ABSTRACT

LOWDER, EVAN MARIE. Risk Assessment and Recidivism in the Context of Mental Health Diversion Programs. (Under the direction of Dr. Sarah L. Desmarais).

Developed in response to the overrepresentation of adults with mental illnesses in U.S. jails, mental health diversion programs are now a common strategy to connect justice-involved adults with mental illnesses to community-based treatment in lieu of incarceration. Although research on such programs has grown substantially in recent decades, findings have not provided consistent support for their effectiveness in improving criminal justice outcomes. Rather, the effectiveness of diversion programs has been shown to vary widely across sites and studies. Given the multifaceted risks and needs of justice-involved adults with mental illnesses, risk management strategies may play a role in improving recidivism outcomes in this population. However, few studies to date have examined the use of risk assessments in this setting or investigated the relative importance of risk and protective factors in predicting recidivism in diversion participants. Additionally, few studies have identified factors (e.g., participant characteristics, program components) associated with diversion program effectiveness, limiting our understanding of mechanisms accounting for the variable effectiveness of diversion programs. In three samples of mental health diversion participants (total $N = 742$), the included manuscripts aim to address these limitations.

In the first manuscript (Chapter 2), I examine the psychometric properties of assessments produced by the Short-Term Assessment of Risk and Treatability (START) and the Level of Service Inventory – Revised (LSI-R) in 95 mental health diversion participants. Overall, results showed good levels of inter-rater reliability for both START and LSI-R
assessments. Although both START and LSI-R assessments demonstrated validity in predicting arrests and jail days across multiple follow-ups, predictive validity estimates were most consistent and strongest for START strength scores, measuring protective factors.

In the second manuscript (Chapter 3), I evaluate five hypothesized models (i.e., Compensatory, Buffer, Challenge, Protective-Protective, and Mediation) describing associations between risk factors, protective factors, and recidivism in 550 mental health diversion program clients. Using hierarchical regression models, results showed direct and independent associations of risk and protective factors on both arrests and jail days over a 1-year period, consistent with the Compensatory model. However, results failed to provide evidence supporting the other four models.

In the third manuscript (Chapter 4), I examine between-group differences in recidivism outcomes as a function of MHC participation and within-group changes in jail days as a function of participant- and process-related factors. Relative to a propensity-scored matched group of offenders receiving treatment as usual ($n = 40$), MHC participants ($n = 57$) had fewer jail days, but not charges or convictions, 1-year following program exit. In MHC participants, factors associated with greater reductions in jail days included graduation from the MHC, presence of co-occurring substance use, and longer time in the program.

Together, findings support the predictive utility of risk assessments administered in mental health diversion programs and suggest that protective factors may be especially relevant to the risk management of diversion participants. Further, findings support the effectiveness of longer-term engagement in mental health diversion programs and suggest that high-risk participants, such as those with co-occurring disorders, may benefit from
participation in structured diversion programs, like MHCs. Future research is needed to inform the development of risk management strategies, including strengths-based approaches, that may increase engagement in diversion services and ultimately improve criminal justice outcomes among mental health diversion participants.
Risk Assessment and Recidivism in the Context of Mental Health Diversion Programs

by

Evan Marie Lowder

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Psychology

Raleigh, North Carolina

2017

APPROVED BY:

Dr. Sarah Desmarais
Committee Chair

Dr. Jennifer Burnette

Dr. Joan Pennell

Dr. Joseph Simons-Rudolph
DEDICATION

To women everywhere.
BIOGRAPHY

Evan Marie Lowder was born to Jody and John Lowder and raised together with her three siblings in Eagan, Minnesota. She attended the College of Saint Benedict in St. Joseph, Minnesota, graduating summa cum laude in 2012 with a Bachelor of Arts in Psychology and Political Science. In 2013, she joined the Applied Social and Community Psychology program at North Carolina State University in Raleigh, North Carolina under the advisement of Dr. Sarah Desmarais. Evan’s research interests have focused broadly on improving behavioral health and criminal justice outcomes for justice-involved adults with mental illnesses. She completed her Master’s degree in 2015, defending her thesis entitled *Models of Protection Against Recidivism in Justice-Involved Adults with Mental Illnesses*. Evan lives with her husband, Kevin, and puppies, Jack and Cinnamon.
ACKNOWLEDGMENTS

I am grateful for many mentors, family members, and friends who have supported me throughout my academic journey. First, I would like to thank the many mentors who fostered my development as a scientist and academic. My advisor and mentor, Dr. Sarah Desmarais, deserves special recognition for her tutelage over the past four years. Your instruction and guidance (in all matters!) has been invaluable, and I am thankful for the many opportunities I received under your mentorship. You set high expectations for your graduate students, and in doing so you have made us all better scientists. I am grateful also to the other members of the North Carolina State University academic community who have provided informal and formal instruction and contributed to my professional development over the past several years. Together, you have provided a model of collegiality that I hope to take with me in my academic career. A special thank you to my dissertation committee—Dr. Joan Pennell, Dr. Joe Simons-Rudolph, and Dr. Jeni Burnette—for your helpful feedback throughout my doctoral work and for your confidence in me.

Second, my family deserves a special shout out for supporting my academic goals from near and far. Thank you especially to my mom, Jody, and my sister, Ryan, for weekly phone calls and Google Hangouts sessions. I am grateful for family members who visited us in North Carolina and helped make our life here a little more like home, especially my in-laws, Diane and Larry, who made Christmas in North Carolina a reality two years in a row. The greatest thanks of all goes to my husband, Kevin, who entertained countless thought experiments, served as my go-to peer reviewer, and spent many Sunday afternoons working
by my side. I am increasingly aware that being in a stable, supportive, and loving relationship is a privilege denied to many, and every day I give thanks that you are in my life.

Third and finally (and last but certainly not least), I would like to thank my friends who have been a source of companionship and motivation over the past four years. I am thankful for my friends from the College of Saint Benedict, who are some of the strongest women I know and motivate me to be good and do good for others. To my peers in the Applied Social and Community Psychology graduate program, thank you for being such a supportive and motivating force in my life for the past four years. To Kiersten and Candalyn, thank you both for your consultation on all matters research, for productivity dates, and for your supportive words and actions. Becoming your peer, collaborator, and friend was among the highlights of my graduate school experience.
# TABLE OF CONTENTS

**LIST OF TABLES** ........................................................................................................... vii

**LIST OF FIGURES** ........................................................................................................ viii

**Chapter 1: Introduction** .................................................................................................. 1
    Purpose............................................................................................................................ 7
    Overview of Research Chapters...................................................................................... 8

**Chapter 2: Reliability and Validity of START and LSI-R Assessments in Mental Health Jail Diversion Clients** ................................................................................................................ 11
    Introduction........................................................................................................................ 12
    Method.............................................................................................................................. 20
    Results.............................................................................................................................. 28
    Discussion.......................................................................................................................... 33

**Chapter 3: Models of Protection against Recidivism in Justice-Involved Adults with Mental Illnesses** ......................................................................................................................... 50
    Introduction........................................................................................................................ 52
    Method.............................................................................................................................. 60
    Results.............................................................................................................................. 66
    Discussion.......................................................................................................................... 70

**Chapter 4: Recidivism Following Mental Health Court Exit: Between and Within-Group Comparisons** ......................................................................................................................... 80
    Introduction........................................................................................................................ 82
    Method.............................................................................................................................. 88
    Results.............................................................................................................................. 94
    Discussion.......................................................................................................................... 99

**Chapter 5: Integrative Review** ...................................................................................... 110
    Summary of Findings........................................................................................................ 111
    Implications for Research............................................................................................... 113
    Implications for Practice................................................................................................. 118
    Implications for Policy..................................................................................................... 119
    Limitations and Future Directions.................................................................................. 121
    Conclusion........................................................................................................................ 123

