Columbo, Wagage, & Brogan, 2017). A common limitation cited for SPJ instruments is that the meaning of their risk levels are often unclear.

Thus, regardless of generation, the benefit of structured risk assessment measures is that they remove human judgment biases and in some cases, increase treatment efficacy by indicating the most appropriate treatment targets (Brown & Singh, 2014). This ability to limit human judgment bias is what makes them extremely useful in legal settings or when decisions regarding public safety are being debated. It is also for this reason that further developing and testing of these measures is a worthwhile academic endeavor. Currently, asking which tool is best is more a question of its purpose and context in the greater scheme of conducting a risk assessment. Campbell et al (2009) conducted a meta-analysis that indicated that all instruments designed to assesses violence risk in adult offenders generally predicted comparably. Mills (2017) expanded on this, reinforcing that a risk assessment instrument is not equivalent to conducting a risk assessment; that multiple risk assessment instruments should be used in the process of conducting a responsible and ethical risk assessment. To rely only on one risk assessment measure and not the integrated assessment of data from multiple measures could be considered irresponsible on the part of the assessor.

1.6.2 Theoretical Background to Contemporary Risk Assessment

A prominent theoretical backdrop for much of the current thinking on risk assessment is based in a general personality and cognitive social learning (GPCSL) perspective. A GPCSL perspective evolved out of decades of research attempting to understand why people commit crime and owes its current existence to research that includes psychodynamic theories (e.g., Freud), learning theories (e.g., Operant and Classical conditioning), and social learning theories

(e.g., Sutherland's Differential Association Theory, Aker's Social Learning Theory). In a GPCSL context "crime cannot be understood without understanding whether the personal, interpersonal, and community supports for human behavior are favourable or unfavorable to crime" (Andrews & Bonta, 2015, p. 55). The GPCSL perspective can also be thought of as sitting within a more general Psychology of Criminal Conduct (PCC) that seeks a "rational and empirical understanding of the variation in the occurrence of criminal acts and, in particular, a rational empirical understanding of individual differences in criminal activity" (Andrews & Bonta, 2015). In other words, the objective of the more broad PCC is to understand variation in delinquent and criminal behaviors in individuals in a way that is holistic, general, and valuable to an interdisciplinary audience.

As mentioned above, the foundation of GPCSL has been built on decades worth of research from various perspectives. One of those perspectives is psychodynamic. Psychodynamic theories are rooted in Freudian thinking and are focused on the internal (and often unconscious) psychological forces that influence human behaviour. From a psychodynamic perspective humans are thought to be inherently antisocial and driven by pleasure-seeking and destructive impulses (Blackburn, 1995). As humans grow they develop internal psychic forces (i.e., ego and super-ego) that regulate such impulses. According to psychodynamic theories, crime occurs when these internal forces are not adequately able to control these antisocial and pleasure seeking impulses. It is hypothesized that early experiences in childhood, like a traumatic experience or ineffective parenting, can have a profound negative impact on personality development and how people behave.

Examples of the research done from a psychodynamic perspective that has informed on risk assessment today includes the work of Sheldon and Eleanor Glueck who were a husband and

wife team from Harvard in the 1950's. Their research represented some of the first systematic empirical explorations of psychodynamic theory (Andrews & Bonta, 2015). The Glueck's were interested in juvenile delinquency, specifically the causes of delinquency and in predicting the likelihood of delinquent behaviour (Brown et al., 2017). Their classic books summarizing their findings *Unravelling Juvenile Delinquency* (Glueck and Glueck, 1950) and *Delinquents and Non-Delinquents in Perspective* (Glueck and Glueck, 1968) identified parenting factors as the primary source for superego development and that there were marked differences between the parenting styles of delinquent and non-delinquent groups. The Glueck's noted that parents of the delinquents were more likely to be less educated, have greater incidences of emotional disturbances, alcoholism and criminality. From a psychodynamic perspective the Glueck's hypothesized that the development of a properly oriented superego or conscience could be greatly hampered by the kinds of parental ideals, attitudes, and temperaments demonstrated in the family unit of a developing child. As a result of their research they proposed a tentative formula that they believed could be used to predict who would become engaged in juvenile delinquency. While their research was not without its criticism (e.g., methodological issues related to causal ordering), their research was important because it was some of the first to highlight the variables related to criminal involvement particularly those variables related to family, peers, and school.

Though not traditionally considered psychodynamic, the control theories of crime emphasize factors that control people's behaviours and prevent them from committing crime. These theories of crime often reflected important psychodynamic themes throughout. Two popular control theories of crime are Travis Hirschi's Social Control Theory (Hirschi, 1969) and Gottfredson and Hirschi's General Theory of Crime (Gottfredson and Hirschi, 1990). Travis

Hischi's Social Control Theory argued that those persons who have a strong abiding connection to conventional society (in the form of investments in conventional pursuits, attachments to significant others, commitment to conventional behaviour, and belief in common rules and systems) are less likely to commit crime than persons who have weak or shallow bonds. This particular research was important because so many of its conclusions clearly overlapped with the findings of the Glueck's and highlighted the importance of variables such as parental supervision, school associates (i.e., delinquent or not), verbal intelligence and attitudes. Later on in his career Hirschi returned to this theory and in collaboration with colleague Michael Gottredson, refined his original theory and proposed a more parsimonious version. This General Theory of Crime (Gottfredson and Hirschi, 1990) focused less on the indirect controlling effects of social bonds as did Hirschi's original theory and instead focused more on self-control. This new theory suggested that people who possess a high degree of self-control do not commit crime and that a person's degree of self-control was largely a result of effective parenting (Brown et al., 2017). More simply, this theory suggested (among other ideas) that low self-control was the variable responsible for individual differences in crime. Continuing to shape and update his ideas, in 2004 Hirschi returned to his theory on control and redefined self-control as reflecting a pattern of considering the range of potential costs to a particular act (Andrews & Bonta, 2015). A perspective that is reflected in the current GPCSL perspective in that crime is partly under the control of many variables and their potential costs or rewards to the individual.

Research done from a psychodynamic perspective and the closely related control theories of crime were some of the first studies to begin to empirically examine crime and identify variables of interest when it came to delinquency and laid important foundational research for future risk assessment variables. Learning theories (i.e., Classical and Operant conditioning) also

play a part in a GPCSL perspective in that they are based in a behavioralist perspective. They are focused on the role of associative learning and stress the importance of environmental factors in shaping criminal behaviour. These particular perspectives often focus on specific conditions that either promote or inhibit criminal behaviour. They differ from psychodynamic theories in that they do not assume that humans have a natural or innate impulse to act antisocially but instead criminal behaviour is learned via operant or classical conditioning. This is important for a GPCSL perspective in that it takes into account the internal or external factors that may be reinforcing or punishing criminal behavior.

Social learning theories are a form of learning theories that focus on the learning that takes place by observing others being either reinforced or punished for prosocial or antisocial behaviour. In other words, it emphasizes the role of vicarious conditioning in the process of acquiring criminal behaviours, and focuses on the cognitive mechanisms that occur in social settings (Brown et al., 2017). Two of the most common social learning theories regarding crime are Edwin Sutherland's Differential Association (DA) Theory (Sutherland, 1947) and Ron Aker's Social Learning Theory (Akers, 1973). What makes Sutherland's Differential Association Theory different from the contemporaneous psychodynamic theories of the time was that DA theory rejected the idea of internal factors as being the driving forces responsible for crime. Instead, for DA theory crime was a result of learning that occurred in social environments (e.g., social interactions and communication) and that overall one's general likelihood of engaging in criminal behaviour was a balance between the contact that one had with both prosocial and antisocial attitudes or others who might influence their behaviour. In other words, it was exposure to the norms (i.e., favourable or not towards following the law) of the particular groups in which one spent their time that would weigh into the likelihood that one would commit crime.

Sutherland's DA was built around nine testable principles (Table 1.5). DA theory stands out for laying part of the foundation of a contemporary GPCSL perspective because it received a decent amount of empirical support and that some of its core concepts (i.e., antisocial associates and attitudes) have currently emerged as some of the best predictors of reoffending today (Andrew & Bonta, 2015). Brown et al. (2017) noted that DA theory was not without its criticisms and many noted that the theory lacks a clear operationalization of the conditions that were favourable or not favourable to crime and that it was hard (if not impossible) to quantify certain variable of interest (e.g., prestige or influence of certain groups over others in a person's life).

Table: 1.5

Principles of Sutherland's Differential Association

	Nine Principles of Sutherland's Differential Association
1	Criminal behaviour is learned.
2	Criminal behaviour is learned in interaction with other persons in a process of communication.
3	The principle part of the learning of criminal behaviour occurs within intimate personal groups.
4	When criminal behaviour is learned, the learning includes: (a) techniques of committing the crime and (b) the specific directions of motives, drives, rationalizations, and attitudes.
5	The specific directions of motives and drives is learned from definitions of the legal code as favourable or unfavourable.
6	A person becomes delinquent because of an excess of definitions favourable to violations of the law of definitions unfavourable to the law.
7	Differential associations vary in frequency, duration, priority, and intensity.
8	The process of learning criminal behaviours by association within criminal and anti-criminal patterns involves mechanisms that are involved in any other learning.
9	Though criminal behaviour is an expression of general needs and values, it is not explained by those general needs and values since non-criminal behaviour is an expression of the same needs and values.

Sutherland (1947) and Brown et al., (2017)

Ronald Akers and Robert Burgess sought to address many of the identified shortcomings in Sutherland's Differential Association theory and combined the principles of Sutherland's (1947) work and the principles of operant conditioning in order to propose the Differential Association-Reinforcement Theory (Burgess & Akers, 1966). In this theory, they suggested that operant conditioning was the process responsible for acquiring criminal behaviour. Akers' would eventually go on to rework this theory and propose a Social Learning Theory of Crime in 1973 (Akers, 1973). In this theory he presented a more general theory of crime (i.e., it expanded its focus passed direct environmental consequences of behaviour) and suggested that crime is learned via group interactions by both operant and vicarious learning (Brown et al., 2017). In other words, it is the reinforcement received for not only engaging or not engaging in crime but also how a person is vicariously reinforced when they see someone else being reinforced for or against their antisocial behaviour and then imitating that particular behaviour. This theory reinforced the importance of considering those variables within an individual's social environment in influencing behaviour.

Some of the most effective modern day assessment tools for predicting reoffending are based on the variables that emerged from the research conducted within the theories mentioned above. Indeed as Andrews and Bonta (2015) state, "the research is clear: personality factors such as weak self-control are best combined with assessments of attitudes and associates in order to enhance predictive accuracy" (p. 130). The GPCSL perspective builds upon the research from various approaches in attempts to better understand the personal, interpersonal and community supports for human behaviour that are either favourable or unfavourable to crime.

The specific theory upon which most modern day research, treatment, and rehabilitative approaches are based is the Personal, Inter-personal, and Community Reinforcement (PIC-R) theory of crime proposed by Andrews and Bonta (Andrew & Bonta, 2015; Ward, Melser, & Yates, 2007). The PIC-R authors acknowledge that it is "one example" of the GPCSL perspective on deviant behaviour (Andrews & Bonta, 2015, p. 139). The PIC-R is centered around principles such as: (a) occurrences of deviant and nondeviant behaviour are under antecedent and consequent control, (b) inter- and intra-individual variations in the likelihood of a certain behaviour (deviant or nondeviant) are due to variations in the rewards and costs for that behaviour, (c) the control effects of the antecedents and consequences are learned through the interaction with the environment (Andrews & Bonta, 2015, p. 139). A cursory glance over the principles of the PIC-R make evident the manner in which it reflects the cognitive, social learning and behavioral approaches from which it was built. Andrews and Bonta (2015) summarized it well when they stated "the PIC-R… encompasses the contributions of many theories" (p.155).

The GPCSL broadly and PIC-R more specifically are the theoretical sources from which the Risk-Needs-Responsivity model (RNR; Andrews, Bonta, & Hoge, 1990) is based. In short, the RNR model is a model of offender rehabilitation that recognizes that rehabilitative efforts will be the most effective when matched to an offender's risk level, when criminogenic needs (i.e., dynamic risk factors; factors that when changed are associated with some change in subsequent criminal behavior) are targeted, and when responsivity factors are addressed. The variables identified in the research above and the perspective taken in the RNR model are some of the best understood ways in which criminal behaviour can predicted and/or be modified.

Considering the theoretical background of risk assessment has been explored in some detail, the considerations of modern day risk assessment will be explored next.

1.6.3 Considerations in Risk Assessment

While structured risk assessment may be an acknowledged improvement over pure clinical judgment, it is not without its criticisms and shortcomings. For example, the debate mentioned above regarding gendered pathways to crime, criminal career trajectories and risks for reoffending are relevant to the application of risk assessment measures as these topics may impact the factors involved in the predictive accuracy of these tools. For example, the Risk-Needs-Responsivity model (RNR; Andrews, Bonta, & Hoge, 1990) would argue that this principle applies equally to males and females. The RNR model holds that while there may be gender differences in general, the gender similarities in the predictive validity of some risk/needs factors outweigh any gender differences. For the RNR model, it would be expected that there would not be any significant differences in the risk factors for males and females and that this should not impact the predictive accuracy of risk assessment measures so long as they are normed correctly. Further support for this idea comes from Andrews and Dowden (2006) who conducted a meta-analysis and found support for the idea that the risk principle applied to both women and youth. Folsom and Atkinson (2007) also found that the same variables that predicted for men also predicted for women and as such it should follow that actuarial tools that work to predict for men should also predict for women.

Although the RNR model is well researched and very well established, there are some within the critical feminist literature that would argue that males and females do not share the same risks for crime or predictive factors (Wattanaporn, & Holtfreter, 2014; Daly, 1992, 1994; Daly, & Chesney-Lind, 1988). These authors would argue that females have a unique pathway to

crime from their male counterparts and that they possess a unique set of criminogenic needs. They claim that increased exposure to experiences such as motherhood, traumas, sexual and domestic abuse are distinct factors that might predict criminal activity for females and that these unique factors may also impact females' ability to desist in criminal activity and impact rates of reoffending. Thus, many would argue that factors such as these should be more closely considered and female specific risk assessment measures are needed as these special considerations could impact the predictive accuracy of male focused tools. A conclusion to this area of investigation is outside the scope of this study. However, Nicholls, Cruise, Greig, Hinz (2015) who have studied female offenders extensively, noted that it could be argued that the general acceptance of gender neutral assessment, at this time, may be the most appropriate course of action, but they did not exclude the possibility that with more research it might be shown that the addition of gender specific factors may improve risk assessment. For a more detailed review of this topic de Vogel and Nicholls (2016) provide a substantial overview and recommendations in their article to researchers, policy makers and practitioners who want to further their understanding of this issue.

A similar concern also mentioned above applies to mentally ill offenders and whether the same risk factors predict their recidivism. Bonta, Law and Hanson (1998) found that within mentally disorder offenders, the central eight dynamic risk factors (i.e., history of antisocial behavior, antisocial personality pattern, antisocial cognition, antisocial associates, family/marital circumstances, school/work, leisure/recreation, and substance abuse) were better predictors than clinical factors (e.g., mood, depression, major psychotic symptoms) for both general and violent recidivism. Bonta, Blais, and Wilson, (2014) conducted a more recent meta-analysis that supported the same interpretation.

While Bonta et al. (1998) found that the central eight were better predictors than clinical factors this discussion is not as straight forward or as clear cut as many might assume. A meta-analysis by Douglas, Guy and Hart (2009) took a nuanced look at the association between one particular clinical factor (i.e., psychosis) and violence. What these researchers found was that the question is not so much as "are individuals with psychosis more likely to be violent than individuals without?" but instead "what particular symptoms of psychosis, under which situational circumstances, and in combination with which personal or situational factors are associated with increased or decreased risk of various kinds of violence?" (Douglas et al., 2009, p. 696). This more nuanced examination of psychosis as a predictor for violence revealed that in certain cases when compared to certain groups, psychosis was found to have a small association with violence.

The implications of the findings of the Douglas et al. (2009) meta-analysis are particularly relevant to the prediction of risk in an NCRMD population in that the authors conclude that psychosis is neither necessary nor sufficient for a determination of high risk; however, they argue that psychosis should be evaluated in all violence risk assessments and furthermore that that each case must be examined in context. For instance, multiple moderators can either increase (e.g., substance use) or decrease the likelihood of violence (e.g., being compared to those with an ASPD diagnosis). Plus, those who experience psychosis are more likely to be violent with someone within their family unit. All of this information would be important considerations for review boards and is particularly salient for the determination of risk for those individuals found NCRMD. Thus, clinical factors such as diagnosis should not be ignored when conducting risk assessments and the reader is encouraged to consult Douglas et al. (2009) for further guidance on the matter.

Another important consideration not yet mentioned in the use of risk assessment measures is base rates. Base rates are an important consideration because the rates of false positives/negatives, true positives/negatives, and the magnitudes of the association between criminal behavior and the risk predictor are all influenced by base rates. Violence is a low base rate event and low base rate events can be difficult to predict. Andrews and Bonta (2015) discussed how nonviolent offending tends to have a base-rate in the range of 40-60% whereas by comparison violent offending is closer to 10-20%. Sexual offending has an even lower base-rate of around 5%. Deviations from base rates of 50% will impact prediction and thus must be considered when examining the predictive accuracy of risk assessment measures (Babchishin & Helmus, 2015). At one time, it was thought that violence base rates were too low and that violence prediction was a futile endeavor (Cocozza & Steadman, 1976). Luckily, many researchers since that time have challenged this idea and over time perspectives have shifted. Andrews and Bonta (2015) capped this old time perspective by arguing anyone who demands perfect predictive accuracy from risk assessment measures is both unreasonable and impossible.

Another consideration regarding forensic risk assessment is matching the purpose of the assessment to the measures used. In other words what is the purpose of the assessment: To predict violence? General recidivism? Violence risk level after treatment? The question of "what measures will best respond to that purpose" must be considered when decided upon a risk assessment battery. Yang, Wong and Coid (2010) demonstrated that nine commonly used risk assessment measures all significantly predicted violence and that none of them outperformed their counterparts. Thus, the authors argued that risk assessment measures are important tools and need to be used in context and in a holistic manner if practitioners want to use them in the most ethical and effective manner.

It is also an ethical obligation for practitioners to maintain knowledge and understanding of the assessment methods they engage in their assessments. Both the American Psychological Association and the Canadian Psychological Association have certain standards to which psychological assessors must abide (American Psychological Association, 2013; Canadian Psychological Association, 2017). Psychologists are obligated to ensure that the measures they choose for their assessments have been validated on the population being assessed or that they can demonstrate sound clinical judgement and decision making in their justification of choice of tools. This is particularly relevant for psychologists as there have been court challenges regarding specific tools and their use within forensic and correctional populations (See Ewert v. Canada, 2018)

1.7 The Violence Risk Appraisal Guide (VRAG) Family of Risk Assessment Measures

The Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993) is a widely-used violence risk assessment tool. The VRAG is generally accepted as an example of a second generation risk assessment measure and is atheoretical in its approach to item selection. It was the first in a family of empirical actuarial measures developed specifically to assess risk of future violence in adult offenders and forensic psychiatric patients and has been shown to predict violent and general criminal recidivism (Harris, Rice, Quinsey, & Cormier, 2015). It was developed and published more than two decades ago and since that time has been independently validated more than 60 times in correctional and psychiatric samples in several countries (Rice, Grant, Harris, & Lang, 2013). Results from various meta-analyses and systemic reviews also support the VRAG's predictive validity for the outcome of violent recidivism (Campbell, French, & Gendreau, 2009; Hanson & Morton-Bourgon, 2009; Yang, Wong, & Coid, 2010) Furthermore, there is evidence that the VRAG is used regularly in the United States, and in other

countries around the globe (Singh et al. 2014; Cox et al. 2018). In Canada, it is often used with those individuals who have been found Not Criminally Responsible on Account of Mental Disorder (NCRMD) to assist with review board decisions (Wilson, Crocker, Nicholls, Charette & Seto, 2015).