**REFERENCES** .................................................................................................................. 125
LIST OF TABLES

CHAPTER 2
Table 1. Descriptive Statistics for Arrests and Jail Days for All Follow-Up Periods.................................................................................................................. 43
Table 2. Distribution of and Concordance between LSI-R and START Risk Estimates.................................................................................................................. 44
Table 3. Validity of LSI-R and START Total Scores in Predicting Arrests and Jail Days across Follow-up Periods........................................................................ 45
Table 4. Validity of LSI-R and START Risk Estimates in Predicting Arrests and Jail Days across Follow-up Periods........................................................................ 46
Table 5. Moderation of Total Scores and Risk Estimates by Race on Recidivism at 3-Month Follow-up ...................................................................................... 47

CHAPTER 3
Table 1. Summary of Negative Binomial Regression Results by Model of Protection .................................................................................................................. 77
Table 2. Model Fit Statistics for Hierarchical Negative Binomial Regression Analyses by Model of Protection .................................................................................. 78
Table 3. Mediation of START Vulnerability Scores and Recidivism by START Strength Scores ........................................................................................................ 79

CHAPTER 4
Table 1. Demographic Characteristics and Recidivism for MHC and TAU Participants .................................................................................................................. 107
Table 2. Group Comparisons between MHC and TAU Participants for Charges, Convictions, and Jail Days................................................................. 108
Table 3. Individual and Process Factors Predicting Changes in Jail Days among MHC Participants ......................................................................................... 109
LIST OF FIGURES

CHAPTER 2

Figure 1. Jail Days at 3-Month Follow-up by START General Offending Risk Estimate and Race.................................................................48
Figure 2. Jail Days at 3-Month Follow-up by LSI-R Risk Estimate and Race.................49
CHAPTER 1

Introduction

Adults with mental illnesses are overrepresented in U.S. jails relative to the general population. Specifically, nearly 60% of jail inmates present with symptoms of any Diagnostic Statistical Manual IV (DSM-IV) disorder (James & Glaze, 2006) compared to 26.2% of adults in the general population (Kessler, Chiu, Demler, & Walters, 2005). Justice-involved adults with mental illnesses are also more likely to recidivate post-release relative to offenders without mental illnesses (Baillargeon, Binswanger, Penn, Williams, & Murray, 2009). Adults with serious mental illnesses (i.e., presence of a major depressive, bipolar, schizophrenia spectrum, or psychotic disorder) are overrepresented in U.S. jails to an even greater extent. To demonstrate, whereas prevalence rates of serious mental illness in the general population stand at 3.2% for men and 4.9% for women (Substance Abuse and Mental Health Services Administration, 2013), up to 14.5% of male inmates and 31.0% of female inmates present with a serious mental illness at the time of intake (Steadman, Osher, Robbins, Case, & Samuels, 2009). Beyond presence of serious mental illness, other factors have been shown to increase risk of justice involvement in this population, including presence of co-occurring substance use (Baillargeon et al., 2010; Castillo & Alarid, 2011; Cox, Morschauer, Banks, & Stone, 2001), homelessness (Burns, Hiday, & Ray, 2013; Fischer, Shinn, Shrou, & Tsemberis, 2008), and prior criminal justice involvement (Bonta, Law, & Hanson, 1998).
High rates of justice involvement in this population spurred the development of mental health diversion programs to connect justice-involved adults with mental illnesses to community-based treatment as an alternative to incarceration. Although these programs are broadly characterized by diversion into treatment, they may take on many forms. For example, a survey of early diversion programs \((N = 18)\) identified programs as either pre-booking (i.e., diversion at the point of police contact) or post-booking (i.e., diversion at the point of jail booking or first court appearance) models (Steadman, Morris, & Dennis, 1995).

Successful diversion models were distinguished by multiple components, including coordinated service referral and delivery, regular meetings with criminal justice and community agency representatives, use of boundary spanners between mental health and criminal justice systems, strong leadership, early identification of participants, and provision of case management services. Mental Health Courts (MHCs), a court-based diversion program, remain the most common post-booking diversion model. These courts are characterized by several key features, including a separate docket of cases, judicial supervision of individually tailored treatment plans, regular judicial status hearings for participants, and terms of participation for successful completion or graduation (Thompson, Osher, & Tomasini-Joshi, 2008). In the past two decades, diversion programs have proliferated across the U.S., from 230 in 1995 (Steadman et al., 1995) to over 500 programs as of 2009 (Case, Steadman, Dupuis, & Morris, 2009). MHCs have grown perhaps even more rapidly, from 250 in 2009 (Almquist & Dodd, 2009) to nearly 350 as of 2015 (SAMHSA’s GAINS Center, 2015).
Despite the dramatic growth in diversion programs, the empirical literature has not consistently supported their effectiveness in improving criminal justice outcomes among justice-involved adults with mental illnesses. To demonstrate, although many single-site investigations have found robust evidence for the effectiveness of diversion programs (e.g., Kubiak, Roddy, Comartin, & Tillander, 2015; McNiel & Binder, 2007), multi-site investigations have produced variable results. In one study, researchers investigated the effectiveness of 14 post-booking jail diversion programs in decreasing recidivism for participants with mental illnesses (Case et al., 2009). Using a within-subjects repeated measures design, results showed that participants (N = 546) experienced reductions in both jail days and arrests over a 12-month pre-enrollment to 12-month post-enrollment period.

In contrast, three multi-site studies employing non-equivalent comparison group designs found limited support for the effectiveness of mental health diversion programs on recidivism. One study examined 12-month recidivism outcomes associated with participation in six pre- and post-booking diversion programs (Steadman & Naples, 2005). Although diversion participants (n = 617) experienced a 16.6% reduction in arrests relative to the 12-month period prior to diversion, non-diverted participants (n = 570) experienced a 42.1% reduction in arrests. In another multi-site study of eight pre- and post-booking criminal justice diversion programs, Broner, Lattimore, Cowell, and Schlenger (2004) compared jail days and arrests at both 3- and 12-month post-enrollment in a sample of 741 diversion program clients with mental illnesses and co-occurring substance use disorders compared to a non-diverted sample with similar behavioral health characteristics (n = 765). Instead of
decreasing recidivism, diversion was associated with a 1.19 increased likelihood of re-arrest at 3-month follow-up and 1.25 increased likelihood of re-arrest at 12-month follow-up. However, these increases were not significant and the regression coefficients for the effect of diversion on number of arrests at 3-month and 12-month follow-up were very small: .02 and .04, respectively. In a final multi-site study, Shafer, Arthur, and Franczak (2004) investigated the effectiveness of two post-booking jail diversion programs on arrests at three and 12 months following program entry for diversion program clients with co-occurring mental illnesses and substance use disorders. The authors found no significant differences in arrests between diverted clients \( (n = 154) \) and a sample of non-diverted justice-involved adults with mental illnesses and co-occurring substance use \( (n = 94) \) three or 12 months following diversion.

To date, there has been one multi-site investigation of MHCs, which employed a non-equivalent comparison group design to examine changes in arrests and jail days as a function of MHC participation across four sites (Steadman, Redlich, Callahan, Robbins, & Vesselinov, 2011). Over an 18-month pre-enrollment to 18-month post-enrollment period, MHC participants \( (n = 447) \) showed greater reductions in arrests relative to offenders experiencing traditional criminal processing \( (n = 600) \). Both groups experienced an increase in jail days over time; however, this increase was substantially lower for MHC participants.

Together, results from multi-site investigations have produced variable support for the effectiveness of diversion programs on recidivism. Moreover, few studies have investigated characteristics of participants or components of mental health diversion
programs that may explain variability in the effectiveness of such programs. In fact, only one study to date has established a potential mechanism for the effectiveness of diversion programs, finding that greater use of mental health treatment services (e.g., individual counseling, family therapy, intensive case management, etc.) was associated with larger reductions in recidivism among diversion participants (Han & Redlich, 2015).

Our limited understanding of the effectiveness of diversion programs for adults with mental illnesses reflects, in part, a lack of attention to the role of general risk factors for recidivism in this population. Briefly, mental health diversion programs were developed under the assumptions of the criminalization hypothesis, which argued that mental health symptoms themselves functioned as risk factors for recidivism and contributed to the incarceration of adults with mental illnesses (Torrey, Stieber, & Ezekiel, 1998). Indeed, there is evidence to support that treatment of mental illness is associated with reductions in recidivism (Han & Redlich, 2015; Van Dorn et al., 2011; Van Dorn, Desmarais, Petrila, Haynes, & Singh, 2013). However, recent evidence shows that justice-involved adults with mental illnesses share similar criminogenic risk factors (e.g., criminal history, procriminal attitudes, antisocial patterns, substance use) as justice-involved adults generally and that such risk factors actually may be better predictors of recidivism than mental illness alone (Skeem, Winter, Kennealy, Louden, & Tatar, 2014). Thus, the factors that contribute to criminal justice involvement in this population are likely complex and interconnected.

Two models proposed to guide offender rehabilitation may assist in clarifying our understanding of the role of criminogenic risks and needs, in addition to protective factors, in
predicting reoffending in this population. The Risk-Need-Responsivity (RNR) model of
effective offender rehabilitation was designed to guide criminal justice practitioners in
addressing the individual risks and needs of offenders while prioritizing use of finite criminal
justice resources (Andrews & Bonta, 2010; Andrews, Bonta, & Hoge, 1990). This model
proposes that criminal justice practitioners: 1) assess an offender’s risk level and target
services toward the highest-risk offenders (the risk principle); 2) identify an offender’s
criminogenic needs and address those needs in treatment and case management planning (the
need principle); and 3) deliver treatment that is responsive to an offender’s individual
background, such as culture or level of education (the responsivity principle).

In offender populations broadly, adherence to RNR principles has been shown to
improve the effectiveness of criminal justice interventions (Andrews, Zinger, et al., 1990;
Andrews & Dowden, 2006). Only one investigation to date has examined RNR adherence in
the context of a mental health diversion program. Campbell and colleagues (2015) conducted
a retrospective review of case management files to examine RNR adherence in a Canadian
MHC. In support of the risk principle, participants with higher risk classifications received
more interventions relative to low- and medium-risk participants. Additionally, although
fewer than half of assessed criminogenic needs (e.g., education and employment, leisure
activities, substance use) were addressed in case management plans, higher number of
matched needs predicted lower recidivism following MHC exit, providing support for the
needs principle.
In contrast to the primarily risk-driven approach of the RNR model, the Good Lives Model (GLM) proposes a strengths-based approach to offender rehabilitation. Originally developed for sex offender populations, this model posits that addressing risk factors is a necessary, but not sufficient, strategy to rehabilitate offenders (Ward, 2002; Ward & Stewart, 2003). Rehabilitation efforts additionally must employ strengths-based case management approaches that develop human goods (e.g., healthy living, knowledge, mastery in leisure and work, autonomy, relationships, etc.) and enhance the lives of offenders. By capitalizing on individual strengths, the GLM framework may better address issues of offender motivation and the responsivity principle of the RNR model (Ward & Fortune, 2013; Ward & Stewart, 2003). However, this model has only recently received attention as a strategy for justice-involved adults with mental illnesses (Barnao, 2013; Gannon, King, Miles, Lockerbie, & Willis, 2011) and has been criticized for its lack of empirical support (Andrews, Bonta, & Wormith, 2011).

**Purpose**

Although models of offender rehabilitation suggest risk factors, criminogenic needs, and protective factors (i.e., individual strengths) may each contribute to the risk management of offender populations, few studies to date have examined the use of risk assessments in mental health diversion settings. Specifically, we know little regarding the unique contributions of risk and protective factors in predicting recidivism outcomes in this population. Yet, knowledge of such factors may inform the delivery of targeted interventions or services in this setting and improve the effectiveness of diversion programs. Further, there
is a need for better understanding of factors (e.g., participant characteristics, program components) associated with the effectiveness of diversion programs. The manuscripts included in this dissertation address these limitations. In three separate samples of mental health diversion participants (total $N = 742$), the included manuscripts examine the psychometric properties of risk assessments, associations between risk and protective factors with respect to recidivism, and participant- and process-related factors associated with recidivism.

**Overview of Research Chapters**

The first manuscript (Chapter 2) examines the reliability and validity of two risk assessment instruments, measuring risk and protective factors, in mental health diversion participants ($N = 95$). Specifically, the inter-rater reliability and convergent validity of two instruments, the Level of Service Inventory – Revised (LSI-R; Andrews & Bonta, 2001) and Short-Term Assessment of Risk and Treatability (Webster, Martin, Brink, Nicholls, & Desmarais, 2009), are assessed at baseline. Negative binomial regression analyses examine the predictive validity of LSI-R and START total scores and risk estimates vis-à-vis arrests and jail days at 3, 6, 9, 12, and 18 months following baseline. Differences in the predictive accuracy of assessments as a function of race are also examined using hierarchical negative binomial regression models.

The second manuscript (Chapter 3) empirically evaluates five models describing hypothesized associations between risk and protective factors with respect to recidivism in a mental health diversion sample ($N = 550$). Risk and protective factors are operationalized
using START vulnerability (i.e. risk factors) and START strength (i.e., protective factors) total scores from assessments conducted at program intake. Recidivism is measured by the number of arrests and jail days one year following the date of START assessment. The incremental validity of the Compensatory, Buffer, Protective-Protective, and Challenge models is examined using hierarchical negative binomial regression models. The Mediation model is tested using conditional process analysis and the indirect effect quantified with bias-corrected bootstrapped confidence intervals.

The third manuscript (Chapter 4), a quasi-experimental evaluation of a mental health court (MHC), investigates participant- and process-related factors associated with decreased recidivism in diversion participants. First, between-group differences in charges, convictions, and jail days are examined using a propensity-score matched comparison group of offenders receiving treatment as usual ($N = 40$) relative to MHC participants ($N = 57$). Generalized estimating equations are used to examine within-subject changes in jail days as a function of participant-related factors (i.e., level of charge, graduation status, and substance use) and process-related factors (i.e., time between program referral and entry, time between program referral and first receipt of services, and time in program). All analyses employ negative binomial regression models to account for overdispersion in dependent variables.

Jointly, the included manuscripts advance the science on mental health diversion programs in three key ways. The first manuscript informs the utility and predictive accuracy of risk assessments administered in diversion settings, including the role of risk and protective factors in predicting recidivism outcomes. Instruments designed to assess and
predict risk of reoffending may contribute to the risk management of diversion clients, but little research to date has been conducted in this setting. The second manuscript clarifies the association between—and the relative importance of—risk and protective factors in predicting recidivism outcomes in this population. These associations inform which of two competing models of offender rehabilitation—the RNR approach or the GLM—may be more appropriate for the risk management of this population. The third manuscript investigates factors associated with reductions in recidivism in MHC participants, providing a framework for future evaluations to identify sources of variability in the effectiveness of MHCs.
CHAPTER 2
Reliability and Validity of START and LSI-R Assessments in Mental Health Jail Diversion Clients

Risk assessment instruments are increasingly used in mental health jail diversion programs. This study examined the reliability and validity of Short-Term Assessment of Risk and Treatability (START) and Level of Service Inventory-Revised (LSI-R) assessments overall and by client race. Research assistants completed START and LSI-R assessments for 95 diversion clients. Arrests and jail days were collected via official records and self-report 3, 6, 9, 12, and 18 months after baseline. Assessments demonstrated good inter-rater reliability and convergent validity. START strength total scores and LSI-R risk estimates were the strongest predictors of recidivism. Total scores and risk estimates did not differ as a function of client race, but there were some differences in accuracy of START vulnerability and LSI-R total scores and risk estimates in predicting jail days (but not arrests), over shorter follow-ups. No such differences were found for START strength total scores across any follow-up period or recidivism measure.