The sample used in the construction of the Violence Risk Appraisal Guide (VRAG) was mostly violent men from Ontario's provincial corrections and forensic mental health systems. As such, mentally disordered offenders were a part of the construction and validation sample, including a significant portion that were found Not Criminally Responsible on Account of Mental Disorder (NCRMD; Harris, Rice, Quinsey, & Cormier, 2015). The VRAG has not traditionally been used with females although some studies have extended its use to a female population and found supportive evidence for this (Eisenbarth, Osterheider, Nedopil, & Stadtland, 2012; Coid et al. 2009; Harris, Rice, & Cromier, 2002). These findings were tempered by the results of Hastings, Krishnan, Tangney, and Stuewing (2011) who advised prudence in the use of the VRAG with female samples due to the low base rates of female violence. Results from Gearaghty and Woodhams' (2015) systematic review also concluded that caution must be exercised in applying the VRAG within female populations. Furthermore, the low base rate for recidivism in general for the NCRMD population, plus the low base rate of violent offending for females in general, does raise questions regarding gendered differences in the predictive validity of the VRAG with NCRMD patients. Possible gendered differences in prediction are not exclusively an issue related to the VRAG. For instance, Nicholls, Ogloff, and Douglas (2004) found that there were gendered differences in the predictive validity of The Historical Clinical Risk Management - 20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997), the Psychopathy Check List – Screening Version (PCL-SV; Hart, Cox, & Hare, 1995) and the

Violence Screening Checklist (VSC; McNiel & Binder, 1994). The systemic review done by Gearagty and Woodhams' (2015) found similar findings in that multiple commonly used risk assessment tools had questionable predictive validity when used within a female population.

Recognizing that sexual offenders may have both common and unique risk factors of reoffending, the Sex Offender Risk Appraisal Guide (SORAG; Quinsey, Rice, Harris, & Cormier, 1998, 2006) was the next measure created in the VRAG family of risk assessment tools. The SORAG is a 14-item tool developed in the same manner as the VRAG and is used to assess risk for general violence among sexual offenders. The SORAG demonstrated similar predictive properties to its predecessor the VRAG and a more fulsome review of the literature focused on the SORAG can be found in Harris, Rice, Quinsey and Cormier (2015).

1.7.1 The Violence Risk Appraisal Guide Revised

The Violence Risk Appraisal Guide Revised (VRAG-R; Rice, Harris, & Lang, 2013) is the newest member of the VRAG family. It is shorter and easier to score and has been shown to predict dichotomous violent recidivism with a high level of predictive accuracy much like its predecessors (Rice et al., 2013). The VRAG-R represents a collapsing of the original VRAG and SORAG into one common violence risk assessment measure. It maintains 12-items that are differentially weighted, much like the original VRAG. The items encompass a collection of static clinical, criminal and behavioral history, and demographic variables that once summed are arranged into nine risk bins. The risk bins are arranged into deciles with relatively equal proportions of cases within. The VRAG-R differentiates itself from its predecessor in the that it no longer includes the Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003) total score as one of its items nor does it utilize variables that were no longer predictive (e.g., female victim or victim injury) or that required formal diagnosis (e.g., personality disorder or schizophrenia

diagnosis). Instead, the VRAG-R utilizes the PCL-R anti-social facet score, the facet which was the most predictive in the construction sample used to develop the VRAG-R.

At the time of writing no study that examined the predictive accuracy of the VRAG-R in an NCRMD population could be found. This was also true for studies that specifically looked at the use of the VRAG-R by gender. That said, other independent validations have begun to be completed with the VRAG-R in a variety of populations and contexts. For example, Glover et al. (2017) examined the use of the VRAG-R in a correctional sample, and both Gregório, Hertz, Rettenberger, and Eher (2019), and Olver and Sewall (2018) examined the use of the VRAG-R within a sexual offender population. Hogan and Olver (2019) noted that the VRAG-R's predictive ability was comparable to its predecessor although in Hogan and Olver (2016) the VRAG-R had less success in predicting inpatient aggression. This finding was in contrast to previous studies done with the VRAG-R's predecessor which indicated that the VRAG had some success in accurately predicting inpatient aggression. Gregório, Hertz, Rettenberger, and Eher (2019) also showed the cross-national utility of the VRAG-R as they validated the tool in the German language and showed empirical support for its use within German speaking populations.

CHAPTER TWO: THE CURRENT STUDY

2.1 Rationale

There is increasing recognition that more research needs to be done regarding the individuals in Canada found Not Criminally Responsible on Account of Mental Disorder (NCRMD) and indeed there are valuable research projects underway across the country (e.g., The National Trajectory Project; Crocker, Nicholls, Seto, & Côté, 2015). Yet Alberta remains a jurisdiction wherein little is known about its NCRMD population. This is particularly true for the female individuals found NCRMD within Alberta. Comparatively, while the Violence Risk Appraisal Guide (VRAG) is utilised around the world and has been independently validated multiple times since its inception almost two decades ago, the Violence Risk Appraisal Guide Revised (VRAG-R) has yet to have the same extensive independent validation as its predecessor (Singh et al, 2014; Rice, Grant, Harris, & Lang, 2013). In addition, studies have been conducted with the VRAG-R within certain populations, for example, sex offenders (Olver and Sewall 2018; Gregório Hertz, Rettenberger, & Eher, 2019) and general offenders (Glover et al., 2017) yet as of this writing no studies have been conducted focusing exclusively on its use within an NCRMD population.

2.2 Research Objectives and Hypotheses

In light of this, the study had two main objectives. First, it examined individuals found NCRMD within Alberta for gender differences. Specifically differences regarding characteristics such as: sociodemographic (e.g., age, sex, education), clinical (e.g., diagnosis at time of index), and criminological (e.g., index offense, rates of conditional discharge, absolute discharge, and recidivism). Much of this portion of the research was exploratory due to the dearth of research

focused on Alberta. However, considering research done in other Canadian provinces (e.g., Nicholls et al., 2015), it was hypothesized that:

1. Males and females will be found to have distinct psychosocial, clinical, and criminological profiles.

Second, it examined the predictive accuracy of the VRAG-R within the Alberta NCRMD population, specifically, discrimination (i.e., the extent to which VRAG-R scores can accurately discriminate recidivists from non-recidivists), and calibration (i.e., what extent recidivism rates are associated with VRAG-R scores and to what extent observed recidivism rates over a defined follow up period are aligned with expected rates predicted through logistic regression). Analyses included comparisons by gender within the Alberta NCRMD population and with the VRAG-R normative sample. Once again while much of this research was exploratory, it was hypothesized that:

2. Females found NCRMD will have lower VRAG-R risk scores and bin number frequency distributions than their male counterparts
3. The VRAG-R will have good predictive accuracy for general and violent recidivism over various fixed and global follow ups (i.e., strong discrimination and calibration properties)

2.3 Method

2.3.1 The Alberta NCR Project Database

The Alberta NCR Project database ("the database") provided the foundation for this study. The database included every person who has been found NCRMD and who has come under the Alberta Review Board's (ARB) jurisdiction. The database is housed at Alberta Hospital Edmonton (AHE) in Edmonton Alberta, Canada and is managed and maintained by Dr. Andrew Haag. It currently has ethics approval through The University of Alberta's Research

Ethics Office that covers assessing prediction in risk assessment instruments. The database currently spans 79 years with the earliest file on record dating back to 1941. The database is continually updated for new entries and recidivism data when available. The information within this database is also continually being used for other ongoing research projects related to risk and violence prediction.

The database contained a wide array of information including, but not limited to, sociodemographic information (e.g., age, sex, race, education level, marital status), clinical information (e.g., psychiatric diagnosis at time of disposition) and criminological information (e.g., date of NCRMD verdict, location of offence(s), index offence information, yearly warrant status, conditional discharge status, and recidivism data). The database also housed a variety of information that was not accessed for this study (e.g., other risk assessment measure outcomes). Sources of information for the database came directly from ARB dispositions, AHE internal database of demographic information on patients, official reports, and the Forensic Assessment Community Services (FACS) psychological and psychiatric reports. An in-depth description of the construction and information contained within the database can be found in Haag, Cheng, and Wirove (2015).

2.3.2 Participants

Participants for this study were chosen from *The Alberta NCR Project* database. This study was archival in nature and as such there was no direct contact with any of the participants. The participants had all been found NCRMD and come under the jurisdiction of the ARB between 1941 and 2018. As of October 2018, there were 574 identified cases of persons found NCRMD within the database. Most of these cases were male (83.1%; $n = 477$) with females

representing a much smaller proportion (16.9%; *n* = 97). The participants were largely white, early middle aged, single, and had not completed high school.

Due to the desire for population level analysis all participant's files were considered for inclusion. Certain participants were excluded due to an inability to examine their files (e.g. unable to be located). Exclusion criteria was also based on file information availability as some of the older files did not contain the needed information or contained inadequate detail to score a VRAG-R.

2.3.3 Measure

The Violence Risk Appraisal Guide Revised (VRAG-R; Rice, Harris, & Lang, 2013) is a 12-item violence risk assessment tool statistically developed from a violent mentally disordered offender population in Ontario, Canada (Harris, Rice, Quinsey, & Cromier, 2015). This tool is a more streamlined and easier to use version of the original VRAG. Similar to its predecessor, items scored are historical and static in nature (e.g., lived with parents until age 16, elementary school maladjustment, conduct disorder before age 15, and age at index offence) and individual items have a positive or negative score depending on the direction of their correlation with recidivism. Individual items are then summed and total scores (range: -34 to +46) are divided into 9 risk bins. During its construction and validation, approximately half of the sample consisted of men who were found NCRMD, with the rest being men who had committed violent crimes and were being assessed in a maximum security psychiatric hospital.

Regarding the prediction of dichotomous violent recidivism, the VRAG-R demonstrates large effect sizes. The development sample contained an AUC value of .76 (*n* = 957) followed by the validation sample with an AUC value of .75 (*n* = 300). When the entire sample was

examined (n = 1,261) an AUC value of .76 was found. Thus, the VRAG-R is in the same range of accuracy as its predecessors (Rice, Harris, & Lang, 2013).

2.3.4 Recidivism Criteria

The Alberta NCR Project database defined recidivism via conviction only and sorted recidivism into four categories: (a) sexual recidivism, including any sexual conviction post NCRMD status, (b) major violent recidivism, including assault causing bodily harm, aggravated assault, assault with a weapon, homicide, and attempted homicide, (c) violent recidivism, including sexual violence and robbery, and (d) general recidivism, which included any criminal code conviction. Conviction data was taken from FPS sheets acquired from the Canadian Police Information Centre (CPIC) for the entirety of the database up to 2015. For the purposes of this study, recidivism was examined within the conventional manner wherein sexual recidivism was nested within violent recidivism and violent recidivism was nested within general recidivism. Sexual recidivism was not examined specifically for two reasons: (1) only four sexual recidivists existed within the entire Alberta NCRMD population at the time of analysis and (2) there are other more appropriate tools to determine risk for sexual recidivism than the VRAG-R. That being said, sexual recidivism was still counted within general and violent recidivism.

2.3.5 Procedure

Regarding the first research objective, this writer and other undergraduate research assistants (RAs) collected data from files located at AHE, FACS, and the ARB. This part of the study is a culmination of a long data collection phase from the existing *Alberta NCR Project*. As this data had been in the process of being collected for years prior to this study's initiation much of this portion of the study was secondary use of data that was already collected for the *Alberta NCR Project*. Dr. A. Haag, the researcher at the helm of the Alberta NCR Project provided

training to research assistants in how to read files and gather the required information. Dr. A. Haag or this writer were always available in person or via text for questions should any come up during the data collection process. The data collected was housed within *The Alberta NCR Project* database (described above).

Regarding the second research objective, a VRAG-R was scored from file information for each of the individuals within the database. This writer and RAs were trained in the scoring of the VRAG-R. Training was provided by either/both Dr. A. Haag or through online training provided by the Global Institute of Forensic Research. This online training was provided by one of the co-creators of the VRAG-R. RAs completed regular scoring validity checks and also had access to senior researchers for questions. To assess inter-rater reliability, 30 files (5.2% of the population) were randomly selected and coded. In the files selected, excellent IRR was found, $ICC_{1,3} = .99$, $p < .000$.

2.2.5 Planned Analysis

Data analysis for the first research objective entailed the use of descriptive statistics, independent and paired sample t-tests, and chi-squared analyses. Independent samples t-tests were used for examining continuous variables (e.g., age at index offense, time spent under the ARB). For those variables that were categorical, a chi-squared test of independence was performed. Regarding the chi-squared analyses in which it was a 2 by 2 analysis (e.g., male/female and yes/no) a Yates Continuity Correction was taken into account in order to compensate for the overestimation of the chi-squared value when used with a 2 by 2 table (Pallant, 2013). A paired samples t-test was conducted in order to examine time to conditional discharge and time to absolute discharge for those who had received both via gender.

For the second research objective, data analysis proceeded over several steps. Where possible all the analyses regarding the VRAG-R were done both via the population as a whole and by gender. To start, descriptive statistics and frequency distributions of VRAG-R scores and their respective bins were examined. This was completed for both scores via the entire population and separated by gender. Next, the discrimination properties of the VRAG-R for predicting 5-year, 10-year, and global violent and general recidivism were examined in order to determine how well the VRAG-R was able to differentiate recidivists from non-recidivists. This was accomplished via examinations of correlations and receiver operator characteristic (ROC) analyses. ROC analyses generate an area under the curve or AUC statistic that can range from 0 to 1. This value represents the probability that a randomly selected recidivist will score higher on a given risk tool than a randomly selected non-recidivist. AUC results of .50 indicate chance levels of predictive accuracy whereas AUC values of .556 represent small, .639 medium, and .714 large effect sizes (Rice and Harris, 2005). In general the larger the effect size the more confidence one can have in the risk assessment tool being utilized.

Finally, calibration analyses of the VRAG-R were undertaken. Calibration analyses examine what recidivism rates are associated with what VRAG-R scores and this analysis was also conducted over several steps. First, the procedure outlined in Olver and Sewall (2018) was followed for the initial portion of the calibration analyses. Logistic regression was used to model 5-year and 10-year estimates of violent recidivism with specific VRAG-R scores. Logistic regression generates a constant (B_0) which is the log odds of the recidivism base rate, and regression coefficients (B_1), each representing the percent increase in the odds of a given outcome between adjacent scores on the measure. Specific scores on the tool under examination, in this case the VRAG-R, in conjunction with base rate information from the sample, can be

used to estimate a specific score through the use of a logistic linking function: $e^{B0+B1xScore}/(1+e^{B0+B1xScore})$ (Tabachnick and Fidell, 2007). After that as a point of further exploration, a visual representation of these results was juxtaposed with the reported actual rates of violent recidivism observed for each of the nine risk bins of the VRAG-R.

As a last step to analyzing the calibration properties of the VRAG-R within a NCRMD population, an E/O index (i.e., expected/observed index) was computed using the method outline in Hanson (2017) and demonstrated in Olver and Sewall (2018). An E/O index results in a ratio of expected number of recidivists to the actual observed number of recidivists. Should a scale have perfect calibration the E/O index will be 1. If an E/O index is below 1, this value would indicate that the risk measure under investigation has underpredicted recidivism. The opposite is true should the E/O index be over 1. A value over 1 would indicate that the risk measure under investigation has over-predicted risk. The E/O index value is also able to indicate the degree to which the measure has over or under estimated recidivism (e.g., an E/O index value of 3 would indicate that the measure has overestimated three times the number of recidivists; an E/O index of 0.6 would indicate that the scale only predicted 60% of the number of observed recidivists). Confidence intervals can also be calculated for E/O indexes (see Hanson 2017 for formula and steps). A confidence interval that does not include 1 would indicate that the expected and observed recidivism rates reported were significantly different ($p<.05$). In other words, good calibration includes upper and lower confidence intervals that include 1 (i.e., indicating that there is not a significant difference). For this study an E/O index was conducted to compare the observed rates of 5-year recidivism for the Alberta NCRMD population to the expected rates of recidivism for the VRAG-R normalization sample.

Lastly, as a final point of investigation the discrimination and calibration (excluding an E/O index) analyses were computed for the four most robust diagnostic categories within the database (e.g., any psychotic disorder, any mood disorder, substance use disorder [present or ever] and antisocial personality disorder including traits). All of the same processes above were followed and results were computed for these categories for the population as a whole.

2.2.6 Ethics

Ethics approval was obtained through the University of Saskatchewan Research Ethics Board (June 2018). The *Alberta NCR Project* which built the database at the foundation of this study has ethical approval through the University of Alberta's Research Ethics Office (REO) and as such ethical approval was also obtained through the University of Alberta's REO office (June 2018).

CHAPTER 3: RESULTS

As of October 15, 2018, there were a total of 574 persons who had been identified for inclusion into the *Alberta NCR Project's* database. The earliest case dated back to 1941. All files were examined for inclusion and exclusion criteria was based on either available file information or available recidivism data which was available up to 2015. Missing information was excluded from analysis. Missing data was not assumed to be missing for a systematic reason, data were either missing due to age of file (i.e., older files contained less information, ink on onionskin paper was illegible) or assessor's discretion (i.e., the reports on file did not comment on needed information).

3.1 Gender Differences within the Alberta NCRMD Population

3.1.1 Socio-Demographic Characteristics

Consistent with the previous NCRMD research conducted in Canada, women represented a minority (N = 97, 16.9%) of the population. No significant differences existed between the genders for age at time of index, marital status, and ethnicity; however women, compared to men, were significantly more likely to have completed high school (Table 3.1). In short, this was a largely White, younger-middle age sample, with the majority reporting to be single and to have not completed high school.

3.1.2 Mental Health Characteristics

Information regarding diagnosis at the time of NCRMD verdict was assessed by either a registered psychologist or psychiatrist. Analyses indicated some significant associations between gender and mental health diagnosis. Among women, compared to men, there were significantly higher base rates of mood disorder (50% vs. 36.1%, respectively), and borderline personality disorder (8.8% vs. 2.7%, respectively), while men had higher rates of antisocial personality

disorder (28.2% vs. 7.6%, respectively) and substance use disorder (59.6% vs. 40.7%, respectively). A significant association between substance use disorder and gender existed both when the categories included “yes”, “no”, and “in remission” as well as when this was further collapsed into “no”, and “yes/ever”. These findings were in line with hypothesized outcomes and previous NCRMD research. It is also interesting to note that the diagnoses mentioned above have also been noted to have gendered prevalence rates within the general population (American Psychiatric Association, 2013).

No significant association between diagnosis and gender existed for categories such as dissociative disorders, obsessive compulsive disorder, posttraumatic stress disorder, phobias, attention-deficit/hyperactivity disorder (ADHD), anxiety disorders, and other personality disorders.

3.1.3 Criminological Characteristics

An examination of the criminological profile of the population demonstrated no significant associations between gender and the number of prior NCRMD verdicts nor between gender and the NCRMD index offence for the categories of homicide, attempted homicide, sexual/violent, and all others (Table 3.2). Males, however, had a significantly higher number of mean number of sentencings (excluding the index offense, $M = 2.93$, $SD = 5.08$) than females ($M = 0.73$, $SD = 1.97$), which was a medium sized effect ($d = .57$).

3.1.4 Conditional Discharge

A person is granted a conditional discharge when the review board, which has been tasked with ensuring public safety, concludes that the individual may be safely supervised within a community setting with certain restrictions placed upon their liberty. No significant differences

were found between males and females on length of time to conditional discharge or the percentage of those who received a revocation of their conditional discharge (Table 3.2).