Assessment, in press.
Introduction

Developed in response to rising numbers of adults with mental illnesses in jails and high rates of reoffending within this population (Baillargeon et al., 2009; Steadman et al., 2009), over 500 mental health diversion programs exist across the United States (Case et al., 2009). Similar programs are found in other countries, such as Canada and Australia, as well (Richardson & McSherry, 2010; Slinger & Roesch, 2010). These programs are distinguished from traditional criminal case processing in that they seek to identify detainees with mental illness, use this information as a factor in case disposition, and connect these detainees to community-based treatment (Steadman, Barbera, & Dennis, 1994). Increasingly, diversion programs are being required to use instruments designed to assess a client’s risk for violence and criminal behavior, yet few studies have investigated the use of risk instruments in these settings (Skeem & Monahan, 2011).

Because risk assessment instruments do not intrinsically hold psychometric properties, validation is critical to establishing support for their use within new settings and with specific populations (AERA, 2014). A growing body of literature suggests justice-involved adults with mental illnesses are a unique population in the criminal justice system that have increased general risk factors for recidivism and behavioral health needs (Skeem et al., 2014). To date, however, very little peer-reviewed research has reported on the reliability and validity of risk assessment instruments in mental health jail diversion settings. In fact, less than a handful of studies have examined the psychometric properties of risk assessment instruments in mental health diversion programs in the United States (Barber-Rioja, Dewey,
Kopelovich, & Kucharski, 2012; Canales, Campbell, Wei, & Totten, 2014; Desmarais, Van Dorn, Telford, Petrila, & Coffey, 2012). Two instruments that have been implemented in these settings are the Short-Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Desmarais, 2009) and the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 2001).

**Short-Term Assessment of Risk and Treatability**

The START is a structured professional judgment instrument that guides assessment of risk for outcomes beyond violence alone, including suicide, self-harm, victimization, substance use, unauthorized leave, and self-neglect. As part of a larger project, the START additionally was being used to inform assessments of risk for general offending in the current study. Although still relatively new, the START has experienced quick uptake into practice. The START has been translated into nine languages and is being used in behavioral health and correctional settings, as well as private practices, in more than a dozen countries (Nicholls, Desmarais, Martin, Brink, & Webster, under review). Compared to other risk assessment tools, the START is unique in its consideration of vulnerabilities and strengths for each item. Additionally, START assessments focus on short-term risk (i.e., weeks to months) rather than long-term risk (i.e., months to years). These distinctive aspects of the START make the approach particularly relevant to community corrections settings, such as mental health jail diversion programs, where treatment planning and intervention are prioritized and the periods of supervision are shorter.
Prior research has established generally high internal consistency and inter-rater reliability of START assessments. To demonstrate, in a recent systematic review, studies \( N = 7 \) examining internal consistency of strength and vulnerability total scores reported Cronbach alpha coefficients of .80 to .95 and .76 to .95, respectively (O’Shea & Dickens, 2014). Findings on inter-rater reliability, when reported, have been similarly high. In the same review, intra-class correlations were found to be high for strength total scores (.78), vulnerability total scores (.86), and risk estimates (.82). Prior research also provides strong evidence of the validity of START assessments in predicting short- and medium-term aggression and violence among civil and forensic psychiatric inpatients (Braithwaite, Charette, Crocker, & Reyes, 2010; Chu, Thomas, Ogloff, & Daffern, 2011a, 2011b; Desmarais, Nicholls, Wilson, & Brink, 2012; O’Shea & Dickens, 2014; Wilson, Desmarais, Nicholls, & Brink, 2010). Moreover, studies have indicated strong predictive validity of START assessments among other vulnerable populations, such as among offenders with an intellectual disability (Inett, Wright, Roberts, & Sheeran, 2014).

Desmarais, Van Dorn, and colleagues (2012) recently investigated characteristics of START assessments completed on 96 mental health jail diversion clients in the context of routine practice. The authors found preliminary support for the use of START assessments in diversion settings; for example, findings revealed good internal consistency of START assessments \( (\alpha = .95 \text{ for strength total scores}; \alpha = 0.90 \text{ for vulnerability total scores}) \) and characteristics of START assessments, such as mean vulnerability total scores, were similar to those conducted in other settings (e.g., forensic hospitals). Moreover, findings suggested
the need for investigation into an eighth risk domain to capture general offending risk in this population. However, to date, no study has examined the predictive validity of START assessments completed on mental health jail diversion program clients or in predicting general, non-violent offending in any population. Further, few studies to date have examined the psychometric properties of START assessments in correctional samples, especially those in community settings, or among racial minorities more broadly.

**Level of Service Inventory-Revised**

The LSI-R is an actuarial risk assessment instrument part of a larger family of Level of Service instruments developed from the Risk-Need-Responsivity model of effective offender rehabilitation and designed to estimate general offending risk (Andrews & Bonta, 2010; Andrews, Bonta, et al., 1990; Andrews, Bonta, & Wormith, 2010). These instruments additionally include the Level of Service/Case Management Inventory (LS/CMI; Andrews, Bonta, & Wormith, 2004), the Level of Service/Risk-Need-Responsivity (LS/RNR; Andrews, Bonta, & Wormith, 2008), and Level of Service-Inventory-Revised: Screening Version (LSI-R: SV; Andrews & Bonta, 1998). Although more recently published, the LS/RNR and LS/CMI are more suitable for correctional practice and case management due to their focus on specific risks, needs, and responsivity considerations. The LSI-R, in contrast, is more readily employed in research. Indeed, whereas few studies have examined the predictive validity of LS/CMI or LS/RNR assessments in correctional settings, at least 25 published studies have reported on the psychometric properties of LSI-R assessments vis-à-
vis recidivism in correctional samples (Desmarais, Johnson, & Singh, 2016), providing context and a point of reference for validation of the LSI-R in unique settings and samples.

The LSI-R includes 54 items measuring both static and dynamic risk factors across 10 separate risk domains (e.g., Financial, Education/Employment, Criminal History, etc.) (Andrews & Bonta, 2001). Total scores on the instrument classify offenders into five distinct risk bins: Low, Low/Moderate, Moderate, Moderate/High, and High. In research and practice, these have been collapsed further into risk bins of Low, Moderate, and High (Canales et al., 2014; Singh, Grann, & Fazel, 2011). The LSI-R is one of the most widely used and well-known risk assessment instruments; to demonstrate, a recent survey by Multi-Health Systems reported that over 900 criminal justice agencies regularly use the LSI-R in practice (Lowenkamp, Lovins, & Latessa, 2009).

Despite its widespread use, surprisingly, limited published research has reported on the reliability of LSI-R assessments. Initial, unpublished studies, including initial validation research conducted by the instrument’s primary author, reported high levels of inter-rater reliability ($r$ range = .87-.92) for LSI-R assessments conducted in correctional samples (for a review, see Andrews et al., 2010). However, in more recent investigations, when reported in published research, levels of inter-rater reliability and internal consistency have been highly variable. For example, Lowenkamp and colleagues (2004) found LSI-R subsections to have varying rates of agreement (61.5% to 97.7%) among correctional practitioners. Furthermore, levels of internal consistency also have been found to be less reliable among racial and ethnic minorities (Schlager & Simourd, 2007), who make up about half of mental health jail
diversion clients (e.g., Case et al., 2009). In contrast, a substantial body of research (i.e., over 100 studies) supports the validity of LSI-R assessments in predicting general recidivism overall (Olver, Stockdale, & Wormith, 2014; Singh et al., 2011; Vose, Cullen, & Smith, 2008), as well as across various subgroups of offenders, including male offenders, female offenders, and ethnic minorities (Olver et al., 2014). However, in some studies, LSI-R assessments have demonstrated poorer predictive validity among racial minorities (Chenane, Brennan, Steiner, & Ellison, 2015; Schlager & Simourd, 2007; Whiteacre, 2006). There is also relatively limited research on the predictive validity of LSI-R assessments among adults with mental illnesses (Ferguson, Ogloff, & Thomson, 2009).