Table 3.1

Alberta NCRMD Population – Sociodemographic and Mental Health

Characteristic	Female n (%)/M (SD)	Male n (%)/M (SD)	t/χ^2	*n*	*p*	*φ*
Age	37.28 (11.66)	34.81 (12.77)	-1.73	565	.084	-
Education						
No high school	52 (58.4)	320 (70.0)	3.97	547	.046	-.09[a]
Highschool and above	37 (41.6)	138 (30.0)				
Marital status						
Single	71 (76.3)	401 (84.6)	3.23	567	.072	-.08[a]
Married	22 (23.7)	73 (15.4)				
Ethnicity						
White	73 (78.5)	337 (71.6)	3.31	567	.191	.08
Indigenous	9 (9.7)	43 (9.1)				
Other	11 (11.8)	94 (19.8)				
Mood disorder						
Absent	44 (50)	239 (63.9)	5.23	462	.016	.11[a]
Present	44 (50)	135 (36.1)				
Psychotic disorder						
Absent	14 (15.4)	72 (16.1)	0.00	539	.870	.01[a]
Present	77 (84.6)	376 (83.9)				
Substance use disorder						
Absent	48 (59.3)	152(40.4)	8.86	457	.003	-.145[a]
Present	33 (40.7)	224 (59.6)				
Antisocial personality disorder			13.78	437	< .001	-.19[a]
Absent	73 (92.4)	257 (71.8)				
Present	6 (7.6)	101 (28.8)				
Borderline personality disorder			4.96	418	.011	.13
Absent	73 (91.3)	329 (97.3)				
Present	7 (8.8)	9 (2.7)				

Note: [a] with Yates Continuity Correction – computed for a 2x2 table

3.1.5 Absolute Discharge and Time Under the Alberta Review Board

When the review board determines that a person is no longer a danger to public safety, the review board is legislatively required to grant an absolute discharge. There were no significant differences between men and women on mean length of time (in months) to be granted an absolute discharge, nor was there any significant association between receiving an absolute discharge and gender (Table 3.2). Moreover, for those still under Alberta Review Board jurisdiction as of October 15, 2018, there was no significant difference in mean length of time spent under jurisdiction of the review board between men and women.

3.1.6 Time to Conditional Discharge and Absolute Discharge

The time to conditional discharge and mean time to an absolute discharge for those who had received both was also examined. A paired-sample t-test demonstrated that the mean amount of time to absolute discharge was significantly longer than the mean time to conditional discharge (M = 31.4 months, SD = 30.6) for the group as a whole, a medium effect, d = .53. The same test was run for each gender separately; significant differences were found between the mean length of time to absolute discharge and conditional discharge for both males (M = 31.0 months, SD = 29.9, d = .51) and females (M = 31.9 months, SD = 33.7, d = .59).

3.1.7 Recidivism

There were only three females who recidivated (3.2%) out of all the females in the population examined in contrast to 92 (19.3%) male recidivists, which was significant (Table 3.2). The three females who recidivated were in their late teens or early twenties when they were found NCRMD, two received conditional discharges, two violently reoffended, all three took more than 10 years to reoffend, and spent a range of 52 to 103 months under the review board

prior to being granted an absolute discharge. A VRAG-R could be scored on only one of the three females due to a lack of information in the others' files.

Table 3.2

Alberta NCRMD Population - Criminological Profile, Conditional Discharge, Absolute Discharge and Recidivism by Gender

Characteristic	Female n (%)/M (SD)	Male n (%)/M (SD)	t/χ^2	*n/df*	*p*	*φ/d*
Prior NCRMD findings			1.37	566	.242	-.07[a]
Yes	0 (0)	12 (2.5)				
No	94 (100)	460 (97.5)				
Index offense			1.18	574	.758	.05
Homicide	19 (19.6)	84 (17.6)				
Attempted homicide	7 (7.2)	47 (9.9)				
Sexual/violent	44 (45.4)	228 (47.8)				
Other	27 (27.8)	118 (24.7)				
Prior sentences	0.73 (1.97)	2.93 (5.08)	7.05	357.45[b]	< .001	.57
Length of time [c] to CD	58.31 (67.31)	61.77 (57.93)	.379	285	.705	-
CD revocations			3.06	513	.217	.08
Yes	7 (8.3)	54 (12.6)				
No	63 (75.0)	280 (65.3)				
Never received CD	14 (16.7)	95 (22.1)				
Time [c] to AD	66.63 (62.84)	70.57 (70.81)	.367	285	.707	-
Gender and AD			.639	500	.424	- .04[a]
Yes	59 (66.3)	251 (61.1)				
No	30 (33.7)	160 (38.9)				
Time [c] under ARB	92.53 (98.46)	119.93 (94.88)	1.44	188	.151	-
Recidivism and gender			13.74	572	< .001	-.16[a]
Yes	3 (3.2)	92 (19.3)				
No	92 (96.8)	385 (80.7)				

Note: [a] with Yates Continuity Correction – computed for a 2x2 table; [b] Levene's test indicated unequal variances ($F = 36.73$, $p < .001$), so degrees of freedom were adjusted from 556 to 357.45; [c] in month. CD = conditional discharge, AD= absolute discharge.

3.2 Predictive Accuracy of the VRAG-R Within the Alberta NCRMD Population

3.2.1 Descriptive Statistics and Frequencies of the Violence Risk Appraisal Guide – Revised (VRAG-R) Scores

Table 3.3 provides the descriptive statistics regarding the total score of the VRAG-R for the Alberta NCRMD population as a whole and by gender. Overall, of the 574 known persons found NCRMD in Alberta's history, 478 (83%) either had a VRAG-R already scored on file or were successfully scored, 96 individuals had files that were either unable to be located or did not contain sufficient information to score the VRAG-R. Omitted items were prorated following the protocol outlined by Rice, Harris and Lang (2013). All scored VRAG-Rs available were used in the analysis.

Table 3.3 provides descriptive statistics for the VRAG-R scores and bin frequencies for the population as a whole and by gender. Overall, the average population total score was low (M = -6.8, SD = 19.0). It was lower risk than that of the sample used by Harris et al. (1993) in the construction of the original VRAG (M = 0.91, SD = 12.9) and by Rice et al. (2013) in the construction of the VRAG-R (M = 3.6, SD = 12.5). In the Alberta NCRMD population there was a significant difference in VRAG-R total scores for males (M = -4.8, SD = 19.1) and females (M = -16.6, SD = 14.9), $t(136.934) = 6.138$, $p < .001$, which approached a large effect $d = .69$.

Over half the population scored within the first four bins, further evidence for the low risk nature of this population in contrast to the VRAG and VRAG-R development samples. A one-way ANOVA determined there was a significant difference between each of the average bin scores overall (F (8, 469) = 3206.44, $p < .001$) and by gender: males (F (8, 389) = 2753.86, $p < .001$) and females (F (7, 72) = 318.22 , $p < .001$).

Table 3.3

Descriptive Statistics and Bin Frequencies for VRAG-R

VRAG-R Bin	Overall		Female		Male	
(Range)	*N (%)*	*M (SD)*	*n (%)*	*M (SD)*	*n (%)*	*M (SD)*
1 (< -25)	109 (22.8)	-28.55 (3.3)	37 (46.3)	-28.86 (3.1)	72 (18.1)	-28.39 (3.4)
2 (-25 to -19)	76 (15.9)	-20.13 (2.1)	9 (11.3)	-20.22 (2.5)	67 (16.8)	-20.12 (2.1)
3 (-18 to -14)	55 (11.5)	-13.51 (1.7)	12 (15.0)	-13.67 (1.5)	43 (10.8)	-13.47 (1.8)
4 (-13 to -8)	56 (11.7)	-6.79 (2.2)	8 (10.0)	-6.25 (2.5)	48 (12.1)	-6.88 (2.2)
5 (-7 to 5)	43 (9.0)	-0.12 (2.0)	3 (3.8)	1.00 (2.0)	40 (10.1)	-0.20 (2.0)
6 (6 to 12)	46 (9.6)	7.43 (2.3)	7 (8.8)	7.71 (2.6)	39 (9.8)	7.38 (2.2)
7 (13 to 18)	25 (5.2)	14.32 (1.9)	1 (0.2)	12.00 (n/a)	24 (6.0)	14.42 (1.9)
8 (19 to 27)	34 (7.1)	21.47 (2.4)	3 (3.8)	22.00 (3.6)	31 (7.8)	21.42 (2.4)
9 (28+)	34 (7.1)	32.22 (3.8)	0 (0)	-	34 (8.5)	32.22 (3.8)
Total	478 (100)	-6.79 (19.0)	80 (100)	-16.61 (14.9)	398 (100)	-4.81 (19.1)
Min, Max	-34, 43		-34, 25		-34, 43	

Nearly half of the female population total scores fell within the first bin, in contrast to half the male total scores being spread between the first four bins. A significant association between gender and VRAG-R bin number assignment was present χ^2 (8, n = 478) = 40.02, p = .000, φ = .29) with females being more strongly associated with having a lower risk bin assignment than males. Small cell *n*s notwithstanding, the results demonstrate the lower risk nature of female NCRMD patients in this population.

3.2.2 Discrimination Properties of the VRAG-R for General and Violent Recidivism

Examination of the discrimination properties of the VRAG-R involves establishing the extent to which VRAG-R scores can differentiate recidivists (i.e., those who reoffend violently or otherwise) from non-recidivists.

3.2.1.1 Overall results for Alberta NCRMD population.

For the Alberta NCRMD population as a whole, for 5-year fixed follow-up (n = 405), 36 individuals (8.8%) received convictions for any new offense (i.e., general recidivism), which included 22 individuals (5.4%) who received convictions for new violent offenses. At ten-year

fixed follow-up (n = 401), 53 individuals (13.2%) received new convictions for any new offense, which included 31 individuals (7.7%) convicted for a new violent offense. For the overall follow-up (n = 476), 71 individuals (14.9%) were reconvicted for any new offense, which included 44 individuals (9.2%) who reoffended violently.

Table 3.4 provides the results of the Receiver Operator Characteristics (ROC) analyses of the VRAG-R for 5-year, 10-year, and overall (unfixed) follow-ups for general and violent recidivism by both VRAG-R total score and bin. Several notable themes arose from these analyses. First, both total score and bin level significantly predicted general and violent recidivism. Second, both total score and bin number demonstrated better predictive accuracy for violent reoffending than general reoffending. Third, following the guidelines laid out by Rice and Harris (2005), AUC magnitudes were consistently within the large or medium to large in magnitude for violence prediction, and AUCs for total score for general reoffending were within the medium to large range and medium range for general recidivism by bin number.

Table 3.4

Discrimination Properties of the VRAG-R for Violent and General Recidivism for 5-year, 10-year, and Overall Follow-up

Recidivism Criterion	Timeframe	AUC	95% CI
Total Score Violent	5-year	.711***	[.604, .818]
	10-year	.714***	[.624, .804]
	Overall	.703***	[.626, .779]
Total Score General	5-year	.682***	[.587, .777]
	10-year	.674***	[.595, .754]
	Overall	.654***	[.586, .722]
Bins Violent	5-year	.691**	[.580, .801]
	10-year	.698***	[.606, .709]
	Overall	.691***	[.613, .770]
Bins General	5-year	.666***	[.571, .762]
	10-year	.663***	[.584,.743]
	Overall	.642***	[.573, .711]

Note: *$p < .05$, **$p < .01$, *** $p < .001$; AUC = area under the curve; CI = confidence interval

3.2.1.2 Results by gender.

Owing to the small number of female recidivists (n = 1 recidivist) with a complete VRAG-R score, formal examination of the discrimination properties of the VRAG-R could only be completed for the male subgroup.

For males, at 5-year follow-up (n = 346), 36 individuals (10.4%) received new convictions for any new offense, which included 22 individuals (6.3%) convicted for new violent offenses. At ten-year follow-up (n = 342), 53 individuals (15.4%) received new convictions for any new offense, which included 31 individuals (9.1%) convicted for new violent offenses. For overall follow-up (n = 389), 70 (18.0%) individuals received new convictions for any new offense which included 43 individuals (11.1%) convicted for new violent offenses.

Table 3.5 provides the results of ROC analyses of the VRAG-R for 5-year, 10-year, and overall follow-up for general and violent recidivism by both VRAG-R total score and by bin. Results closely mirrored the findings reported for the overall population. First, both total score and bin number significantly predicted general and violent recidivism. Second, both total score and bin number demonstrated better predictive accuracy for violent reoffending than general reoffending. Third, per Rice and Harris (2005), VRAG-R AUC magnitudes were broadly in the medium range for violent and general recidivism. Fourth, AUCs tended to be slightly higher for fixed, vs. overall, follow-ups.

Table 3.5

Discrimination Properties of the VRAG-R for Violent and General Recidivism for 5-year, 10-year, and Overall Follow-up (Males Only)

Recidivism Criterion	Timeframe	AUC	95% CI
Total Score Violent	5 Year	.691**	[.579, .803]
	10 Year	.694***	[.600, .788]
	Overall	.672***	[.590, .754]
Total Score General	5 Year	.662***	[.564, .760]
	10 Year	.653***	[.571, .736]
	Overall	.621***	[.549, .693]
Bins Violent	5 Year	.670**	[.555, .784]
	10 Year	.677***	[.581, .773]
	Overall	.659***	[.575, .742]
Bins General	5 Year	.645**	[.547, .744]
	10 Year	.641***	[.559, .724]
	Overall	.608**	[.536, .681]

Note: *$p < .05$, **$p < .01$, *** $p < .001$; AUC = area under the curve; CI = confidence interval

3.2.3 Calibration Properties of the VRAG-R

Calibration examines what recidivism rates are associated with VRAG-R scores, and to what extent observed recidivism rates over a defined follow-up period from one sample or follow-up period are aligned with those rates expected from the normative sample. For the NCRMD population as a whole, the 5-year and 10-year violent recidivism rates were examined.

3.2.3.1 Overall results for Alberta NCRMD population.

Logistic regression was conducted to estimate the rates of recidivism associated with specific VRAG-R scores over a given follow-up. To do this, the Hosmer–Lemeshow goodness of fit tests were all nonsignificant suggesting that the logistic distributions provided a reasonable approximation of violent recidivism rates to warrant modelling. Results of the logistic regression generated the following terms for 5-year (B_0 = -2.856, B_1 = .040, $p < .001$) and 10-year (B_0 = -2.471, B_1 = .041, $p < .001$) violent recidivism. Using a log linking function, these values can be employed to estimate rates of recidivism associated with all possible scores. Figure 3.1 illustrates all possible estimated VRAG-R scores for 5-year and 10-year violent recidivism as estimated through logistic regression.

Figure 3.1

Logistic Regression Estimated 5 and 10-Year Violent Recidivism for all Possible VRAG-R Scores

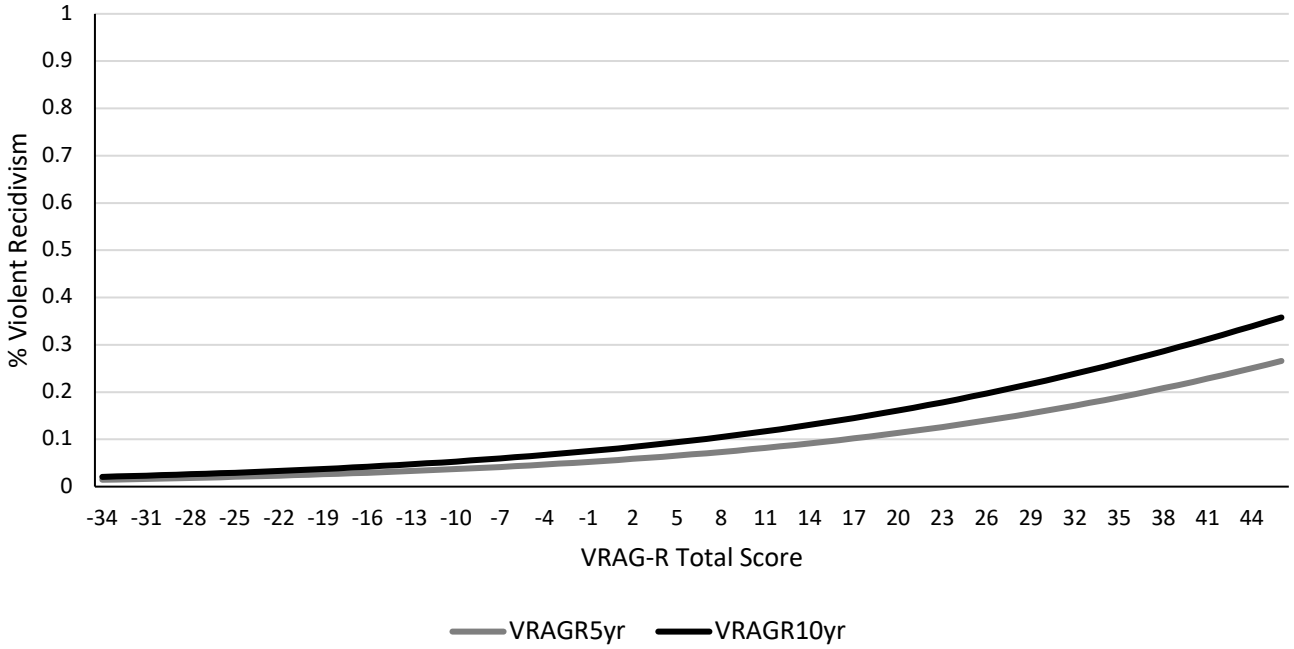


Figure 3.2 represents a visual analysis of estimated VRAG-R scores for 5-year and 10-year violent recidivism juxtaposed with the observed rates of recidivism for the same time periods across the VRAG-R scores. It should be noted that there are some fluctuations in observed rates due to small *ns* of recidivists in some bins, underscoring the utility of logistic regression for estimating rates of recidivism associated with specific risk scores.

Following the procedures outlined in Hanson (2017) and Olver and Sewall (2018) an E/O Index was computed as a formal examination of calibration. Table 3.6 provides the recidivism rates observed in the Alberta NCRMD population compared with those expected from the VRAG-R (2013) normative sample. The results indicated that the VRAG-R scores substantially overestimated risk for the Alberta NCRMD population as all E/O indexes were greater than 1 (i.e., a value of 1 indicates perfect calibration), with some values reaching as high as 15 (Bin 6). Overall, the VRAG-R normative sample from Rice et al. (2013) overestimated risk by 4.6 times. In general, there was poor agreement between the expected and observed recidivism rates. Only two of the E/O index values were non-significant (Bin 1 and 4), given that the confidence interval included 1.0. This demonstrated that for all other values there was a statistically significant difference between the observed and expected rates of recidivism. Data for the 10-year violent recidivism outcomes for the VRAG-R normative sample was not available at the time of investigation (Helmus, personal communication, March 13, 2019) and as such a 10-year E/O index could not be computed.

Figure 3.2

VRAG-R Calibration: Actual Rates of Violent Recidivism for the Nine-Bin Structure and Estimated Rates of Violent Recidivism Associated with Individual Scores over Fixed 5 and 10-Year Follow-Ups.