To our knowledge, only one study has examined the validity of LSI-R assessments in the context of mental health jail diversion programs. Specifically, Canales, Campbell, Wei, and Totten (2014) investigated the predictive validity of Level of Service/Risk-Needs-Responsivity (LS/RNR; Andrews, Bonta, & Wormith, 2008) assessments conducted in a Canadian mental health court. Results showed that assessments of General Risk/Need produced strong levels of predictive validity for general recidivism, particularly among male clients. However, this study was limited in its investigation of the LS/RNR scales and not the more frequently used LSI-R. To our knowledge, no research to date has examined the predictive validity or inter-rater reliability of LSI-R assessments completed on mental health jail diversion clients in the United States.
Racial Bias in Risk Assessment

In the context of risk assessment, there are growing concerns about the potential for racial bias in both the predictive accuracy and utilization of these instruments in correctional practice. Concerns have been raised by notable figures such as U.S. Attorney General Eric Holder, who has argued that by focusing on static factors like socioeconomic status or criminal history, risk assessments may obstruct the administration of justice by biasing minority offenders toward higher-risk classifications (Holder, 2014; Starr, 2014). Although race is not measured overtly in modern risk assessments, risk factors may serve as proxies for race in risk assessments (Harcourt, 2015) due to systematic racial disparities in wealth accumulation (Proctor, Semega, & Kollar, 2016; U.S. Census Bureau, 2014; Vornovitsky, Gottschalk, & Smith, 2011), homeownership (Callis & Kresin, 2016), and educational attainment (U.S. Census Bureau, 2016). Furthermore, racial minorities are more likely to be stopped by police (Gelman, Fagan, & Kiss, 2007) and may receive harsher criminal sentences (Sweeney & Haney, 1992) relative to their non-minority counterparts, contributing to lengthier and more substantial criminal histories. The inclusion of these factors in risk assessment suggests that racial minorities will receive higher-risk scores and classifications relative to their rate of reoffending, resulting in differences in the predictive accuracy of risk assessments as a function of race.

Consequently, a growing number of research studies have investigated racial differences in the predictive accuracy of risk assessments in adults (e.g., Lowenkamp, Holsinger, & Cohen, 2015; Varela, Boccaccini, Murrie, Caperton, & Gonzalez, 2013). To
date, no studies have investigated racial differences in the predictive validity of START assessments. However, several studies have investigated racial differences in the predictive validity of LSI-R assessments in offenders. For example, one study found that LSI-R assessments produced weaker predictive validity estimates for institutional misconduct risk among non-White prison inmates compared to White inmates (Chenane et al., 2015). In another study, LSI-R assessments of re-arrest and parole revocation risk produced weaker predictive validity for Black prisoners upon release relative to White prisoners (Ostermann & Salerno, 2016). In a final study, LSI-R assessments produced overclassification errors for Black offenders when a lower cutoff score for risk classification was employed (Whiteacre, 2006). Together, these findings suggest the potential for racial bias in the ability for a risk assessment to accurately predict likelihood of reoffending. However, no study to date has examined racial bias in the predictive validity of START assessments, and to our knowledge, no study to date has compared the potential for racial bias in predictive validity between structured professional judgment and actuarial risk assessment instruments.

The Present Study

There is limited data on the reliability and validity of risk assessments, and START and LSI-R assessments specifically, in mental health jail diversion programs. Furthermore, despite the national dialogue on race and risk assessment, relatively few studies have examined the predictive validity of risk assessments as a function of client race, in this setting or otherwise (Holder, 2014; Skeem & Lowenkamp, 2015). To those ends, the present study examined the reliability and validity of START and LSI-R assessments completed on
mental health jail diversion program clients, piloting a START general offending risk estimate for the first time. Our study aims were to investigate: 1) inter-rater reliability of START and LSI-R assessments; 2) convergent validity between START and LSI-R assessments; 3) validity of START total scores, START general offending risk estimates, LSI-R total scores, and LSI-R final risk classification in predicting recidivism across five follow-up periods; and 4) differences, if any, in START and LSI-R assessments and their predictive validity by client race.

**Method**

**Participants**

The sample includes 95 clients (77 male; 18 female) participating in mental health jail diversion programs in a large metropolitan county in the southern United States. Participants were drawn from post-booking jail diversion programs targeting misdemeanor- and felony-level offenders with mental illnesses. Through the diversion programs, participants received case management services for up to a 1-year period, which included referral to and coordination of community-based treatment and housing services. However, there were no ongoing status hearings for participants; revocation of diversion status occurred only in the case of a new offense. At baseline, participants were an average age of 36.05 (SD = 12.46) years old. The sample comprised a mix of felony (47.4%) and misdemeanor (52.6%) offenders. Just under half (45.3%) were African American, while slightly more than half (53.7%) were Caucasian; one client identified as Asian. More than half of clients (54.7%) identified their ethnicity as Hispanic or Latino and less than half identified as non-Hispanic
or Latino (44.2%); data were unavailable for one client. Participants had extensive criminal histories, with an average of 6.69 (SD = 5.42, range = 1-31) jail bookings in the three years prior to baseline and 16.07 (SD = 14.99, range = 3-76) lifetime bookings. More than half of participants (53.7%) reported graduating from high school or receiving a GED. Consistent with study inclusion criteria, primary diagnoses at baseline included schizophrenia (35.8%), psychosis NOS (25.3%), and bipolar disorder (20.0%). Nearly one-third of the sample (30.5%) had a co-occurring substance use diagnosis at baseline. Diagnostic information was pulled from clinical records maintained by the jail diversion program.

**Procedures**

The Institutional Review Board of the University approved and monitored this study for human participant protections. Participants provided consent to participate in this study prior to each interview. As part of this process, interviewers assessed a client’s comprehension of the study purpose and procedures and also queried external pressures (i.e., coercive influences) to participate. Participants who were judged by the interviewer to have poor comprehension of the study purpose and procedures or who experienced undue external pressures to participate would not be enrolled in the study nor would they be interviewed at follow-up.

**Data collection.** Semi-structured interviews were conducted by one of four research assistants with clients at baseline and at five follow-up periods (3-month, 6-month, 9-month, 12-month, and 18-month). At baseline, research assistants completed the START and LSI-R assessments based on participant self-report, observation, and review of official records. At
each follow-up, participants self-reported arrests and days incarcerated in the previous three months. These self-reported recidivism data were used to supplement official county-wide arrest and incarceration records, which were queried by the diversion program coordinator in the County’s jail records database and provided in raw form to the research team.

**Inter-rater reliability.** Research assistants with Bachelor’s degrees attended a two-day workshop prior to the study start that including training on the START and LSI-R. The workshops and subsequent training included completion of practice cases \( n = 5 \) upon which the research assistants were required to achieve adequate agreement with the trainers. Inter-rater reliability was conducted on a subset of 25% of the total sample \( n = 24 \). Inter-rater reliability procedures were identical to the interview procedures. In all instances, an interview was conducted with a client, and both research assistants attended the interview and coded the assessment independently.

**Measures**

**START.** The START is an SPJ guide for assessing risk of seven adverse outcomes: violence to others, suicide, self-harm, victimization, substance use, unauthorized leave, and self-neglect (Webster et al., 2009). In the present study, the START additionally was used to inform an eighth estimate of risk for general offending (Desmarais, Van Dorn, et al., 2012). For each of the 20 items, strength and vulnerability demonstrated in the past two to three months are coded on a 3-point ordinal scale, from 0 (*minimally present*) to 2 (*maximally present*). Strength and vulnerability are rated independent of one another; thus, an individual may receive the same high (or low) score for both strength *and* vulnerability on any given
item. Additionally, assessors may indicate key and critical items to highlight those items of particular importance to the individual. Assessors subsequently estimate risk extending to the next three months as low, moderate, or high for each of the eight outcomes. These specific risk estimates are formed through structured professional judgment which considers strength and vulnerability ratings, the presence of key and/or critical items, physical health problems, and other historical factors (e.g., past violence or victimization). Strength and vulnerability total scores were calculated by summing the item ratings, with a possible range of 0 to 40. When there were five or fewer missing item ratings, we prorated total scores by summing the item ratings, dividing by the total possible scale score (i.e., 40), multiplying that amount by the number of omitted items, and adding the obtained value to the original total (see Webster et al., 2009, p. 34). Assessments missing more than five item ratings were not prorated, leading to the exclusion of two START assessments from our analytic sample. Levels of internal consistency were good for START strength total scores (Cronbach’s α = .87) and START vulnerability scores (Cronbach’s α = .81).

**LSI-R.** The LSI-R is an actuarial instrument designed to assess risk of criminal offending and technical violations in adult offenders (Andrews & Bonta, 2001). We selected this version of the Level of Service assessments due to its widespread use in research (Desmarais et al., 2016), which allowed us to investigate the predictive validity of assessments in relation to the extant literature. Additionally, because we selected this instrument for use in a research study, the case management and specific needs and
responsivity components of the LS/CMI (Andrews et al., 2004) and LS/RNR (Andrews et al., 2008) assessments were not relevant to our investigation.

The LSI-R includes 54 items representing both static and dynamic risk factors for recidivism. Out of the 54 items, 13 items measure the severity of specific risk factors and are coded from 0 indicating a very unsatisfactory situation with a very clear and strong need for improvement (most severe) to 3 indicating a very satisfactory situation with no need for improvement (least severe). For these items, scores were recoded as 0 or 1 to indicate either a generally satisfactory situation (i.e., absence of the risk factor, 0) or a generally unsatisfactory situation (i.e., presence of the risk factor, 1). All other items measure the absence or presence of a risk factor and are coded as 0 or 1. Scores from the 54 items were totaled to create the LSI-R total score. Predetermined cut-offs were used to classify offenders into risk categories: Low = 0-13; Low-Moderate = 14-23; Moderate = 24-33; Moderate-High = 34-40; High = 41-54. Due to low cell counts, these five risk bins were collapsed into three categories following the procedure used in prior research (Singh et al., 2011); specifically, Low and Low-Moderate were coded as Low, Moderate was coded as Moderate, and Moderate-High and High were coded as High. This strategy also afforded better comparison with the three categories of the START General Offending risk estimate. For all predictive validity analyses involving risk estimates, High risk was used as the reference group. Insufficient data were available to code an LSI-R assessment for one participant. Levels of internal consistency were acceptable for LSI-R total scores (Cronbach’s α = .72).
**Race.** Race was coded categorically and measured independently from ethnicity. All participants except one identified as either Caucasian or African American. As a result, one participant identifying as Asian was excluded from race analyses.

**Recidivism.** Recidivism measures were coded from raw booking data provided by jail diversion staff and included count measures of number of arrests and days incarcerated at each follow-up period. Although days incarcerated is not frequently employed as a measure of general recidivism in risk assessment validation research, this measure is meaningful in the context of jail diversion for several reasons. Primarily, jail diversion programs were created to decrease time incarcerated among justice-involved adults with mental illnesses because jails provide limited access to mental health care (Wilper et al., 2009) and may exacerbate symptoms of mental illness (Steadman et al., 1994; Torrey et al., 1998). Accordingly, days incarcerated is a common and accepted indicator of recidivism in jail diversion research (Case et al., 2009; Sirotich, 2009; Steadman & Naples, 2005; Steadman et al., 2011). Additionally, days incarcerated may be an appropriate measure of reoffending severity for groups that are at particularly high-risk for reincarceration because it captures a broader range of severity in reoffending (Urban Institute, 2016). As detailed earlier, our study involves a sample of justice-involved adults with mental illnesses with substantial criminal histories (i.e., an average of 16.07 lifetime jail bookings).

To code recidivism variables, each unique booking date was treated as a separate arrest and counts of jail days were tabulated based on date of booking and date of release from jail. Additionally, participants self-reported number of arrests and days incarcerated
during the follow-up periods. Because attrition rates were relatively high for self-reported data (i.e., 30% at 3-month follow-up to 81.0% at 18-month follow-up), official records were used as the primary source of recidivism data. However, when available, we incorporated self-reported data from follow-up interviews by recording client self-report estimates when a client overestimated recidivism in comparison to official records. Across follow-up periods, agreement between self-reported recidivism and official records was high for both arrests (86.8%, range: 83.3% to 91.2%) and days incarcerated (85.7%, range: 81.2% to 88.1%). Official recidivism data were unavailable for one participant who had START and LSI-R assessments.

**Analyses**

First, descriptive statistics were calculated on all study variables. Second, inter-rater reliability of START Strength and Vulnerability total scores and LSI-R total scores were evaluated using two-way mixed single effects Intraclass Correlation Coefficients (ICC$_{3,1}$) which estimates reliability of a single rating using a fixed effects model (McGraw & Wong, 1996; Shrout & Fleiss, 1979). For intraclass correlation coefficients (ICC), values less than .40 indicate slight, between .40 and .59 fair, between .60 and .74 good, and greater than .75 excellent agreement (Cicchetti et al., 2006). Third, convergent validity was assessed via bivariate correlations between LSI-R and START total scores and via Cohen’s Kappa percentage agreement between LSI-R and START risk estimates. Based on Cohen’s criteria, correlation values of .10 indicate a small, .30 a medium, and .50 a large effect size (Cohen, 1988). Kappa values of less than .20 reveal slight, between .21 and .40 fair, between .41 and
.60 moderate, between .61 and .80 substantial, and greater than .81 almost perfect agreement (Landis & Koch, 1977). Fourth, a series of bivariate negative binomial regression analyses were conducted between total scores and recidivism at each follow-up. Because criminal justice records represent count data, generalized linear approaches are necessary to account for positive skew (i.e., excess of zeros) and absence of a normal distribution in the dependent variables (Walters, 2007). Data in this study failed to meet the variance assumptions of Poisson distributions (i.e., the variance is equivalent to the mean), thus, resulting in the negative binomial regression approach. In negative binomial models, the Wald chi-square statistic provides a significance test of the hypothesis that the regression coefficient is different from zero, controlling for other model terms. The incidence rate ratio (IRR) provides a measure of effect size and represents a factor of the dependent variable associated with a 1-unit increase in the independent variable. Additional negative binomial regression analyses were conducted using dummy-coded risk estimates (with high risk as the reference group) and recidivism measures. Fifth and finally, hierarchical negative binomial models were conducted to test for interaction effects of race by total scores and risk estimates on recidivism at 12-month follow-up. In block one, the total score or risk estimate under investigation was entered together with race. In block two, a risk by race effect was added. Improvement in model fit in block two was determined by change in -2 Log(Likelihood) statistics, which were evaluated by computing a p value for a chi-square distribution with 1-2 degrees of freedom.
Results

Descriptive Statistics

Total scores. The mean START vulnerability total scores was 19.13 (SD = 6.30, range = 2-30), and the mean START strength total scores was 14.78 (SD = 7.31, 0-29), suggesting that participants presented with greater vulnerabilities than strengths at baseline. Mean LSI-R total scores were 28.08 (SD = 6.08, range = 15-41), suggesting that participants presented with a moderate amount of recidivism risk at baseline. Between-group comparisons showed no significant differences between African American and Caucasian participants on START strength total scores (t [90] = 1.08, p = .282), START vulnerability total scores (t [90] = -0.85, p = .397), or LSI-R total scores (t [91] = -0.57, p = .571).