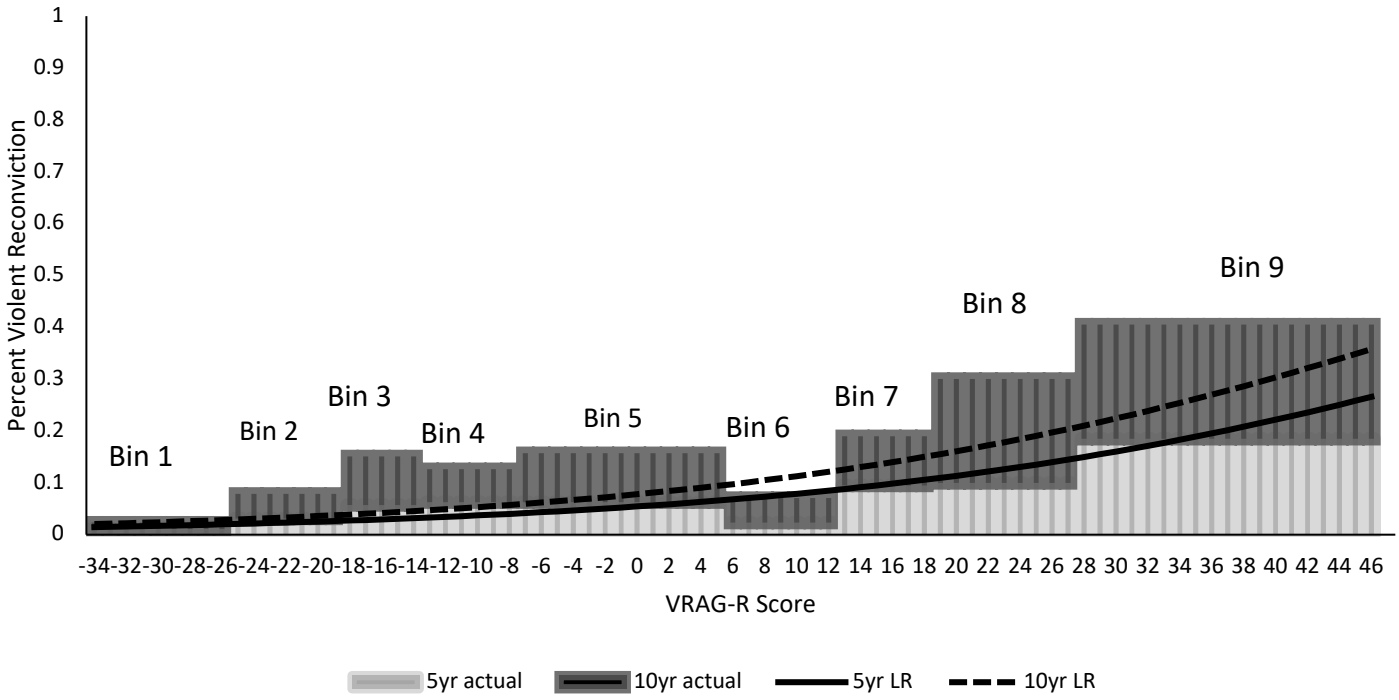


Table 3.6

E/O Index: Five-year Rates of Violent Recidivism for Normative Sample (Rice, Harris, & Lang, 2013) Compared with the Alberta NCRMD Population

Risk bin and score range		5-year violent recidivism					
		Expected rates: VRAG-R norms		Observed rates: Current sample			
	n	%	n	%	n	E/O Index	95% CI
1 (≤ -25)	87	8.0	7.0	1.1	1	7.0	.98, 49.7
2 (-24 to -19)	64	9.0	5.8	3.1	2	**2.9**	**1.89, 4.47**
3 (-18 to -14)	53	18.0	9.5	5.7	3	**3.2**	**1.06, 9.79**
4 (-13 to -8)	48	19.0	6.7	6.3	3	2.2	.73, 6.73
5 (-7 to 5)	32	25.0	8.0	6.3	2	**4.0**	**2.60, 6.16**
6 (6 to 12)	43	37.0	15.9	2.3	1	**15.0**	**2.23, 112.89**
7 (13 to 18)	21	45.0	9.5	9.5	2	**4.7**	**3.06, 7.24**
8 (19 to 27)	30	58.0	17.4	10.0	3	**5.8**	**1.91, 17.75**
9 (28+)	27	80.0	21.6	18.5	5	**4.3**	**1.76,10.36**
Total	383	26.5	101.4	5.4	22	**4.6**	**3.04,6.95**

Note: Bolded E/O index and 95%CIs denote significance.

3.2.3.2 Results by gender.

Finally, the same calibration analyses were completed by gender. Logistic regression analyses and an E/O index could not be computed for females due to the low number of females who recidivated, as noted previously.

Logistic regressions for males were conducted for 5-year and 10-year fixed follow ups. The Hosmer - Lemeshow goodness of fit tests were both nonsignificant and suggested that the logistic distributions provided a reasonable approximation of violent recidivism rates to warrant modelling. Results of logistic regression generated the following terms for 5-year (B_0 = -2.717, B_1 = .036, p = .002) and 10-year (B_0 = -2.328, B_1 = .037, p < .001) violent recidivism for males (Figure 3.3). Figure 3.4 juxtaposes the previous overall results for all possible estimated VRAG-R scores for 5-year and 10-year violent recidivism with the same results for males. Visual inspection indicates little change in values when the values for females are removed. Figure 3.5, in turn, provides a visual inspection of the correspondence between estimated and observed rates for 5-year and 10- year violent recidivism. Again, there are some fluctuations due to small *n*s of recidivists in some bins.

Figure 3.3

Logistic Regression Estimated 5 and 10-year Violent Recidivism for all Possible VRAG-R Scores (Males Only)

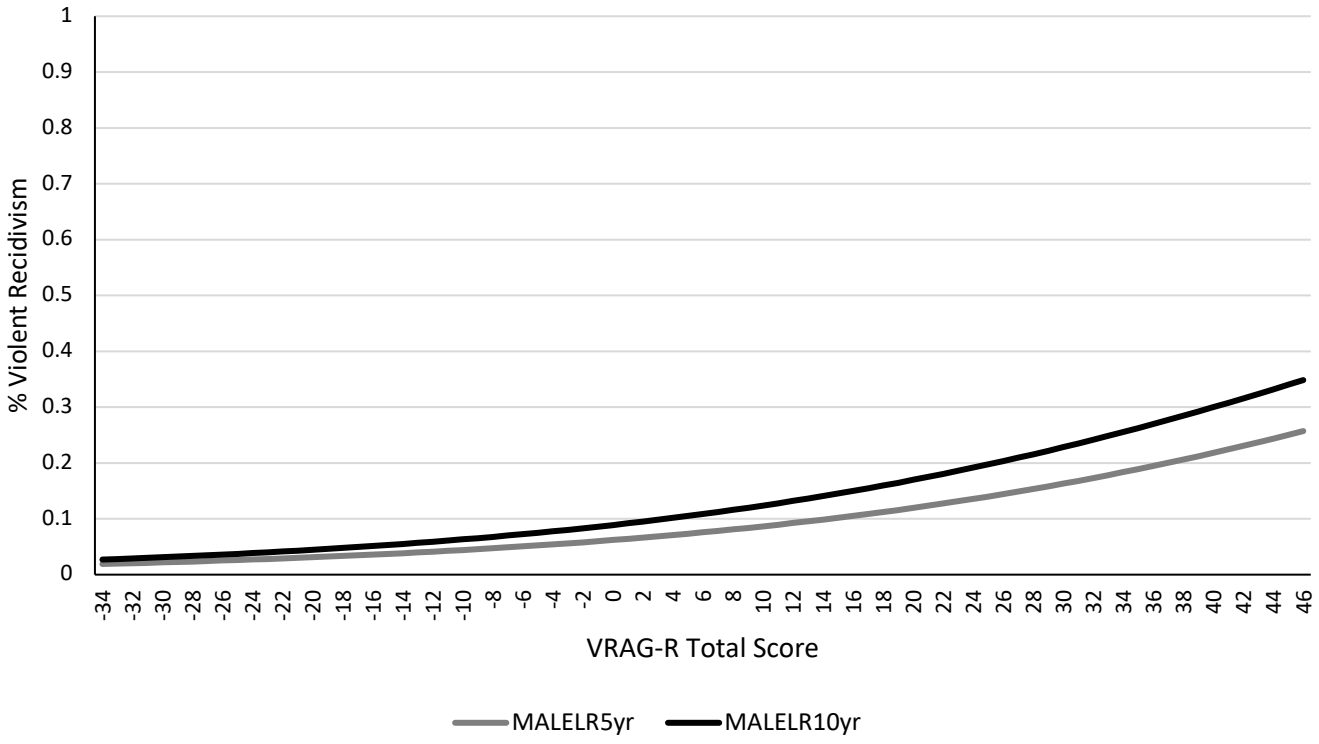


Figure 3.4

Logistic Regression Estimated 5 and 10-Year Violent Recidivism for all Possible VRAG-R Scores (Males Only) Juxtaposed with Predicted 5 and 10-year Rates for the Overall Population.

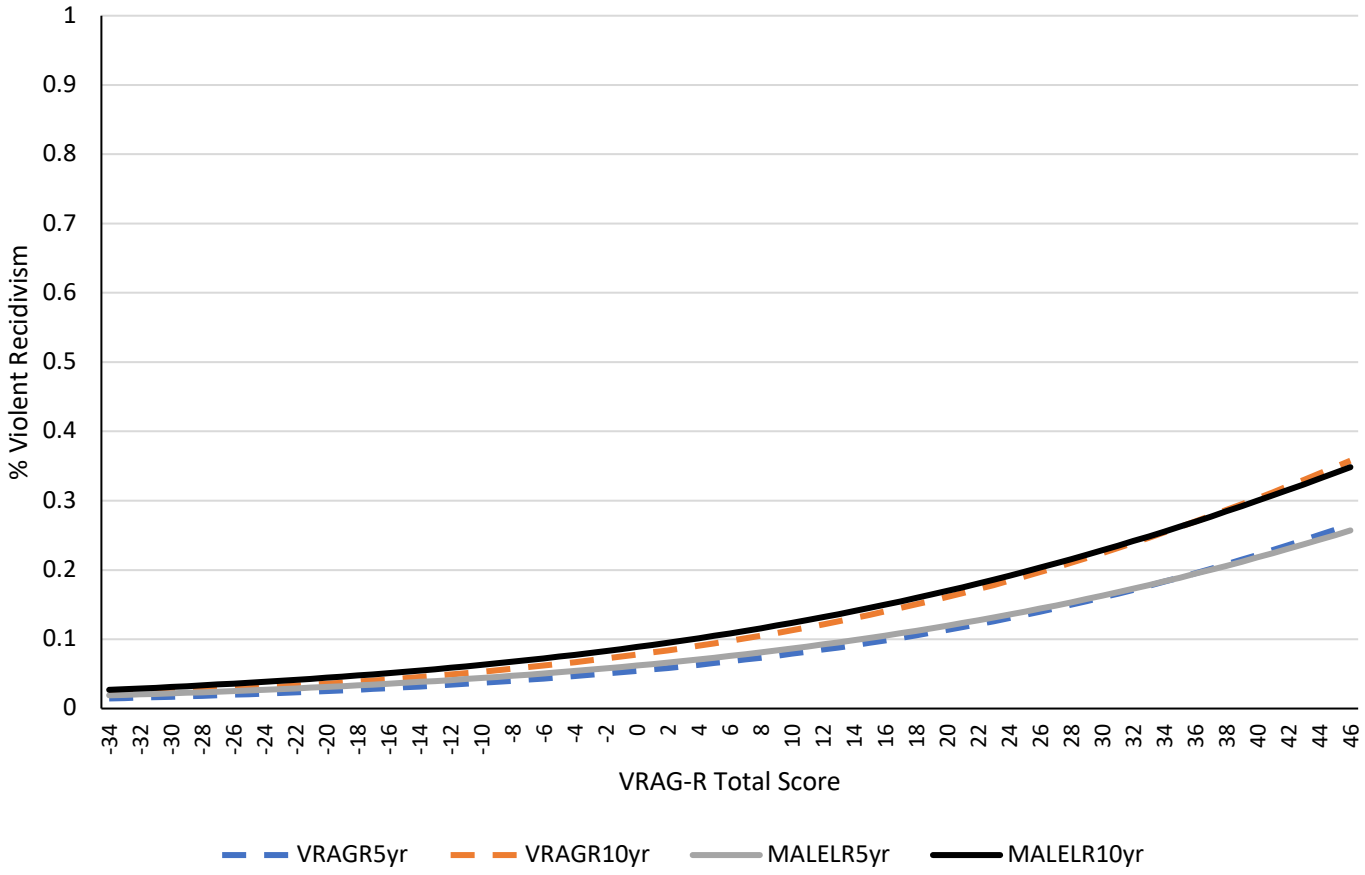


Figure 3.5

VRAG-R Calibration: Actual Rates of Violent Recidivism for the Nine-Bin Structure and Estimated Rates of Violent Recidivism Associated with Individual Scores over Fixed 5 and 10-Year Follow-ups (Males Only)

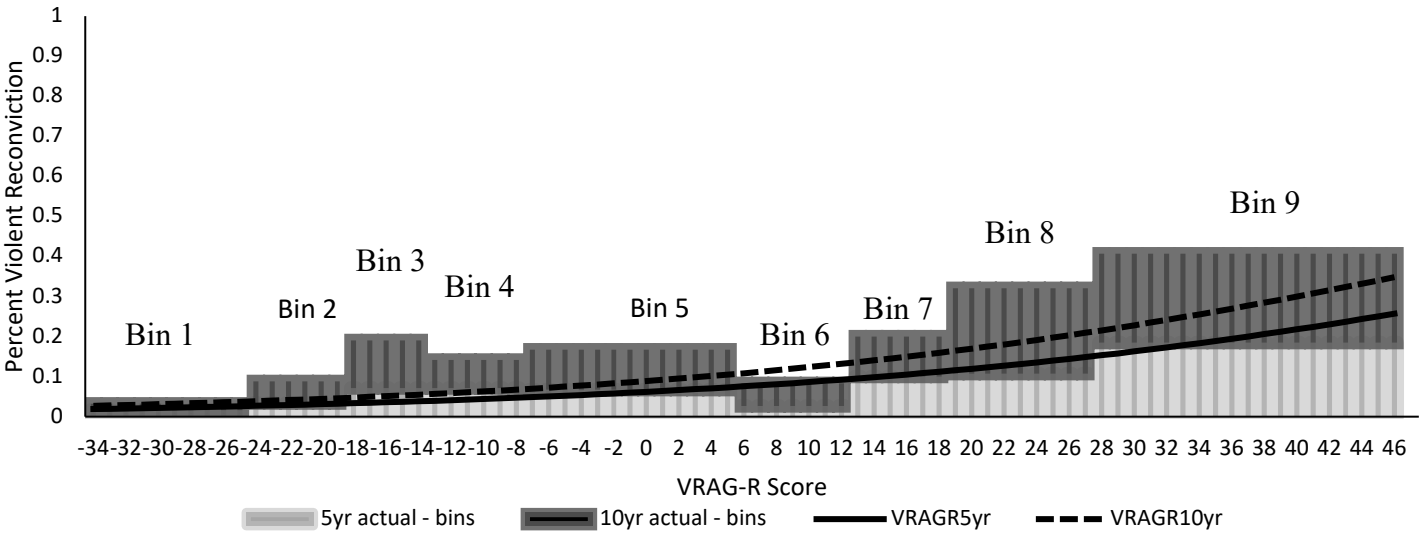


Table 3.7 provides the E/O index computed based on the number of recidivists from the observed 5-year violent recidivism rates for males only from the Alberta population compared to the 2013 VRAG-R construction sample. As noted previously, data for the 10-year outcomes was not available at the time of investigation. The results indicated that VRAG-R scores substantially overestimated risk for males in the Alberta NCRMD population. All E/O indexes were greater than 1.0 (i.e., where 1.0 indicates perfect calibration) with some values reaching as high as 13.7 (Bin 6). However, the very small cell *n*s for some of the observed recidivist frequencies may have contributed to this overestimation. In addition, there was still a significant overestimation of risk among males, with the VRAG-R total score overestimating risk by 4.2 times. Similar to the combined values above, in general there was poor agreement between the expected and observed recidivism rates, however, to a slightly lesser degree than when males and females were combined. The confidence intervals for the first five bins were non-significant, indicating that there was not a significant difference between the expected and observed rate of recidivism. For bins 6 and above including the overall observed rates of violent recidivism, there was a significant difference between expected and observed rates of future violence.

Table 3.7

E/O Index: Five-Year Rates of Violent Recidivism for Normative Sample (Rice, Harris, Lang, 2013) Compared with Alberta Population (Males Only)

Risk bin and score range	5-year violent recidivism						
		Expected rates: VRAG-R norms		Observed rates: Current sample			
	n	%	n	%	n	E/O Index	95% CI
1 (≤ -25)	62	8.0	5.0	1.6	1	5.0	.7, 35.5
2 (-24 to -19)	58	9.0	5.2	3.4	2	2.6	.65, 10.56
3 (-18 to -14)	42	18.0	7.6	7.1	3	2.5	.83, 7.65
4 (-13 to -8)	42	19.0	5.9	7.1	3	2.0	.66, 6.12
5 (-7 to 5)	30	25.0	7.5	6.7	2	3.8	.95, 15.43
6 (6 to 12)	37	37.0	13.7	2.7	1	**13.7**	**1.92, 97.27**
7 (13 to 18)	20	45.0	9.0	10.0	2	**4.5**	**1.13, 18.27**
8 (19 to 27)	28	58.0	16.2	10.7	3	**5.4**	**1.78, 16.52**
9 (28+)	27	80.0	21.6	18.5	5	**4.3**	**1.76, 10.36**
Total	346	26.5	91.7	6.4	22	**4.2**	**2.77, 6.34**

Note: Bolded E/O index and 95%CIs denote significance

3.2.4 Discrimination and Calibration Properties of the VRAG-R for Violent Recidivism by Diagnostic Category

As a point of interest, the same analyses were conducted across the following diagnostic categories: any psychotic disorder, any mood disorder, substance use disorder (i.e., present or ever) and antisocial personality disorder (including traits). These represented the most stable categories for examination. Such analyses were conducted to determine if diagnostic categories were predictive, a meaningful question given the gendered differences regarding diagnosis found within the population.

3.2.4.1 Discrimination.

For those individuals who received a diagnosis on the psychotic disorder spectrum at their first NCRMD review board meeting, 5-year fixed follow-up (n = 334) included 27 individuals (8.0 %) who received convictions for any new offense (i.e., general recidivism), which included 18 individuals (5.4 %) who received convictions for new violent offenses. At 10-year fixed follow up *(*n = 330) 36 individuals (11%) received new convictions for any new offenses, which included 22 individuals (6.7%) convicted for new violent offenses. For the overall follow-up for individuals diagnosed on the psychotic spectrum (n = 387), 51 individuals (13.2%) were reconvicted for any new offenses, which included 32 individuals (8.3%) who reoffended violently.

For those individuals who received a diagnosis of a mood disorder at their first NCRMD review board meeting, 5-year fixed follow up (n = 130) included 6 individuals (4.6 %) who received convictions for any new offense, which included 2 individuals (1.5%) who received convictions for new violent offenses. At 10-year fixed follow up *(*n = 129), 11 individuals (8.5%) received new convictions for any new offenses, which included 4 individuals (3.1%) convicted

for new violent offenses. For the overall follow-up for individuals who received a mood disorder diagnosis (*n* = 153), 19 individuals (12.4%) were reconvicted for any new offenses, which included 9 individuals (5.9%) who reoffended violently.

For those individuals whose file indicated a substance use disorder (either present or ever), 5-year fixed follow up (*n* = 216) included 23 individuals (10.6%) who received convictions for any new offense, which included 12 individuals (5.6%) who received convictions for new violent offenses. At 10-year fixed follow up *(n* = 215), 33 individuals (15.3%) received new convictions for any new offenses, which included 17 individuals (8.0%) convicted for new violent offenses. For the overall follow-up (*n* = 251), 42 individuals (16.7%) were reconvicted for any new offenses, which included 22 individuals (8.8%) who reoffended violently.

For those individuals whose file indicated antisocial personality disorder and/or traits, 5-year fixed follow-up (*n* = 83) included 15 individuals (18.1%) who received convictions for any new offense, which included 13 individuals (15.7%) with convictions for new violent offenses. At 10-year fixed follow-up *(n* = 83), 23 individuals (27.7%) received new convictions for any new offenses, which included 18 individuals (21.7%) convicted for new violent offenses. For the overall follow-up for individuals whose file indicated antisocial personality disorder and/or traits (*n* = 97), 26 individuals (26.8%) were reconvicted for any new offenses, which included 19 individuals (19.6%) who reoffended violently.

Table 3.8 provides the results of the Receiver Operator Characteristic (ROC) analyses for the four diagnostic categories (e.g., psychotic disorders, mood disorders, substance use disorder including present or ever, and antisocial personality disorder plus traits) for 5-year, 10-year, and overall follow-up by both VRAG-R total score and by bin number. Following in a similar manner to the analyses above, several prominent themes arose from the analyses. Again, the

guidelines laid out by Rice and Harris (2005) for interpreting AUC magnitudes were utilized. First, regarding the results for psychotic disorders, both total score and bin level significantly predicted general and violent recidivism. Consistent with the previous analyses, better predictive accuracy was observed in both total score and bin number for violent reoffending than general reoffending; however, overall total score demonstrated better predictive accuracy over bin numbers. Furthermore, for psychotic disorders, AUC magnitudes were lower than previous analyses with most being classified as small or medium. None of the AUC values for psychotic disorders reached a magnitude large enough to be classified as large. Second, regarding the ROC analyses for mood disorders, the AUC values were unstable at fixed follow-ups due to so few violent recidivists. As such, AUC magnitudes tended to be more stable across diagnostic categories using unfixed follow-ups. Third, VRAG-R total scores and bin groupings for each of antisocial personality disorder/traits and substance use disorder (present or ever) did not significantly predict any recidivism outcomes and AUCs were generally small in magnitude; an interesting finding, given both diagnoses are criminogenically relevant.

Table 3.8

Discrimination Properties of the VRAG-R for Violent and General Recidivism for 5-year, 10-year, and Overall Follow-up by Diagnostic Category

Recidivism Criterion	Timeframe	Psychotic Disorders		Mood Disorders		Substance Use		ASPD	
		AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI
Total Score Violent	5 Year	.666*	[.546, .787]	.902	[.761, 1.00]	.646	[.468, .824]	.632	[.450, .814]
	10 Year	.658*	[.548, .768]	.771	[.576, .966]	.610	[.466, .754]	.587	[.423, .751]
	Overall	.658**	[565, .751]	.710*	[.539, .881]	.604	[.482, .725]	.566	[.411, .711]
Total Score General	5 Year	.636*	[.527, .745]	.886***	[.792, .979]	.608	[.473, .743]	.662	[.498, .825]
	10 Year	.633**	[.536, .731]	.697*	[.515, .880]	.586	[.476, .697]	.595	[.450, .741]
	Overall	.609*	[.528, .690]	.642*	[.504, 780]	.570	[.475, .665]	.550	[.449, .685]
Bins Violent	5 Year	.645*	[.520, .771]	.873	[.719, 1.00]	.625	[.449, .801]	.599	[.419, .788]
	10 Year	.637*	[.523, .752]	.766	[.605, .927]	.597	[.455, .738]	.555	[.394, .716]
	Overall	.647**	[.551, .743]	.692	[.514, .870]	.597	[.477, .717]	.538	[.387, .689]
Bins General	5 Year	.620*	[.512, .729]	.870	[.767, .973]	.593	[.457, .728]	.635	[.471, .799]
	10 Year	.621*	[.526, .717]	.695*	[.518, .871]	.572	[.462, .682]	.560	[.416, .705]
	Overall	.599*	[.518, .680]	.624	[.481, 767]	.558	[.463, .653]	.522	[.390, .654]

Note: *$p < .05$, **$p < .01$, *** $p < .001$; AUC = area under the curve; CI = confidence interval

3.2.4.2 Calibration

Logistic regression was conducted to estimate the rates of recidivism associated with specific VRAG-R scores over a given follow-up (e.g., 5-year or 10-year) for the four diagnostic categories identified above (i.e., psychotic disorders, mood disorders, antisocial personality disorder and/or traits, and substance use disorder). The Hosmer-Lemeshow goodness of fit tests were all nonsignificant for all four diagnostic categories suggesting that the logistic distributions provided a reasonable approximation of violent recidivism rates to warrant modelling. Results of the logistic regression generated for each of the diagnostic categories for 5-year and 10-year recidivism are displayed in Table 3.9. Using a log linking function, these values were employed to estimate rates of recidivism associated with all possible VRAG-R total scores. Figures 3.6 to 3.9 illustrate all possible estimated VRAG-R scores for 5-year and 10-year violent recidivism as estimated through logistic regression for each of the diagnostic categories. Notably, the results for both mood 5-year and ASPD 10-year demonstrated inflated curves. Such results are potentially due to the small *n*s. Figure 3.10 provides a visual for the logistic regression for estimated 5-year and 10-year violent recidivism for all possible VRAG-R scores by diagnostic category. The logistic regression for 5-year mood disorders is excluded due to only having 2 recidivists within the category. Figure 3.11 provides a visual of the logistic regressions estimated for 5-year and 10-year violent recidivism for the observed VRAG-R scores by diagnostic category.

Table 3.9

Binary Logistic Regression of VRAG-R Scores for Violent Recidivism for 5-year, and 10-year Follow-up by Diagnostic Category

	Regression model by diagnostic group					
Regression model outcome	*B*	*SE*	Wald	*p*	Exp(B)	95% CL [LL, UL]
			Psychotic Disorders			
5-year						
VRAG-R score	.032	.013	6.026	.014	1.032	[1.006,1.059]
Constant	-2.798	.247				
10-year						
VRAG-R score	.030	.012	6.471	.011	1.030	[1.007, 1.054]
Constant	-2.566	.224				
			Mood Disorders			
5-year						
VRAG-R score	.089	.046	3.755	.053	.1.093	[.999, 1.1196]
Constant	-4.528	1.126				
10-year						
VRAG-R score	.050	.026	3.658	.056	1.051	[.999, 1.106]
Constant	-3.287	.537				
			Substance Use			
5-year						
VRAG-R score	.031	.017	3.498	.061	1.032	[.998, 1.067]
Constant	-2.972	.335				
10-year						
VRAG-R score	.023	.014	2.699	.100	1.023	[.996, 1.052]
Constant	-2.523	.269				
			ASPD			
5-year						
VRAG-R score	.028	.021	1.837	.175	1.028	[.988, 1.071]
Constant	-2.103	.471				
10-year						
VRAG-R score	.020	.017	1.288	.256	1.020	[.986, 1.055]
Constant	-1.558	.378				

Note. CL=confidence interval; LL= lower limit; UL= upper limit

Figure 3.6

Logistic Regression Estimated 5 and 10-year Violent Recidivism for all Possible VRAG-R Scores (Psychotic Disorders)

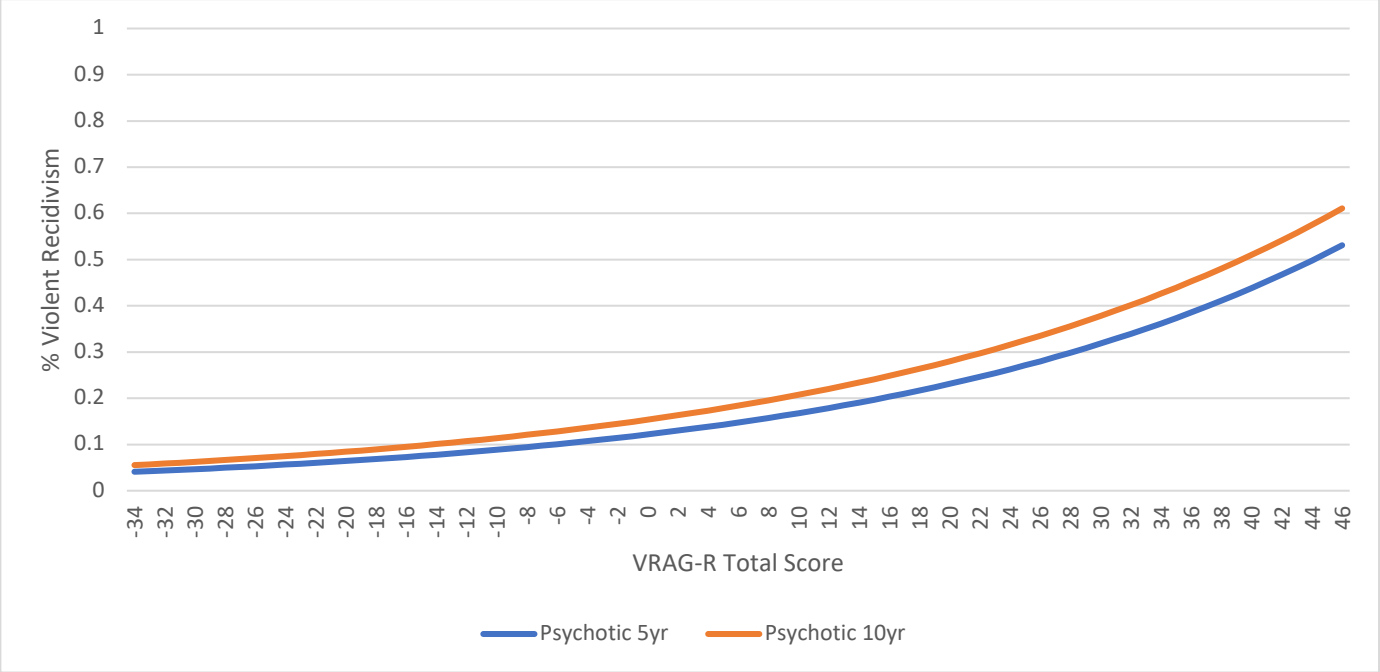


Figure 3.7

Logistic Regression Estimated 5 and 10-year Violent Recidivism for all Possible VRAG-R Scores (Mood Disorders)

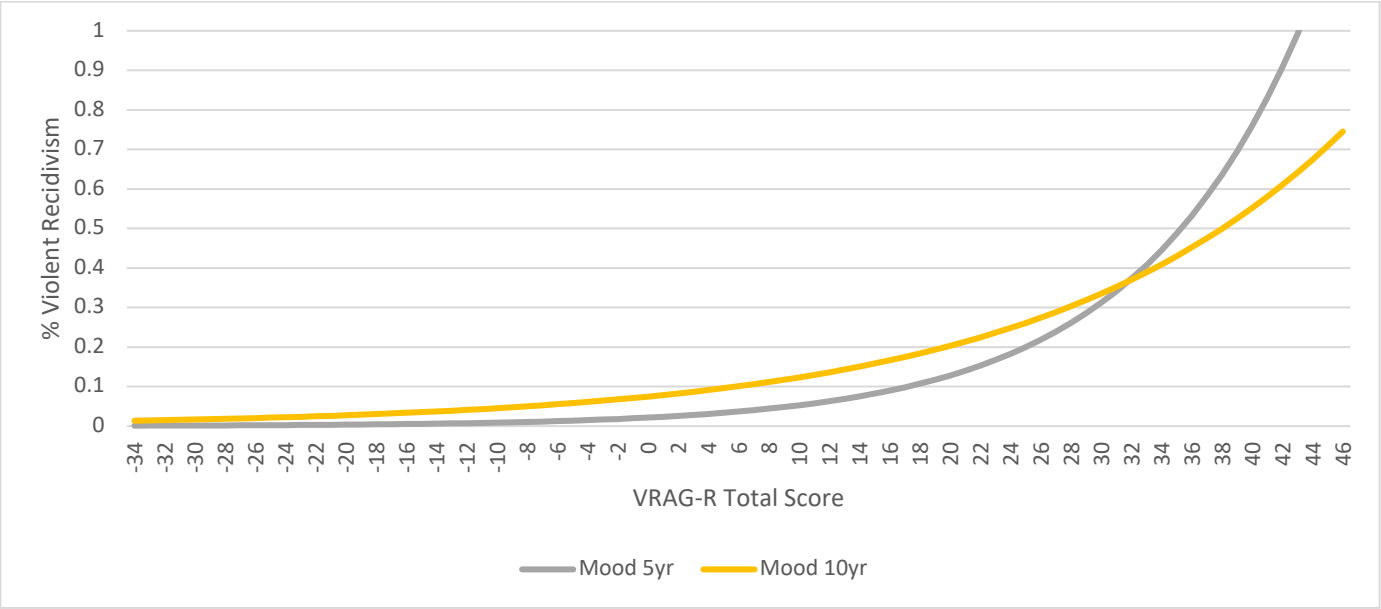


Figure 3.8

Logistic Regression Estimated 5 and 10-year Violent Recidivism for all Possible VRAG-R Scores (ASPD)

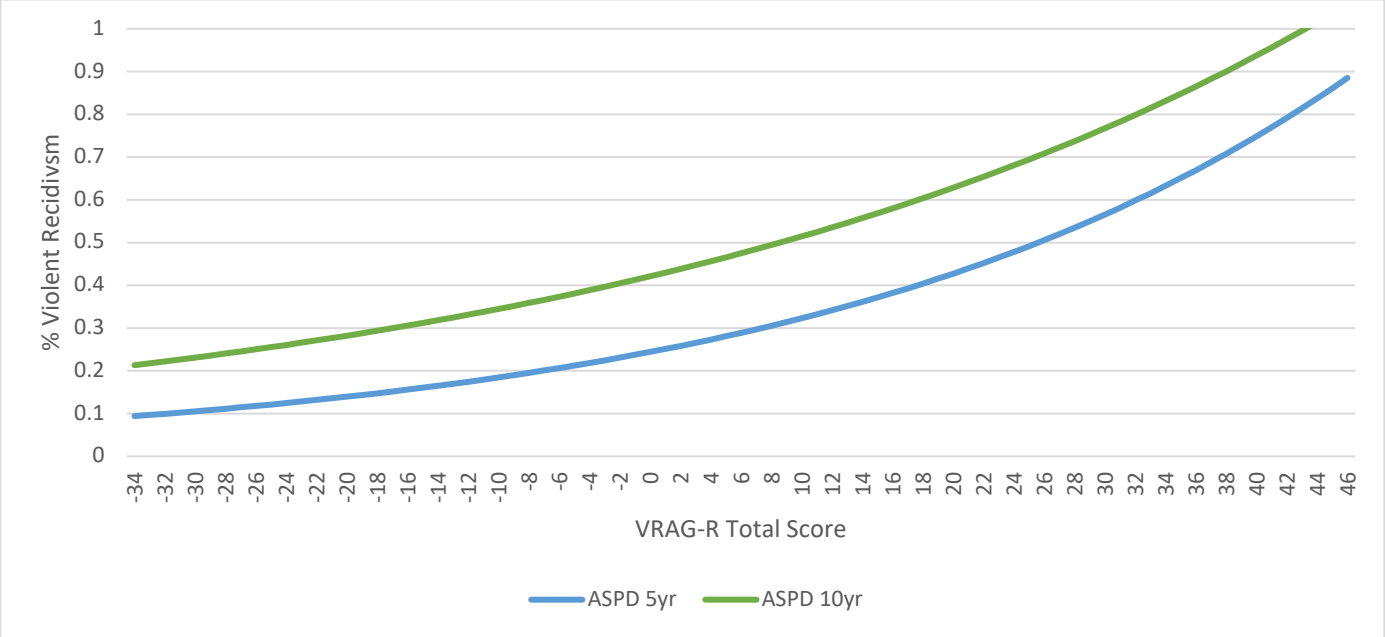


Figure 3.9

Logistic Regression Estimated 5 and 10-year Violent Recidivism for all Possible VRAG-R Scores (SUD)

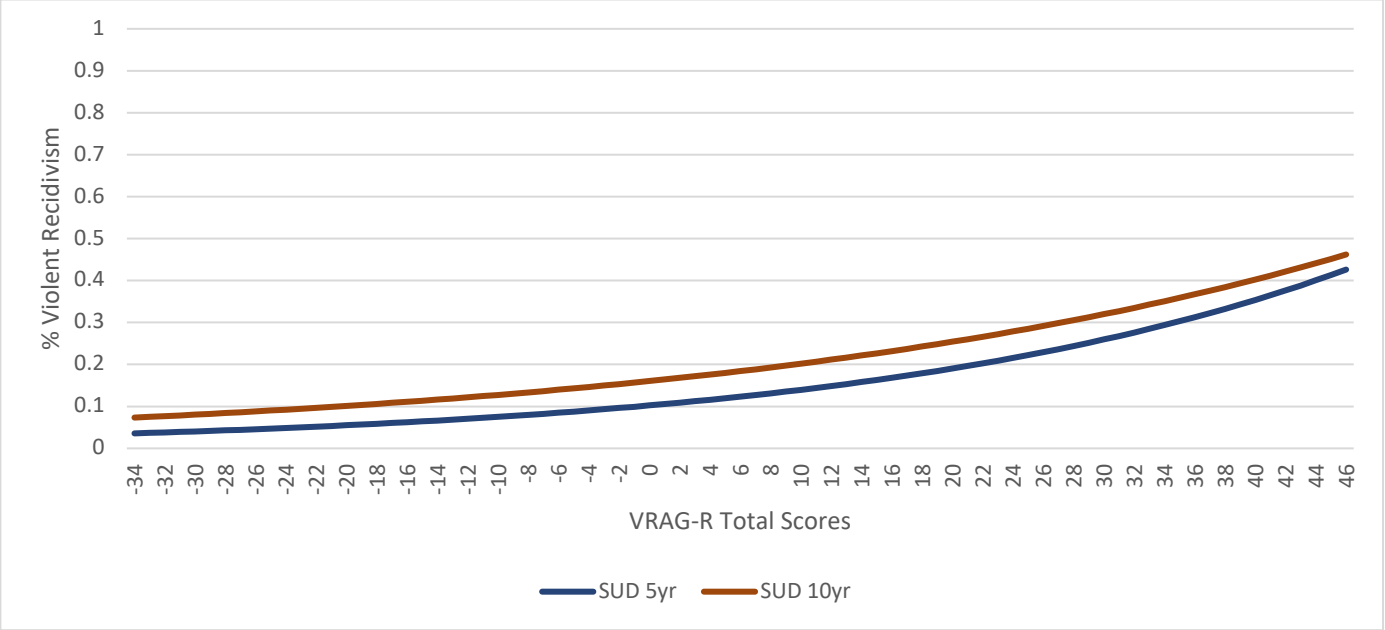


Figure 3.10

Logistic Regression Estimated 5 and 10-year Violent Recidivism for all Possible VRAG-R Scores by Diagnostic Category (excluding mood 5-year)