Risk estimates. Using START general offending risk estimates, less than a quarter of participants were rated as low risk (24.7%, n = 23), 53.8% as moderate risk (53.8%, n = 50), and 21.5% at high risk (n = 20) for recidivism. Using the three LSI-R risk bins, less than a quarter of participants were categorized at low risk (20.2%, n = 19), 61.7% as moderate risk, (n = 58), and 18.1% as high risk (n = 17) risk. Similar to total scores, between-group comparisons showed no significant differences between African-American and Caucasian participants on START general offending risk estimates ($\chi^2[2] = 1.53, p = .465$), or LSI-R risk estimates ($\chi^2[2] = 1.13, p = .569$).

Recidivism. Full descriptive statistics for recidivism across all follow-up periods are presented in Table 1. Overall, the average number of arrests for the sample was generally
low: less than two arrests per person over the 18-month follow-up period. However, there
was considerable variability across participants; for example, the number of arrests ranged
from 0 to 15. Findings were similar for the number of jail days across follow-up periods (see
Table 1). Further, each arrest was associated with an average of 15.35 jail days.

Reliability

To address Aim 1, inter-rater reliability was calculated on a subset of 24 participants,
using interclass correlation coefficients (ICC) for START strength and vulnerability total
scores and LSI-R total scores, and Cohen’s Kappa (and percentage agreement) for START
and LSI-R risk estimates. Inter-rater reliability was good for START strength (ICC$_{3,1}$ = .61,
95% CI = .29-.81, $p = .001$) and vulnerability total scores (ICC$_{3,1}$ = .63, 95% CI = .31-.82, $p$
< .001), and was excellent for LSI-R total scores (ICC$_{3,1}$ = .76, 95% CI = .53-.89, $p < .001$)
(Cicchetti et al., 2006). Inter-rater reliability was lower for START general offending risk
estimates (58.3% agreement, $p = .209$) and adequate for LSI-R risk bins (66.7%, $p = .017$)
(Landis & Koch, 1977).

Validity

Convergent validity. To address Aim 2, we investigated the convergent validity
between LSI-R and START assessments. Results indicated a significant, positive association
between START vulnerability and LSI-R total scores, $r(91) = .37$, $p < .001$, providing
evidence for convergent validity. Similarly, START strength scores were negatively and
significantly associated with LSI-R total scores, $r(91) = -.37$, $p < .001$. Cross-tabulation of
START risk estimates and LSI-R risk bins are presented in Table 2. Results showed weak agreement between START general offending risk estimates and LSI-R risk bins, $\kappa = .04$ ($SE = .08$), $p = .623$.

**Predictive validity.** To address Aim 3, we investigated the predictive validity of START and LSI-R assessments using negative binomial regression analyses, the results of which are provided in Table 3. START strength total scores significantly and inversely predicted number of arrests (IRR range: 0.88-0.95) and jail days (IRR range: 0.83-0.96) across all follow-up periods. The strongest effect was observed for the prediction of jail days at 6-month follow-up (IRR = 0.83); specifically, each 1-point increase in START strength scores was associated with 1.20 times fewer jail days. START vulnerability total scores significantly and positively predicted arrests (IRR range: 1.04-1.06) and jail days (IRR range: 1.03-1.07) over follow-up periods of six months and longer for both. START vulnerability total scores were most robust predictors of jail days at 6- and 9-month follow-up (IRRs = 1.07). Each 1-point increase in vulnerability scores was associated with 1.07 times greater jail days served. LSI-R total scores demonstrated predictive validity over most follow-up periods, with the exception of the 18-month follow-up (see Table 3). LSI-R total scores were more robust predictors of jail days (IRR range: 1.06-1.13) than arrests (IRR range: 1.05-1.06) (see columns 3 and 6 in Table 3). LSI-R total scores produced the strongest predictive validity estimates for jail days at 6-month follow-up (IRR = 1.13); each 1-point increase in LSI-R total scores was associated with 1.13 times greater jail days served.
Table 4 presents the results of analyses testing the predictive validity of START general offending risk estimates and LSI-R risk bins. Overall, START risk estimates showed relatively limited validity in predicting number of arrests across follow-up periods ($ps \geq .025$). In contrast, START risk estimates significantly predicted number of jail days from 3-month to 12-month follow-up periods. Estimates of low versus high risk showed greater discrimination (IRR range: 0.11-0.19) than found between estimates of moderate versus high risk (IRR range: 0.48-0.52), as would be expected. LSI-R risk bins similarly showed greater validity in predicting jail days than arrests, though effect sizes were generally stronger relative to START risk estimates. With respect to arrests, we found discrimination between participants in the low vs high risk bins (IRR range: 0.24-0.33), but not the moderate vs high risk bins ($ps \geq .120$). For jail days, however, there was discrimination between both low vs high risk (IRR range: 0.07-0.31) and moderate vs high risk (IRR range: 0.39-0.48) participants across most follow-up periods (see Table 4). Effect sizes were strongest at the 6-month period with respect to jail days; participants rated at high risk had 2.56 and 14.29 times more jail days relative to those rated at moderate and high risk, respectively.

To address Aim 4, we examined whether client race moderated the predictive validity of LSI-R and START assessments. Although we conducted these analyses across all five follow-up periods, we found the most significant assessment by race interaction effects at 3-month follow-up. In contrast, there were few significant interaction effects at the 6-, 9-, and 12-month follow-up periods and none at the 18-month follow-up periods. Of note, race did
not moderate predictive validity of START strength total scores across any follow-up period or recidivism measure ($p_s > .071$). (Full results not presented, but are available upon request.)

Overall, there were no significant moderating effects of race on validity in predicting arrests at 3-month follow-up for any of the assessment measures ($p_s > .186$). However, we did find some moderating effects of race on validity in predicting jail days at 3-month follow-up. Specifically, there was a significant START vulnerability total scores by race interaction ($p < .001$), such that a 1-point gain in vulnerability total scores was associated with 1.18 times more jail days among Caucasian participants relative to African American participants. LSI-R total scores similarly were moderated by race ($p = .040$), such that 1-point increase in LSI-R total scores was associated with 1.10 times more jail days among Caucasian participants relative to African American participants. Results also showed significant effects of START general offending and LSI-R risk estimates by race on jail days at 3-month follow-up. For the START risk estimate, Caucasian participants were incarcerated for more days relative to African American participants in both the low and high risk categories; however, this effect was greater among participants classified at low risk (see Figure 1). For the LSI-R risk estimate, at the moderate risk level, African American participants had more jail days relative to Caucasian participants; however, at the high risk level, Caucasian participants had more jail days relative to African American participants (see Figure 2). This interaction was also observed for START general offending risk estimates, but only trended toward significance ($p = .051$).
Discussion

Structured risk assessments are increasingly used in mental health jail diversion programs across the U.S. and internationally to facilitate community-based treatment and case management of justice-involved adults with mental illnesses. However, few studies have evaluated the reliability and validity of such instruments used with mental health jail diversion clients. Thus, the goal of the present study was to examine the START and LSI-R with regard to inter-rater reliability, convergent validity, and predictive validity of their total scores and risk estimates over multiple follow-up periods. We additionally explored differences in START and LSI-R assessments, as well as their predictive validity, as a function of client race. Overall, findings support the use of these instruments to inform risk assessment and management of mental health jail diversion clients and provide limited evidence of racial bias. Below, we discuss the study findings in further detail.