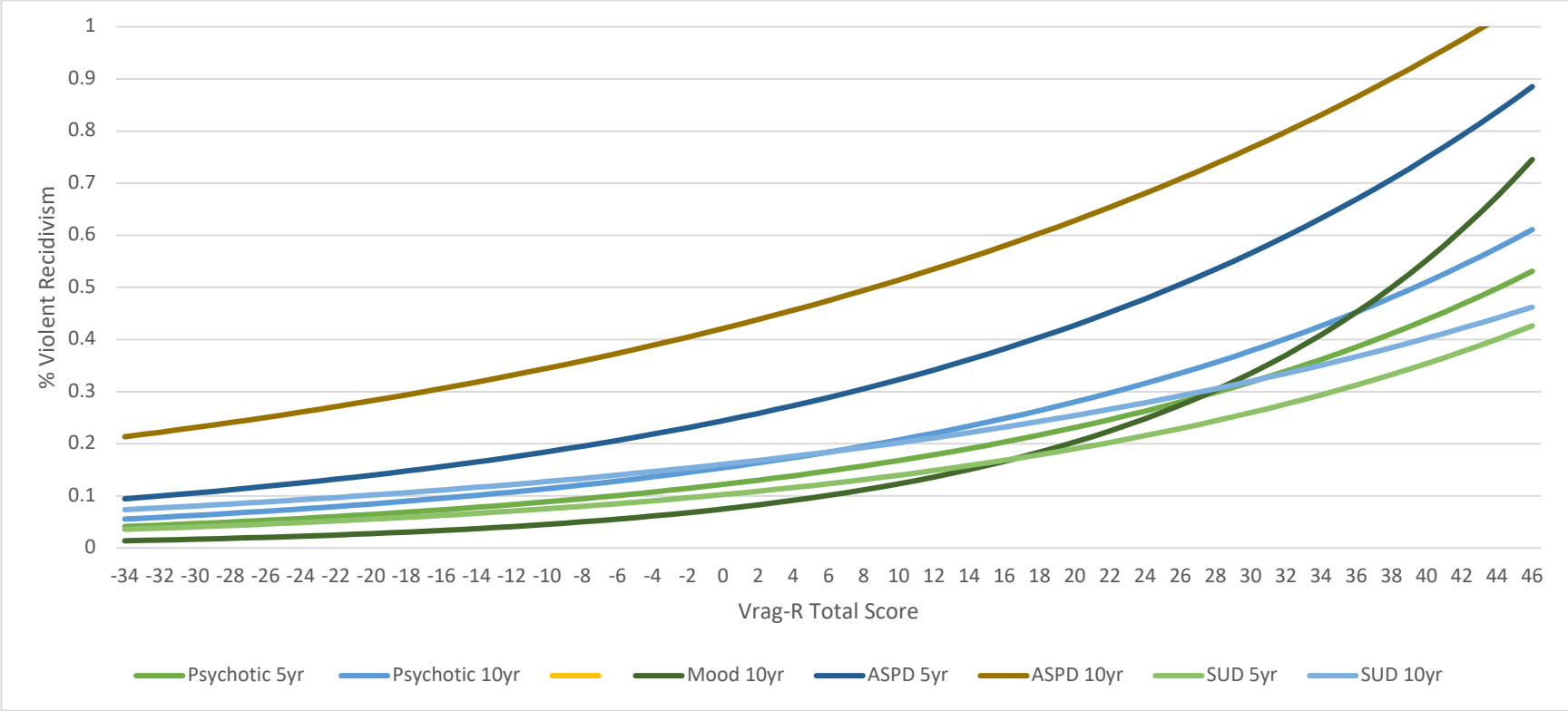
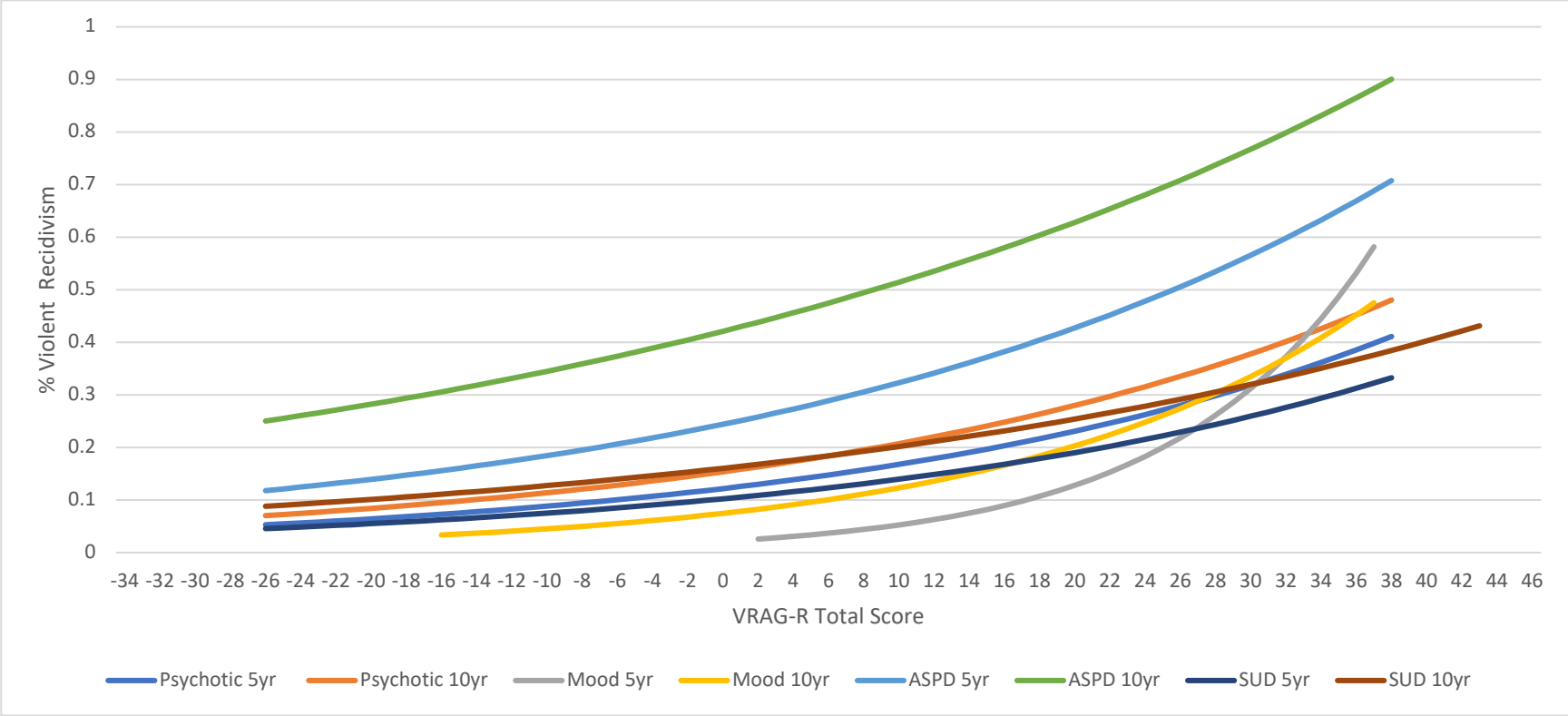


Figure 3.11

Logistic Regression Estimated 5 and 10-year Violent Recidivism for Observed VRAG-R Scores by Diagnostic Category (all)



CHAPTER FOUR: DISCUSSION

The current study sought to examine the risk profiles and recidivism outcomes of individuals who had been found Not Criminally Responsible on Account of Mental Disorder (NCRMD) in the province of Alberta. The study had two goals: first to investigate the population for potential gender differences, and second to examine the predictive accuracy (i.e., the discrimination and calibration properties) of the VRAG-R. Given the previous research on individuals found NCRMD in Canada, it was expected that some unique sociodemographic, clinical, and criminological factors would be found between the genders. Regarding the VRAG-R it was expected that: (a) females would have lower VRAG-R total scores and bin number frequency distributions, and (b) that the VRAG-R would have strong discrimination and calibration properties for both the population and separated by gender (i.e., male/female). Past research supported such findings, however, had little to say specifically regarding the NCRMD population in Alberta, the potential gender differences therein, and the use of the VRAG-R within this population. Thus, the results of this study worked towards addressing these gaps in the forensic mental health and violence risk assessment literatures. The results are discussed below, in turn, in addition to some considerations regarding the clinical and correctional implications of this study, strengths and limitations, and future directions.

4.1 Gender Differences within the Alberta NCRMD Population

4.1.1 Sociodemographic Characteristics

Results for the sociodemographic characteristics for both the overall population and by gender were consistent with previous research (Crocker, Nicholls, Seto, Charette, et al., 2015; Nicholls et al., 2015; Statistics Canada, 2014). For example, overall, the Alberta NCRMD population was largely white (71%), single (82%), and the majority had not completed high

school (65%). Such a description was in line with both the findings of the National Trajectory Project (NTP) and the findings of a 2014 Statistics Canada report on persons who had been found NCRMD. Regarding gender differences, there were fewer sociodemographic differences than similarities. For example, no difference in age at index offense, marital status, or ethnicity was found. However, there was a significant difference in the number of females and males who completed high school, with females more likely to have completed high school or higher levels of education versus their male counterparts (e.g., 41.6% versus 30% respectively). These findings can be contrasted with those of the NTP wherein there were more robust sociodemographic differences between the genders (Nicholls et al., 2015). For example, females in the NTP sample tended to be older at the time of index offense, to have completed high school, and were more likely to be married than their male counterparts.

In both this study and the NTP, there does appear to be some evidence that females may have slightly higher psychosocial functioning than their male counterparts, such as having completed high school, yet neither this study nor the NTP collected the necessary information in order to formally test this assumption. Overall, the population of persons found NCRMD in Alberta, appears to have less sociodemographic differentiation via gender than in other provinces such as British Columbia, Ontario, and Quebec. That said, most other provinces have yet to conduct similar research regarding gender, and warrants further investigation.

4.1.2 Mental Health Characteristics

As noted above, the sample had high base rates of serious mental health disorders, consistent with the spirit of the NCRMD legislation. In terms of diagnosis and clinical profile, results were consistent in that differences existed between the genders. For example, females were more likely to be diagnosed with a mood disorder or borderline personality disorder while

males were more likely to be diagnosed with antisocial personality disorder or substance use disorder. These results have been found in other research on persons found NCRMD and it is well established in the *Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition* (DSM-5) that these specific disorders often have gendered prevalence rates (Nicholls et al. 2015; American Psychiatric Association, 2013). An unexpected result was the lack of gender difference between the occurrence of diagnoses on the psychotic spectrum. In research done by the NTP gender differences have been found with a significantly larger number of males being diagnosed with a psychotic spectrum disorder (Nicholls et al., 2015). The reason for the lack of occurrence is unknown yet represents an interesting aspect of the Alberta NCRMD population compared to other provinces in Canada.

4.1.3 Criminological Characteristics

Criminological variables revealed some differences between the genders. For instance, while there was no difference between number of prior NCRMD findings and index offense category, there was a significant difference in the mean number of prior sentences with males having a higher number of prior sentences than their female counterparts. Females having a lower number of previous criminal convictions or sentences is consistent across the literature for general offenders, mentally disordered offenders, and forensic patients (Nicholls, Cruise, Greig, & Hinz, 2015). Considering criminal history is often a strong predictor of future reoffending, it has been suggested that the lower rate of recidivism for females could potentially be connected to their frequently lower rates of criminal history.

Regarding conditional discharge, the genders were treated similarly as there was no difference between the mean length of time to receive a conditional discharge from the review board (RB) nor were there any difference in the number of conditional discharge revocations.

Concerning absolute discharge (AD), again, there was no difference between the genders in both receiving an AD and also in the length of time it took to receive an AD.

In terms of recidivism, the results were both consistent with the literature (i.e., lower rate of recidivism for females) but surprising in absolute value. Only three females in the scope of time covered by the recidivism data held within the database (i.e., 1941 – 2015) had reoffended after being granted an AD; this is in contrast to 92 males in the same time period. In fact, the base rate of females who recidivated was so low that the predictive accuracy of the VRAG-R could not be examined with this subgroup (i.e., due to instability of findings from small N). While this finding is in line with other areas of criminological research indicating that females are often found to reoffend at a lower rate than males, the extent to which females in this population did not reoffend is surprising (Nicholls, Cruise, Greig, & Hinz, 2015). Further research in other provincial jurisdictions, building on the work of the NTP and current study, will establish if this is indeed the norm within the NCRMD population. A further possibility is that it may be a combination of factors including the lower recidivism rates of females compared to males, that RBs are successful at managing persons found NCRMD, and possible emerging evidence suggesting that females might have a slightly higher level of psychosocial functioning than their male counterparts—all of which mitigate risk. Taken together this might explain the intensely lower rate of recidivism for females within this population. However, it is also possible that other criminogenically relevant factors that are outside the current scope of examination are involved in the much lower rate of recidivism.

4.1.4 Overall Gender Profiles

Taken together the results of this section of the study created a profile of the Alberta NCRMD population as a whole that was a largely white, Euro-Canadian, single, younger middle

age sample of adults, with incomplete education, and diagnoses of serious mental illness that were in line with the spirit of the NCRMD legislation. The resulting gender profiles within the Alberta NCRMD population indicated there were more similarities than differences between the genders, yet the differences that did exist were of significant importance. For example, female individuals who have been found NCRMD in Alberta were more likely than their male counterparts to have completed high school, to have fewer previous sentences, and to have had a greater likelihood of being diagnosed with a mood disorder or borderline personality disorder. In contrast, males, were more likely to have had a greater mean number of previous sentences, to have been diagnosed with antisocial personality disorder (ASPD) or substance use, and to have had a higher rate of recidivism. For certain criminological variables (e.g., time under the RB, time to CD or AD, CD revocations), there were no significantly notable differences other than number of previous sentences and rate of recidivism with males having higher values in both. In fact, the absolute value in the difference between the rates of recidivism between the genders, with females having a drastically lower rate of recidivism, was one of the most outstanding findings of this section of the study. Given all this information it would have been expected that due to the sociodemographic characteristics, fewer criminological factors indicating higher risk and shockingly low rates of recidivism demonstrated by females within the Alberta NCRMD population that there would have been more differentiation in how they were treated by the review board. However, this was not demonstrated by the results.

Overall, while the differences between the genders were not as robust (i.e., fewer differences between the genders) as in some studies (i.e., the NTP), the differences that were found were of significant importance and provide further evidence for the unique

sociodemographic, clinical, and criminological profiles of the genders in this population that have also been found in previous research.

4.2 Violence Risk Appraisal Guide Revised

4.2.1 Descriptive Statistics and Frequencies of the Violence Risk Appraisal Guide Revised (VRAG-R) Scores

The population of persons found Not Criminally Responsible on Account of Mental Disorder (NCRMD) in Alberta was much lower risk than the construction or validation sample used to develop the VRAG (Harris et al., 1993) and VRAG-R (Rice et al., 2013*).* The current sample scored about a half standard deviation lower overall than the normative group, and the differences were even more marked for females. As expected, females, overall, did have a much lower total score and bin number frequency distribution than their male counterparts at approximately two thirds of a standard deviation difference. Indeed, nearly half of all the females' total scores classified their risk within the first risk bin of the VRAG-R, while an approximate similar proportion of males was distributed among the first three risk bins.

The findings of lower overall risk scores for females within this population is congruent with the above findings regarding the sociodemographic, clinical, and criminological profiles of the genders within this population. The findings indicated few criminological variables for females that would indicate a higher risk score (e.g., criminal history) in addition to more sociodemographic characteristics that may indicate stronger prosocial functioning. In all, it was apparent that females commit crime and violence at lower rates and have lower risk scores than their male counter parts, a result that was clearly demonstrated within this study

4.2.2 Discrimination

One of the key aims of the present dissertation was to examine the discrimination properties of the VRAG-R for violent and general recidivism in an NCRMD sample; that is, to what extent do VRAG-R scores differentiate discharged persons who eventually reoffend violently or generally from the vast majority who do not. Overall the discrimination results were in line with what was predicted, and consistent with past research wherein strong discrimination results have been found for the VRAG-R (Rice et al., 2013; Olver & Sewall, 2018; Gregório, Hertz, Rettenberger, & Eher, 2019; Olver & Hogan, 2019). Strong discrimination, as evidenced by broadly large AUC magnitudes, was observed for both the overall sample and male subgroup. Both VRAG-R total score and bin classification appeared to be equally predictive, with slightly higher AUC magnitudes for violent vs. general recidivism, consistent with past research. Although this pattern was observed for males, the very small number of female recidivists precluded conducting discrimination analyses for this subgroup. Of note, the AUCs were slightly higher for the sample overall than in the male subgroup. This likely reflects the fact that the AUC is a rank ordered statistic, such that higher AUC magnitudes would reflect a greater concentration of recidivists at the top end of scores. Given that few or no females reoffended and also had lower scores, this would stretch out the bottom half of the distribution of risk scores, populated largely by non-recidivists, thereby improving discrimination.

Strong discrimination as evidenced by large AUC magnitudes were not found when examining the discrimination properties via mental health diagnoses. Only the results for psychotic disorders for total score and bin number significantly predicted general and violent recidivism and even then the AUC magnitudes were lower (i.e., small to medium) than the above mentioned results for the overall and by gender analyses. The VRAG-R did not significantly predict for mood disorders, substance use disorders or antisocial personality disorder including

ASPD traits. Interestingly, antisocial personality disorder had almost double the rate of recidivism in comparison to the other diagnostic categories. It is important to note that ASPD is inherently criminogenic - as such the higher rates of recidivism would be expected as ASPD is going to drive up recidivism rates.

4.2.3 Calibration

The calibration analyses, in turn, sought to address the crucial question: "what recidivism rates are associated with VRAG-R scores in an Alberta NCRMD sample?" This was examined through frequency analysis of VRAG-R bin recidivism rates, and more specifically, through comparison of observed frequencies of 5-year violent recidivism from the current sample, with those expected from the Rice et al. (2013) normative sample, via the E/O index. Results for the calibration properties of the VRAG-R within this population were unexpected. Previous research such as the study done by Olver and Sewall (2018) had found strong calibration properties in tandem with strong discrimination; however, the results of the calibration analysis indicated that the VRAG-R scores were substantially overestimating risk for the Alberta NCRMD population at each risk bin, and in general there was poor agreement between the expected and observed rates of recidivism. The E/O index was significant in 7 out of 9 bin comparisons, with the overall rate of 5-year violent recidivism being nearly 5 times higher for the normative sample. Violent recidivism was overpredicted in each bin by 2 to 15 times (or by 200 to 1500%)!

Such substantial over calibration by the VRAG-R in the Alberta NCRMD population was surprising given that most of the, albeit limited number of, other studies examining the calibration of the VRAG-R have not found this. For example as noted above, in Olver and Sewall (2018) strong calibration properties were found in the use of the VRAG-R with persons who had committed sexual offenses. This population was higher risk than that of the Alberta

NCRMD population; as such, potentially the VRAG-R is better calibrated with a higher risk population. Olver and Sewall (2018) examined a federal correctional sample of men referred to a high intensity sexual violence treatment program—most of the men had a previous history of sexual violence or other personal or psychological characteristics that would prompt referral to such a program. The sample was even actuarially higher risk than the Rice et al. (2013) sample. But of note, although the Olver and Sewall (2018) sample had closer calibration to the VRAG-R norms, the E/O index values indicated non-significant overprediction of 5-year violent recidivism, even though the Olver and Sewall sample was still higher risk overall. Such findings suggest there may be something distinct about the VRAG-R Penetanguishine forensic mental health sample that sets them apart, from even a risky federal sex offender sample (e.g., unmeasured dynamic risk factors, criminogenically relevant mental health symptoms, etc.).

It may also be worth examining the calibration properties of the VRAG-R in another low risk population to ascertain the stability of recidivism estimates. The results for the gendered analysis also mirrored this surprising result with the calibration results for the males again substantially overestimating risk and in general having poor agreement between the expected and observed rates, though to a lesser degree.

4.3 Implications for Research and Practice

Taken together the results of this study have several important clinical and correctional implications regarding policy, practice, treatment, and assessment within the Alberta NCRMD population. First, the study supports previous research indicating that unique sociodemographic, clinical and criminological profiles exist for the genders. Evidence supports the idea that females found NCRMD in Alberta appear to have a higher level of psychosocial functioning as evidenced by their lower number of previous crimes, higher likelihood of completing school and

lower rates of recidivism. Both genders within the Alberta NCRMD population are more likely to have certain mental health diagnoses which can have implications for treatment and long term prognosis. Results such as these support the importance of at minimum attending to gender as a specific responsivity issue when dealing with this population; particularly with respect to treatment and risk assessment.

Second, results also indicated that while certain differences existed between the genders within the Alberta NCRMD population, the genders were not being treated differently via review board dispositions regarding decisions such as conditional release and absolute discharge. Some consideration may need to be given to the fact that females within this population have a remarkably lower rate of recidivism and much lower risk scores overall, yet are spending nearly equivalent amounts of time under Alberta Review Board jurisdiction as compared to their male counterparts. This finding runs counter to the general ethos of the *Winko v. British Columbia (Forensic Psychiatric Institute;* 1999) decision wherein if an individual does not pose a significant threat to the safety of the public then they must be granted an absolute discharge; although the possibility exists that other unmeasured variables outside the scope of this study accounted for the time becoming equivalent. It would be a worthwhile research endeavor to determine whether a similar outcome is occurring in other jurisdictions across Canada as it is currently unknown.

Third, the findings from the discrimination and calibration analyses indicated strong discrimination properties, but significant issues with the calibration properties of the VRAG-R on the present sample. However, caution should be noted in interpreting these results as the Alberta NCRMD population was much lower risk than the construction and validation sample of the VRAG-R. Despite the caution needed in the interpretation of the results, it is sobering that

overprediction occurred even at the lowest risk VRAG-R bands, suggesting that baseline actuarial risk does not explain the whole picture. It is also worth noting that in the construction and validation sample of the VRAG-R, new charges, as opposed to convictions (used in the present study) were used as the outcome measure of recidivism. Thus, use of new charges over convictions may have led to more inclusive recidivism counts for the original VRAG-R studies, thereby increasing base rates. In other words, the present study used more stringent criteria for what counted as recidivism (i.e., convictions, not charges). As a result there may have been fewer counts towards recidivism than in comparison to the original VRAG-R studies. This change in how recidivism was defined could have impacted the base rates used within the analyses.

Furthermore, caution should be noted in interpreting these results as there are multiple important systemic issues that need to be considered when reflecting on the results of the discrimination and calibration outcomes of this study. For instance, what an individual is charged with may not be the same as what that individual is eventually convicted of (e.g., aggravated assault being pled down to common assault). This drift in categorization could have resulted in index offenses being counted within the database that were not necessarily representative of the events that occurred in order to bring that individual to the attention of the criminal justice system. This may have an influence on what conclusions are being drawn (e.g., regarding the ratio of general to violent index offenses and the subsequent analyses). Additionally, it is difficult to say how the multiple factors involved in the interactions between law enforcement and individuals who struggle with mental health issues could impact how and with what an individual was charged. For instance factors such as whether the individual was known to the law enforcement officer and whether this impacted the law enforcement officer's willingness or

unwillingness to charge them with a criminal offense; whether the law enforcement officer had any training in interacting with those who struggle with serious mental health issues as this could escalate or de-escalate the situation or impact the kinds of charges being laid. This list is not exhaustive and other unmeasured factors could be involved that are outside of the scope of this study. Another systemic issue that could impact overall rates of persons found NCRMD is access to legal representation. This is a common difficulty for those with lower sociodemographic backgrounds (common for those persons found NCRMD). In many provinces in Canada, Legal Aid is frequently drastically underfunded thus limiting the access for many individuals to a lawyer. Additionally, there is a stigma attached to being found NCRMD (e.g., easily evidenced by the media's treatment of persons found NCRMD) and this stigma may dissuade people from wanting to raise or agree with their lawyer in bringing forth the NCRMD defense. All of these systemic issues (and potentially others not mentioned) could have impacted the results found within this study and caution should be noted in interpreting the results of the discrimination and calibration outcomes of this study.

Fourth, given the study's results regarding the sociodemographic, clinical, and criminological differences between the genders and the fact that the discrimination and calibration properties of the VRAG-R could not currently be examined within the Alberta NCRMD population due to a small number of females who recidivated, examining the predictive validity of VRAG-R with a female population in the future would be a worthwhile research endeavor. Currently, the VRAG-R is not endorsed to be used within female populations and continued research on whether this tool could be used with females who are found NCRMD (or otherwise) would not only increase what is known about the predictive validity of the VRAG-R

but also potentially increase the number of tools available to practitioners working with female offenders.

Given the VRAG-R is not currently endorsed for use with a female population yet there have been 92 females to come under the jurisdiction of the Alberta Review Board it is apparent that some tool is required to assist in the measure of risk determination for review board decisions. It would be unethical to leave such a decision to pure clinical judgement given the evidence provided over how poorly professionals are at predicting violent events (Monahan, 1988). Given this, there are a number of well validated measures that can be use either in place of or in conjunction with the VRAG-R. For instance a clinician may choose to not use the VRAG-R with their female clients or they may choose to couch the VRAG-R within a battery of tools that have been well validated within a female population (e.g., PCL-R, Level of Service measures). Continuing to use the VRAG-R couched within a battery of tools would provide data for future research that could provide the support for or against the VRAG-R's use within a female population.

Additionally, given the differences found within the female NCRMD population a more gender responsive review board would consider additional aspects of a female individual's life history. For instance perhaps, a more serious consideration of her level of risk as determined by well validated measures, trauma history, education level, substance use history, and social support network. Motherhood can complicate treatment and rehabilitation efforts and should be carefully considered, especially given separation can cause additional stressors but also given that females found NCRMD were more likely to offend within private relationships.

Fifth, this study's results highlighted the importance of local norms for risk assessment measures and the importance of conducting risk assessments as an integrated, multi-measure,

multi-source exercise that does not rely upon one single measure to appraise risk or to make decisions. It is clear that the VRAG-R norms would generate considerably higher projections of rates of future violence attached to scores. As such, the present findings may be taken as a set of local Alberta NCRMD norms, which likely represent more realistic portrayals of risk. Even still, the VRAG-R should not be used in isolation and is likely best complemented by a dynamic measure. For instance, Olver and Sewall (2018) found that a measure of sexual violence risk incremented VRAG-R predictions, and that logistic regression could be employed to model recidivism estimates incorporating treatment change information. Future research should extend and replicate the findings with other tools and samples.

Given the issues raised by the outcomes of this study (e.g., over-estimation of risk and inability to run validation analyses with the female population), it begs the question of whether the VRAG-R should be used with this population at all. A simple answer to this question does not exist given what we know about clinicians' poor ability to predict violence. Without the use of the VRAG-R the number of tools left available specific to violence prediction dwindles - though this is not to say there are no other well validated violence risk scales. The Violence Risk Scale (Wong & Gordon, 1999) is a prime example of a possible alternative. It has been demonstrated that when review boards are left to their own devices (i.e., not guided by a structured risk assessment) they frequently fall back on heuristics that have little to do with the prediction of violent outcomes. In fact, in some cases the best predictors of clinician recommendations for release were factors such as psychotropic medication, or patient physical attractiveness (Hilton & Simmons, 2001)! As such, it could be reckless to simply abandon the use of the VRAG-R all together. Instead, with increased research along with using local norms and logistic regression estimates, the VRAG-R may be improved in its use with a specific

NCRMD population. As mentioned before the VRAG-R couched within a multi-tool, multi-source battery of risk assessment (especially when matched with a more dynamic tool) may be a way forward.

Sixth, the results of this study also support the assertion of multiple other researchers in Canada (Grantham, 2014; Charette et al., 2015; Lacroix, O'Shaughnessy, McNiel, & Binder, 2017; Goossens, Nicholls, Charette, Wilson, Seto, & Crocker, 2019) that there is very little evidence that supports the idea that legislation change was needed in order to protect public safety (i.e., Bill C-14). The Alberta NCRMD population has overall low rates of recidivism, and is overall lower risk than many other offending populations (e.g., general offenders). These low rates of recidivism and risk are especially true for those females within the population. As noted above, these results even provide support for the idea that females found NCRMD in Alberta are spending as much time under the jurisdiction of the Alberta Review Board (ARB) despite having fewer criminological characteristics and a much lower recidivism rate than their male counterparts. Such a finding has implications for the delicate balance that review boards have been tasked with in the sense of balancing individual's civil rights with the priority of protecting public safety.

Seventh, the findings of this study have also supported the idea put forth by Demarais et al. (2008) and Grossi and Green (2017) of highlighting the importance of looking at each individual province in order to get the most accurate information regarding those persons found NCRMD in Canada. While national statistics are important in their own right, due to the fact that in Canada NCRMD legislation is federally defined and provincially administered, there may be important differences between the provinces that could be overlooked or missed if all data from the provinces was only ever grouped together. As demonstrated by the results of this study, there

exist some interprovincial differences that have relevance towards how to most effectively meet the unique needs of persons found NCRMD. At the time of this writing, Alberta was the only province to look at population level data for their NCRMD population. Furthermore, not all Canadian provinces are homogenous, each having their own characteristics (i.e., more Liberal or Conservative; population density; rural versus urban divide) which conceivably could impact the way in which the federal legislation is administered. By having both the national level statistics plus provincial level data could help to ensure the most fulsome image of persons found NCRMD in Canada today.

Eighth, the study results fell well within the theoretical background described above regarding risk assessment. Despite being atheoretical in its development the items on the VRAG-R reflect those variables identified as being predictors of crime within the Psychology of Criminal Conduct (PCC) framework. Furthermore, the evidence of potentially higher psychosocial functioning for females within this population in addition to their lower rates of previous criminal sentences and recidivism was directly in line with what would be expected regarding their outcome for risk (i.e., lower scores on the VRAG-R).

Lastly, the findings also highlight psychologists' ethical duty to maintain an appropriate understanding and research support for the measures used within their assessments. Despite the limitations noted in this study of the VRAG-R, it is critical to bear in mind that as an actuarial tool, it is substantially more accurate than unstructured clinical judgment, and the results still support its use within a population of persons found NCRMD when embedded within a comprehensive risk assessment battery of tools.

4.4 Strengths and Limitations

The present study has strengths and limitations, and as such, the results of this study should be interpreted accordingly in light of the following considerations.

This study was archival and retrospective in nature and as such was not free from the limitations that are often associated with such research designs. Being at the mercy of data quality and quantity were tangible issues for this study. The quality and quantity of information in some files was limited and insufficient. This was especially true for the older files (e.g., circa 1940 – 1970). As a result, some data was missing and some files while identified within the database were not able to be included within the VRAG-R analysis due to inadequate file information. Related to this is the issue of missing files. In certain cases, some individuals were identified as being found NCRMD within Alberta but their files could not be located for inclusion and analysis.

Another potential limitation to this study is that any single individual determined NCRMD by the provincial courts within Alberta and who was given an immediate absolute discharge was not captured by this database. This is because these persons would not have come under the jurisdiction of the Alberta Review Board (ARB) and as such there would be no file existing for inclusion. As such this portion of the NCRMD population was unaccounted for in this research. It is believed that such cases were quite rare, however, there was no way of verifying how often such situations occurred.

Another possible limitation to this study is the potential for both legislative and diagnostic cohort effects. Throughout the span of 1941 (i.e., the first person determined NCRMD in the database) to the current day, multiple legislative changes have occurred and the *Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition* (American Psychiatric Association, 2013) has undergone multiple iterations since its inception in 1952. Whether any of these

changes actually resulted in cohort effects was unknown, but the possibility remains and as such the results may need to be interpreted with prudence.

These issues aside, a unique strength of this study is that it has access to the entire population of persons found NCRMD in Alberta's history who have come under the ARB (with the above noted exception). Therefore, questions of sample representativeness should not be at issue. The present study is also the first cross validation of the predictive accuracy of the VRAG-R in an NCRMD population outside of Ontario specifically. A further strength is the examination of potential gender differences within this population within Alberta. At the time of this writing no, other study had sought to examine this facet of research. There are also core methodological strengths that give confidence in the integrity of findings. High quality VRAG-R data with strong interrater agreement was collected, with comprehensive long-term outcome data subsequently captured. The methodological and data conditions are thus ideal for rigorous examination of the predictive properties of VRAG-R scores.

4.5 Future Directions

While this study has sought to fill many of the gaps regarding those individuals in Alberta who have been found NCRMD and the use of the VRAG-R within this population, further research is warranted. For example, while it does appear that males and females are being treated the same via review board dispositions despite their differences, it is unknown what factors are actually going into these decisions. It would be worthwhile to examine the factors review boards are actually taking into consideration when making their decisions. Furthermore, the differences between the genders within this study were not as robust as within other studies in Canada (e.g., the National Trajectory Project) and it might be worthwhile to investigate the reasons for such a finding. Considering the fact that NCRMD legislation is federally defined and

provincially administered this is not entirely surprising yet knowing what the provincial differences in how the laws are applied and administered may help to further inform on policy and treatment.

While it was outside the scope of this research, it would have been noteworthy to examine the victim profiles of this population. Previous research has shown that females were more likely to offend against a family member or close friend while males were more likely to offend against an acquaintance or stranger (Nicholls, Crocker, Seto, Wilson, Charette, & Côté, 2015). Given the consistency of the findings in this study it would be noteworthy to see if this finding would have been maintained within the Alberta population as well. Such a finding could have implications for decision making regarding review board decisions.

A aspect that was not answered by this study but would be a worthwhile research endeavor would be to determine if the number of females being found NCRMD is increasing in both Alberta and in Canada. Currently in the federal correctional system there has been a steady increase in the number of women entering the system. It could be worthwhile to determine if the same was occurring for women coming under the jurisdiction of the provincial and territorial review boards (Sapers, 2016).

Little is known about the phenomenological experience of those persons who are found NCRMD. While much of the current research done on this population is quantitative, a more qualitative approach to the lived experiences of those persons found NCRMD might provide some insight into their lived experience and could potentially better inform on treatment, motivation to engage treatment or even how to better connect individuals to supports (i.e., clarify perceived barriers to services or supports).

In general, psychologists have an ethical duty to protect and promote the wellbeing of those with whom they work (Canadian Psychological Association, 2017*).* Part of this responsibility is maintaining knowledge and competence within the populations with which they work and the measures that they use as a part of their assessments. Continued study on the use of the VRAG-R within certain populations and ensuring that the most accurate up to date tools are available for use within this population is good practice. Thus, in general more research is needed for those individuals found NCRMD and the tools used within. It is useful to note that *The Alberta NCR Project* is continuing to work towards this goal of a greater understanding of this population and the risk assessment tools that are used within it.

4.6 Conclusion

This was the first large scale investigation to examine both potential gender differences and the predictive accuracy of the VRAG-R within the population of persons found Not Criminally Responsible on Account of Mental Disorder (NCRMD) in Alberta. Gendered differences exist within this population and while these differences were not as robust as in other provinces in Canada, they are important characteristics that need to be taken into account when deciding upon the adjudication of this population. For example, females who have been found NCRMD in Alberta, appear to have a slightly higher likelihood of completing high school, lower overall risk scores and drastically lower rates of recidivism. Males on the other hand had a higher number of previous sentences than their female counterparts. Both genders had a unique clinical profile, a finding that was not surprising given many *DSM-5* diagnoses have gendered prevalence rates. Despite these differences, males and females were treated similarly via review board dispositions. A finding that was surprising given the task of the review boards to protect public safety but also respect civil liberties by not retaining those who are deemed to no longer be a

danger to the public. The VRAG-R demonstrated strong discrimination but concerning calibration properties. The results demonstrated the importance of psychologists maintaining their ethical duty to ensure that the tools used for treatment and assessment are valid and reliable within the populations in which they are applied. The VRAG-R may still be used within this population so long as its limitations are noted and it is embedded within a comprehensive risk assessment battery of tools. That said, the present findings contraindicate the use of the VRAG-R norms from Rice et al. (2013), as they simply inflate violent recidivism estimates in an Alberta NCRMD population. The implications are: a) to use the present findings as a set of local violent recidivism norms for use of the VRAG-R in making violence risk appraisals; and b) to use the VRAG-R in combination with a dynamic risk tool such as the HCR-20 V3, the VRS, and/or the LSI measures. Research demonstrates that dynamic tools increment predictions of future violence beyond the VRAG-R (e.g., Olver & Sewall, 2018), in addition to being able to fulfill the other essential risk management functions of the RNR framework, not the least of which includes being able to capture changes in risk.

In conclusion, the more that is known about individuals who are found NCRMD, the better policy or legislation, treatment, and assessment can be tailored to effectively meet their unique needs. Given that review boards have been recently instructed to place public safety as their paramount consideration, the results of this study can help to refine how much risk should be allocated to those persons found NCRMD and create a better balance between entrenched civil liberties and public safety.

REFERENCES

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., … Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*. *(34)*3, 341-382.

Akers, R. L., (1973). *Deviant behavior: A social learning approach.* Belmont, CA: Wadsworth.

Allnutt, S., Samuels, A., & O'driscoll, C. (2007). The insanity defense: From wild beasts to M'Naghten. *Australasian Psychiatry: Bulletin of Royal Australian and New Zealand College of Psychiatrists*, *15*(4), 292-298.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct.* Washington, DC: Author.

Andrews, D. A., & Bonta, J. (1995). *LSI-R: The Level of Service Inventory manual.* North Tonawanda, NY: Multi-Health Systems.

Andrews, D. A., & Bonta, J., (2015). *The psychology of criminal conduct*. (5th ed.) New York, NY: Anderson Publishing.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19-52.

Andrews, D. A., Bonta, J., & Wormith, S. J. (2004). *The Level of Service/Case Management Inventory (LS/CMI)*. Toronto, Canada: Multi-Health Systems.

Andrews, D.A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency*, *52*(1), 7–27.

Appelbaum, P.S. (2013) Does the constitution require an insanity defence? *Psychiatric Services 64*(10), 943-945.

Babchishin, K. M., & Helmus, L.-M. (2016). The influence of base rates on correlations: An evaluation of proposed alternative effect sizes with real-world data. *Behavior Research Methods*, *48*, 1021–1031.

Baillie, P. (2015). A valuable (and ongoing) study, The National Trajectory Project addresses many myths about the verdict of not criminally responsible on account of mental disorder. *Canadian Journal of Psychiatry*, *60*(3), 93–96.

Bal, P., & Koenraadt, F. (2000). Criminal law and mentally ill offenders in comparative perspective. *Psychology, Crime & Law*, *6*(4), 219–250.

Bill C-10: An Act to amend the Criminal Code (mental disorder) and to make consequential amendments to other Acts. (2005). Royal Assent May 19, 2005. 38th Parliament, 1st session, Retrieved from the parliament of Canada website: http://www.parl.ca/LegisInfo/BillDetails.aspx?Language=E&billId=1395895

Bill C-14: An Act to Amend the Criminal Code and the National Defense Act. (2014). Royal Assent April 11, 2014. 41st Parliament, 2nd Session. Retrieved from the parliament of Canada website: http://www.parl.gc.ca/HousePublications/Publication.aspx?Language=E&Mode=1&DocId=6540830&File=4

Bill C-30: An Act to amend the Criminal Code and to amend the National Defence Act and the Young Offenders Act in consequence thereof. (1992). Royal Assent 1991(assented to 1991, c. 43, s. 4, proclaimed in force February 4, 1992).

Blackburn, R. (1995). *The psychology of criminal conduct: Theory, research and practice*. Chichester, UK: Wiley.

Blanchette, K. (2004). Revisiting effective classification strategies for women offenders in Canada. *Feminism & Psychology*, *14*(2), 231–236.

Block, R. C., Blokland, A. A. J., van der Werff, C., van Os, R., & Nieuwbeerta, P. (2010). Long-term patterns of offending in women. *Feminist Criminology*, *5*(1), 73–107.

Bonta, J., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior, 19*(3), 278–287.

Bonta, J., Dauvergne, M., Rugge, T. (2003). The reconviction rate for federal offenders 2003–02. (Catalog number: JS42-100/2002E-IN). Retrieved from Public Safety Canada: https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/rcnvctn-rt-fdrl/index-en.aspx

Bonta, J., Law, M., & Hanson, R. K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: a meta-analysis. *Psychological Bulletin*, *123*(2), 123–142.

Brodsky, G. G., Defence, C. R., & Information, C. (2017). Proceed with extreme caution: The not criminally responsible defence. *Manitoba Law Journal*, *40*, 89–116.

Brook, M. (2017). Structured approaches to violence risk assessment: A critical review. *Psychiatric Annals*, *47*(9), 454–459.

Brown, J., & Singh, J. P. (2014). Forensic risk assessment: A beginner's guide. *Archives of Forensic Psychology*, *1*(1), 49–59.

Brown, S., Serin, R., Forth, A., Nunes, K., Bennell, C., Pozzulo, J., (2017). *Psychology of criminal behaviour: A Canadian perspective* (2nd ed.). Toronto, Canada: Pearson.

Bureau of Federal Prisons. (2017). *Inmate gender*. Retrieved from https://www.bop.gov/about/statistics/statistics_inmate_gender.jsp

Burgess, R. L., & Akers, R. L., (1966). A differential association-reinforcement theory of criminal behaviour. *Social Problems*, *14*(2), 128-147.

Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders. *Criminal Justice and Behavior*, *36*(6), 567–590.

Canadian Charter of Rights and Freedoms, s 2, Part I of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), 1982, c 11.

Canadian Psychological Association. (2017). *Canadian code of ethics for psychologists (4th Ed.)*. Retrieved from: https://cpa.ca/docs/File/Ethics/CPA_Code_2017_4thEd.pdf

Charette, Y., Crocker, A. G., Seto, M. C., Salem L., Nicholls, T. L., Caulet, M. (2015). The National Trajectory Project of individuals found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 4: Criminal recidivism. *Canadian Journal of Psychiatry*, *60*(3), 127–134.

Chesney-Lind, M. (2000) 'What to do about girls? Thinking about programs for young women'. In M. McMahon (Ed.) *Assessment to Assistance: Programs for Women in Community Corrections*, (pp. 139–70). Arlington, VA: American Correctional Association.

Cocozza, J. J., & Steadman, H. J., (1976). The failure of psychiatric predictions of dangerousness: Clear and convincing evidence. *Rutgers Law Review*, *29*(5), 1084-1101.

Coid, J., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Roberts, C., Farrington, D. P., & Rogers, R. D. (2009). Gender differences in structured risk assessment: Comparing the accuracy of five instruments. *Journal of Consulting and Clinical Psychology*, *77*(2), 337–348.

Cox, J., Fairfax-Columbo, J., DeMatteo, D., Vitacco, M. J., Kopkin, M. R., Parrott, C. T., & Bownes, E. (2018). An update and expansion on the role of the Violence Risk Appraisal

Guide and Historical Clinical Risk Management-20 in United States case law. *Behavioral Sciences and the Law*, *36*(5), 517–531

Criminal Code, S.C. 1892, c. 29, s.11.

Criminal Code, R.S.C 1985

Criminal Code, R.S.C 1985, c. C-46, s. 2.

Criminal Code, R.S.C 1985, c. C-46, s. 16.

Crocker, A. G., Charette, Y., Seto, M. C., Nicholls, T. L., Côté, G. ,& Caulet, M. (2015). The National Trajectory Project of individuals found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 3: Trajectories and outcomes through the forensic system. *Canadian Journal of Psychiatry*, *60*(3), 117–126.

Crocker, A. G., Nicholls, T. L., Charette, Y., & Seto, M. C. (2014). Dynamic and static factors associated with discharge dispositions: The National Trajectory Project of individuals found Not Criminally Responsible on Account of Mental Disorder (NCRMD) in Canada. *Behavioral Sciences and the Law*, *32*, 577-595.

Crocker, A. G., Nicholls, T. L., Seto, M. C., & Côté, G. (2015). The National Trajectory Project of individuals found Not Criminally Responsible on Account of Mental Disorder in Canada. *Canadian Journal of Psychiatry*, *60*(3), 96-97.

Crocker, A. G., Nicholls, T. L., Seto, M. C., Côté, G., Charette, Y., & Caulet, M. (2015). The National Trajectory Project of individuals found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 1: Context and methods. *Canadian Journal of Psychiatry*, *60*(3), 98–105.

Crocker, A. G., Nicholls, T. L., Seto, M. C., Charette, Y., Côté, G., & Caulet, M. (2015). The National Trajectory Project of individuals found Not Criminally Responsible on Account

of Mental Disorder in Canada. Part 2: The people behind the label. *Canadian Journal of Psychiatry*, *60*(3), 106–116.

Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *Southern California Review of Law and Women's Studies*, *2*, 11-52.

Daly, K. (1994). *Gender, crime, and punishment*. New Haven, CT: Yale University Press.

Daly, K., & Chesney-Lind, M. (1988). Feminism and criminology. *Justice Quarterly*, *5*, 497-538.

Delisi, M. (2002). Not just a boy's club: An empirical assessment of female career criminals. *Women & Criminal Justice*, *13*(4), 27–45.

de Vogel, V., & Nicholls, T. L. (2016). Gender matters: An introduction to the special issues on women and girls. *International Journal of Forensic Mental Health*, *15*(1), 1–25.

Desmarais, S. L., Hucker, S., Brink, J., & De Freitas, K. (2008). A Canadian example of insanity defence reform: Accused found Not Criminally Responsible before and after the Winko decision. *International Journal of Forensic Mental Health*, *7*(1), 1–14.

Derkzen, D., Booth, L., Taylor, K., & McConnell, A. (2013). Mental health needs of federal female offenders. *Psychological Services*, *10*(1), 24–36.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk of violence – User guide*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Dowden, C. (2015). Risk principle of case. *International Journal of Offender Therapy and Comparative Criminology*, *50*(1), 88–100.

Eisenbarth, H., Osterheider, M., Nedopil, N., & Stadtland, C. (2012). Recidivism in female offenders: PCL-R lifestyle factor and VRAG show predictive validity in a German sample. *Behavioral Sciences and the Law*, *30*(5), 575–584.

Ewert *v.* Canada, 2 S.C.R. 165 (2018)

Ferguson, M., & Ogloff, J. (2011). Criminal responsibility evaluations: Role of psychologists in assessment. *Psychiatry, Psychology and Law*, *18* (1), 79-94.

Fazel, S., Fimińska, Z., Cocks, C., & Coid, J. (2016). Patient outcomes following discharge from secure psychiatric hospitals: Systematic review and meta-analysis. *British Journal of Psychiatry*, *208*(1), 17–25.

Friendship, C., McClintock, T., Rutter, S., & Maden, A. (1999). Re-offending: Patients discharged from a regional secure unit. *Criminal Behaviour and Mental Health*, *9*(3), 226–236.

Folsom, J., & Atkinson, J. L. (2007). The generalizability of the LSI-R and the CAT to the prediction of recidivism in female offenders. *Criminal Justice and Behavior*, *34*(8), 1044–1056.

Geraghty, K. A., & Woodhams, J. (2015). The predictive validity of risk assessment tools for female offenders : A systematic review. *Aggression and Violent Behavior*, *21*, 25–38.

Glover, A. J. J., Churcher, F. P., Gray, A. L., Mills, J. F., & Nicholson, D. E. (2017). A cross-validation of the Violence Risk Appraisal Guide—Revised (VRAG–R) within a correctional sample. *Law and Human Behavior*, *41*, 507-518.

Gleuck, S., & Glueck, E., (1950). *Unravelling juvenile delinquency.* Cambridge, MA: Harvard University Press.

Gleuck, S., & Glueck, E., (1968). *Delinquents and non-delinquents in perspective.* Cambridge, MA: Harvard University Press.

Goossens, I., Nicholls, T. L., Charette, Y., Wilson, C. M., Seto, M. C., & Crocker, A. G. (2019). Examining the high-risk accused designation for individuals found Not Criminally Responsible on Account of Mental Disorder. *Canadian Psychology*, *60*(2), 102–114.

Gottfredson M. R., & Hirschi, T., (1990). *A general theory of crime*. Standford, CA: Standford University Press.

Grantham, L. (2014). Bill C-14: A step backwards for the rights of mentally disordered offenders in the Canadian criminal justice system. *Appeal: Review of Current Law and Law Reform*, *19*(1), 63-81.

Grann, M., Danesh, J., & Fazel, S. (2008). The association between psychiatric diagnosis and violent re-offending in adult offenders in the community. *BMC Psychiatry*, *8*, 1–7.

Grant, I. (1997). Canada's new mental disorder disposition provisions. *International Journal of Law and Psychiatry*, *20*(4), 419–443.

Gregório Hertz, P., Eher, R., Etzler, S., & Rettenberger, M. (2019). Cross-validation of the revised version of the Violence Risk Appraisal Guide Revised (VRAG-R) in a sample of individuals convicted of sexual offenses. *Sexual Abuse: Journal of Research and Treatment*. Advance online publication.

Grossi, L. M., & Green, D. (2017). An international perspective on criminal responsibility and mental illness. *Practice Innovations*, *2*(1), 2–12.

Haag, A. M., Cheng, J., & Wirove, R. (2016). Describing the not criminally responsible population in Alberta's history: Sociodemographic, mental health and criminological profiles. *Journal of Community Safety and Well-Being*. *1* (3), 68 – 74.

Hanson, R. K. (2017). Assessing the calibration of actuarial risk scales. *Criminal Justice and Behavior*, *44*(1), 26–39.

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: comparison of three actuarial scales. *Law and Human Behavior*, *24*, 119-136.

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist* (2nd ed.). Toronto, ON, Canada: Multi-Health Systems.

Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior*, *26*(4), 377-394.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20,* 315-355.

Harris, G. T., Rice, M. E., Quinsey, V. L., & Cormier, C. A. (2015). *Violent offenders: Appraising and managing risk* (3rd ed.). Washington, DC: American Psychological Association.

Hastings, M.E., Krishnan, S., Tangney, J.P., & Stuewig, J. (2011). Predictive and incremental validity of the Violence Risk Appraisal Guide scores with male and female jail inmates. *Psychological Assessment, 23*, 174-183.

Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *The Hare Psychopathy Checklist: Screening Version* (PCL:SV). Toronto: Multi-Health Systems.

Hayes, H., Kemp, R. I., Large, M. M., & Nielssen, O. B. (2014). A 21-year retrospective outcome study of New South Wales forensic patients granted conditional and unconditional release. *Australian and New Zealand Journal of Psychiatry*, *48*(3), 259–282.

Heilbrun, K., Fairfax-columbo, J., Wagage, S., & Brogan, L. (2017). Risk assessment for future offending: Value and limits of expert evidence at sentencing. *Court Review*, *53*, 116–125.

Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 64-101.

Hiday, V. A. (1997). Understanding the connection between mental illness and violence. *International Journal of Law and Psychiatry*, *20*(4), 399–417.

Hilton, N. Z., & Simmons, J. L. (2001). The Influence of Actuarial Risk Assessment in Clinical Judgments and Tribunal Decisions about Mentally Disordered Offenders in Maximum Security. *Law and Human Behavior*, *25*(4), 393–408.

Hirschi, T., (1969). *Causes of delinquency*. New Brunswick, NJ: Transaction Publishers.

Hogan, N. R., & Olver, M. E. (2016). Assessing risk for aggression in forensic psychiatric inpatients: An examination of five measures. *Law and Human Behavior*, *40*(3), 233–243.

Hogan, N. R., & Olver, M. E. (2019). Static and dynamic assessment of violence risk among discharged forensic patients. *Criminal Justice and Behavior*, *46*(7), 923–938.

Johnson S.L., & Grant, B. (2000). Release outcomes of long-term offenders. *FORUM on Corrections Research*. *12*(3), 16–20.

Kong, R., AuCoin, K., (2008). Female Offenders in Canada. *Juristat: Canadian Centre for Justice Statistics.* (Catalogue No. 85-002-XIE, 28[1]). Retrieved from Statistics Canada website: http://www5.statcan.gc.ca/olc-cel/olc.action?objId=85-002-X200800110509&objType=47&lang=en&limit=0

Lacroix, R., O'Shaughnessy, R., McNiel, D. E., & Binder, R. L. (2017). Controversies concerning the Canadian not criminally responsible reform act. *Journal of the American Academy of Psychiatry and the Law*, *45*(1), 44–51.

Latimer J, & Lawrence A. (2006). *The review board systems in Canada: Overview of results from the mentally disordered accused data collection study.* Ottawa (ON): Department of Justice Canada.

Livingston, J. D., Wilson, D., Tien, G., & Bond, L. (2003). A follow-up study of persons found Not Criminally Responsible on Account of Mental Disorder in British Columbia. *Canadian Journal of Psychiatry*, *48*(6), 408–415.

Logan and Blackburn (2009). Canada's new mental disorder disposition provisions: A case study of the British Columbia criminal code review board. *International Journal of Law and Psychiatry*, *20*(4), 419-443.

Maden, A., Skapinakis, P., Lewis, G., Scott, F., Burnett, R., & Jamieson, E. (2006). Gender differences in reoffending after discharge from medium-secure units. *British Journal of Psychiatry*, *189*(2), 168–172.

McNiel, D. E., & Binder, R. L. (1994). Screening for risk of inpatient violence: Validation of an actuarial tool. *Law and Human Behavior*, *18*, 579–586.

Mills, J. F. (2017). Violence risk assessment: A brief review, current issues, and future directions. *Canadian Psychology*, *58*(1), 40–49.

Monahan, J. (1988). Risk assessment of violence among the mentally disordered: Generating useful knowledge. *International Journal of Law and Psychiatry*, *11*, 249-257.

Monahan, J. (2002). The MacArthur studies of violence risk. *Criminal Behaviour and Mental Health*, *12*(S1), S67–S72.

Monahan J, Steadman H, Silver E, Appelbaum P, Robbins P, Mulvey E, Roth L, Grisso T, Banks S (2001) *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence.* New York: Oxford University Press.

Nicholls, T. L., Crocker, A. G., Seto, M. C., Wilson, C. M., Charette, Y., & Côté, G. (2015). The National Trajectory Project of individuals found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 5: How essential are gender-specific forensic psychiatric services? *Canadian Journal of Psychiatry*, *60*(3), 135–145.

Nicholls, T. L., Cruise, K. R., Greig, D., & Hinz, H. (n.d.). Female offenders. In *APA handbook of forensic psychology, Vol. 2: Criminal investigation, adjudication, and sentencing outcomes.*(pp. 79–123). Washington: American Psychological Association.

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing risk for violence among male and female civil psychiatric patients: The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law*, *22*(1), 127–158.

Nicholls, T. L., Wilson, C., Charette, Y., Crocker, A. G., & Seto, M. C. (2015, March). *Aboriginal participants in the National Trajectory Project: Examining profiles, processing, and recidivism of persons found Not Criminally Responsible on Account of Mental Disorder through a cultural lens*. Paper presented at the American Psychology-Law Society, San Diego, CA.

Norko, M.A., Tobias, W., Magro, H., Leavitt-Smith, E., Morton, F.J., & Hollis, T. (2016). Assessing insanity acquittee recidivism in Connecticut. *Behavioral Sciences and the Law. 34*, 423-443.

Olver, M. E., & Sewall, L. (2018). Cross-validation of the discrimination and calibration properties of the VRAG-R in a treated sexual offender sample. *Criminal Justice and Behavior, 45*, 741-761.

Pallant, J. (2013) *SPSS Survival Manual* (5th ed.). Berkshire, England: McGraw-Hill.

Penney, S. R., Seto, M. C., Crocker, A. G., Nicholls, T. L., Grimbos, T., Darby, P. L., & Simpson, A. I. F. (2018). Changing characteristics of forensic psychiatric patients in Ontario: A population-based study from 1987 to 2012. *Social Psychiatry and Psychiatric Epidemiology*, *54*(5), 627–638.

Pilon, M. (2002). *Mental Disorder and Canadian Criminal Law.* (PRB 99-22E). Retrieved from Library of parliament website: http://publications.gc.ca/collections/Collection-R/LoPBdP/PRB-e/PRB9922-e.pdf

Privy Council Office. (1956). *Report of the royal commission on the law of insanity as a defence in criminal cases.* Hull: Queen's Printer and Controller of Stationery. Retrieved from: http://publications.gc.ca/site/eng/472632/publication.html

Public Safety Canada Portfolio Corrections Statistics Committee. (2017). Corrections and conditional release statistical overview: 2016 annual report (Cat. No.: PS1-3E-PDF). Retrieved from: https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ccrso-2016/ccrso-2016-en.pdf

Quinsey, V. L., Rice, M. E., Harris, G. T., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd Ed.). Washington, DC: American Psychological Association.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk (*3rd Ed.). Washington, DC: American Psychological Association.

Reitano, J. (2017). *Adult correctional statistics in Canada , 2015 / 2016.* Retrieved from http://www.statcan.gc.ca/pub/85-002-x/2016001/article/14318-eng.htm

R. v. Demers, 2 S.C.R. 489. (2004).

R. v. Swain, 1.S.C.R. 933 (1991).

Raaflaub, W. (2005). *The mental disorder provisions of the criminal code*. (PRB 05-05E). Retrieved from The Library of Parliament website https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/LegislativeSummaries/381LS481E

Report of the Standing Committee of Justice and Human Rights. (2002). *Review of the mental disorder provisions of the criminal code*. Retrieved from http://www.ourcommons.ca/Content/Committee/371/JUST/Reports/RP1032130/justrp14/justrp14-e.pdf

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide-Revised (VRAG-R). *Psychological Assessment*, *25*(3), 951–65.

Richer, K., Cheng, J., & Haag, A. M. (2018). Historical recidivism rates of Alberta's Not Criminally Responsible population. *Journal of Community Safety and Well-Being*, *3*(2), 59.

Sapers, H. (2016). Annual report of the office of the correctional investigator (Cat. No.: PS100E-PDF) Retrieved from: http://www.oci-bec.gc.ca/cnt/rpt/pdf/annrpt/annrpt20152016-eng.pdf

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., … Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health*, *13*(3), 193–206.

Simpson, A. I. F., Chatterjee, S., Duchcherer, M., Ray, I., Prosser, A., & Penney, S. R. (2018). Short-term outcomes for forensic patients receiving an absolute discharge under the Canadian Criminal Code. *Journal of Forensic Psychiatry and Psychology*, *29*(6), 867–881.

Statistics Canada (2014). *Verdicts of not criminally responsible on account of mental disorder in adult criminal courts, 2005/2006-2011/2012*. Retrieved from http://www.statcan.gc.ca/pub/85-002-x/2014001/article/14085-eng.htm

Statistics Canada (2016) *Parole, Pardons and Clemency.* Retrieved from: https://www.canada.ca/en/parole-board/corporate/publications-and-forms/fact-sheets/statistics-parole-pardons-and-clemency.html

Statistics Canada (2017) *Aboriginal peoples in Canada: Key results from the 2016 census*. Retrieved from http://www.statcan.gc.ca/daily-quotidien/171025/dq171025a-eng.pdf

Statistics Canada (2017), *Court, adult cases by type of sentence, total guilty cases by province and territory*. Retrieved from: https://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/legal22a-eng.htm

Sutherland, E. H., (1949). *Principles of Criminology* (4th ed). Philadelphia, PA: J.B. Lippincoot Company.

Tabita, B., de Santi, M. G., & Kjellin, L. (2012) Criminal recidivism and mortality among patients discharged from a forensic medium secure hospital, *Nordic Journal of Psychiatry*, *66*(4), 283-289.

Taylor, K. N., & Blanchette, K. (2009). The women are not wrong: It is the approach that is debatable. *Criminology & Public Policy*, *8(*1), 221–229.

Tye, C. S., & Mullen, P. E. (2006). Mental disorders in female prisoners. *Australian and New Zealand Journal of Psychiatry*, *40*, 266–271.

Villeneuve, D. B., & Quinsey, V.L., (1995). Predictors of general and violent recidivism among mentally disordered inmates. *Criminal Justice and Behavior*, *22*(4), 387–410.

Viljoen, J. L., McLachlan, K., & Vincent, G. M. (2010). Assessing violence risk and psychopathy in juvenile and adult offenders: A survey of clinical practices. *Assessment*, *17*(3), 377–395.

Viljoen, J. L., Roesch, R., Ogloff, J. R. P., & Zapf, P. A. (2003). The role of Canadian psychologists in conducting fitness and criminal responsibility evaluations. *Canadian Psychology*, *44*(4), 369–381.

Vitacco, M. J., Erickson, S. K., Kurus, S., Apple, B. N., Lamberti, J. S., & Gasser, D. (2011). Evaluating conditional release in female insanity acquittees: A risk management perspective. *Psychological Services*, *8*(4), 332–342.

Vitacco, M. J., Vauter, R., Erickson, S. K., & Ragatz, L. (2014). Evaluating conditional release in not guilty by reason of insanity acquittees: A prospective follow-up study in Virginia. *Law and Human Behavior*, *38*(4), 346–356.

Vitacco, M. J., Van Rybroek, G. J., Erickson, S. K., Rogstad, J. E., Tripp, A., Harris, L., & Miller, R. (2008). Developing services for insanity acquittees conditionally released into the community: Maximizing success and minimizing recidivism. *Psychological Services*, *5*(2), 118–125.

Wattanaporn, K. A., & Holtfreter, K. (2014). The impact of feminist pathways research on gender-responsive policy and practice. *Feminist Criminology*, *9*(3), 191–207.

Ward, T., Melser, J., & Yates, P. M. (2007). Reconstructing the Risk -Need - Responsivity model : A theoretical elaboration and evaluation. *Aggression and Violent Behavior*, *12*, 208–228.

Wilson, C. M., Crocker, A. G., Nicholls, T. L., Charette, Y., & Seto, M. C. (2015). The use of risk and need factors in forensic mental health decision-making and the role of gender and index offense severity. *Behavioral Sciences and the Law*, *33*(1), 19–38.

Winko v. British Columbia (Forensic Psychiatric Institute), 2 S.C.R. 625 (1999).

Wong, S., & Gordon, A. E. (1999-2003). *The Violence Risk Scale.* Unpublished manuscript, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Xie, L. (2000). Gender difference in mentally ill offenders: A nationwide Japanese study. *International Journal of Offender Therapy and Comparative Criminology*, *44*(6), 714–724.

Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*(5), 740–767.

Zinger. I. (2018). The office of the correctional investigator: Annual report. (Cat. No.: PS100) Retrieved from: https://www.oci-bec.gc.ca/cnt/rpt/pdf/annrpt/annrpt20172018-eng.pdf