Integration of Findings

Our first aim was to investigate the inter-rater reliability of START and LSI-R assessments completed on mental health jail diversion clients. Results showed good to excellent levels of inter-rater reliability for START vulnerability, START strength, and LSI-R total scores. To our knowledge, only one published study conducted in a correctional sample has investigated the inter-rater reliability of LSI-R total scores, reporting slightly lower levels of inter-rater reliability compared to the present findings (Rocque & Plummer-Beale, 2014). For START assessments, these findings are consistent with levels of inter-rater reliability reported across seven studies with institutionalized patients (O’Shea & Dickens,
2014), but provide the first evidence supporting inter-rater reliability of the START strength and vulnerability total scores completed in a community corrections setting. Levels of inter-rater reliability were adequate for LSI-R risk estimates, but lower for START general offending risk estimates. These findings suggest the need for more training to consensus on the START general offending risk estimate prior to implementation in other criminal justice settings and echo clinicians’ reports that the final risk estimates are one of the most challenging aspects of the structured professional judgment approach, and START assessments, specifically (Desmarais, 2009; Michael Doyle, Lewis, & Brisbane, 2008). These findings also likely reflect the pilot nature of the START general offending risk estimate and the need for more detailed coding instructions.

A recent meta-analytic review of validation research conducted in the United States identified only two studies that reported on the inter-rater reliability of risk assessments completed in correctional settings (Desmarais et al., 2016). Though the current results are promising, further research is needed to establish the reliability of START and LSI-R assessments, as well as risk assessments more generally, in mental health jail diversion and other U.S. correctional settings.

Our second aim was to examine convergent validity between START and LSI-R assessments. Results showed strong associations between START total scores and LSI-R scores, suggesting similar constructions and operation of factors associated with recidivism risk. In contrast, there was limited convergent validity between START general offending and LSI-R risk estimates as demonstrated by poor agreement between ratings of low,
moderate, and high risk. These discrepancies may reflect, at least in part, inherent differences in the assessment time frames of these two instruments. Specifically, the START risk estimates are designed to forecast outcomes over a 90-day period (Webster et al., 2009), whereas the LSI-R risk estimates forecast risk over several months to years (Fass, Heilbrun, DeMatteo, & Fretz, 2008; Kelly & Welsh, 2008; Manchak, Skeem, & Douglas, 2008). Indeed, in the present study, the LSI-R assessments classified more participants as moderate risk relative to the START assessments, whereas START assessments classified more participants as low risk than did LSI-R assessments.

Taken together, convergent validity findings suggest there is moderate overlap between the instruments, even though the START was not originally designed to estimate risk for general offending. Moreover, these findings support the START’s comprehensive approach to risk assessment and management that focuses on shared predictors of multiple adverse outcomes (Webster, Nicholls, Martin, Desmarais, & Brink, 2006). Findings are also consistent with recent assertions that risk assessment instruments are more or less comparable to one another and, thus, that tool selection should be informed by other issues, such as the purpose of the evaluation (i.e., classification, treatment planning, etc.), the population of interest, and resources available (Desmarais et al., 2016; Skeem & Monahan, 2011).

Our third aim was to evaluate the validity of START and LSI-R assessments in predicting general offending across multiple time frames. Overall, findings support for the predictive validity of START and LSI-R assessments across both shorter and longer follow-
up periods. Interestingly, although START is designed to predict outcomes within three months of assessment, START vulnerability total scores predicted recidivism across all follow-up periods, except three months. In previous studies, START assessments have demonstrated validity in predicting institutional aggression in psychiatric inpatients over 1-month to 12-month follow-up periods (Chu et al., 2011b; Desmarais, Nicholls, et al., 2012). Because this is the first study to investigate the validity of START assessments in predicting general offending, further research is needed to explore the time frame over which START vulnerability total scores predict recidivism in general and in mental health diversion clients specifically.

Similarly, the predictive validity of the LSI-R is typically tested with respect to longer follow-up periods (i.e., 12 or more months; Fass et al., 2008; Kelly & Welsh, 2008; Manchak et al., 2008); however, LSI-R total scores were most predictive at 3- and 6-month follow-ups. In fact, LSI-R total scores did not significantly predict recidivism over the 18-month follow-up, though LSI-R total scores have predicted recidivism as many as 3.79 years following assessment in other research (Vose, Smith, & Cullen, 2013). Discrepancies between prior and current findings may be attributable to the unique nature of the current sample (i.e., mental health jail diversion clients), but also may reflect measurement and analytic issues. In particular, we operationalized recidivism as number of arrests and jail days, and not reconviction (e.g., Manchak et al., 2008). Additionally, we used negative binomial regression models to analyze count recidivism data whereas prior studies typically have relied on analytic strategies, such as Receiver Operating Curve analyses, that require researchers to
dichotomize recidivism measures (Fass et al., 2008; Manchak et al., 2008). Our strategy is appropriate given the skewed distribution of recidivism data (Walters, 2007) and allowed us to more accurately account for the frequency of recidivism across participants. Nonetheless, as with the START vulnerability total scores, findings suggest the need for more research examining the time frame during which LSI-R total scores predict recidivism in mental health jail diversion clients.

START strength total scores emerged as the most consistent and robust predictor of arrests and jail days across follow-up periods. These results are consistent with prior research demonstrating the validity of START strength total scores in predicting institutional aggression in forensic psychiatric patients (Desmarais, Nicholls, et al., 2012) and suggest that protective factors may be especially relevant to the assessment and management of recidivism risk among community-based, justice-involved adults with mental illnesses. Consideration of protective factors may help differentiate among individuals in this population who present with many similar risk factors (e.g., chronic homelessness, substance abuse problems, mental health problems) and may guide risk management strategies by acting as the foundation or cornerstone for treatment (de Ruiter & Nicholls, 2011). Indeed, consideration of protective factors have demonstrated validity in predicting adverse outcomes, such as violence, above and beyond risk factors (Desmarais, Nicholls, et al., 2012; de Vries Robbé, de Vogel, & Douglas, 2013; Lodewijks, Ruiter, & Doreleijers, 2010). However, questions remain regarding how risk and protective factors are related both to each other and to risk for adverse outcomes.
Although the need for a START general offending risk estimate has been articulated in practice and in previous research (Desmarais, Van Dorn, et al., 2012), this is the first study to provide evidence supporting its predictive validity. Specifically, START general offending risk estimates demonstrated validity in predicting jail days across all follow-up periods, but were somewhat less accurate in predicting arrests. LSI-R risk estimates similarly demonstrated greater validity in predicting jail days compared to arrests, but nonetheless were able to discriminate between participants at low compared to high risk of arrests across all but the 18-month follow-up period. Overall, both START and LSI-R risk estimates produced more robust effect sizes when significant compared to their total scores, suggesting they were more discriminating than 1-point gains in total scores, which is consistent with prior research demonstrating the superiority of risk estimates over total scores (Desmarais et al., 2016; Desmarais, Nicholls, et al., 2012; Douglas, Ogloff, & Hart, 2003; Douglas, Yeomans, & Boer, 2005). Though findings of the current study are encouraging, further research is needed to establish the validity of a START general offending risk estimate.

Finally, our fourth aim was to explore differences in LSI-R and START assessments, as well as their predictive validity, as a function of client race. Our findings showed no differences between Caucasian and African American mental health jail diversion clients in the total scores or risk estimates for either the LSI-R or START. However, we did find some evidence of differences in the predictive validity of START and LSI-R assessments for jail days, mainly over a shorter follow-up period (i.e., 3-month). Specifically, both LSI-R and START general offending risk estimates underclassified African American clients for
moderate risk and overclassified African American clients for high risk relative to their time incarcerated over a 3-month follow-up period. Similarly, both START vulnerability total scores and LSI-R total scores demonstrated weaker predictive validity for African American clients relative to Caucasian clients. These findings are consistent with previous studies that similarly found LSI-R assessments overclassified African American offenders in a community corrections sample (Whiteacre, 2006), and demonstrated racial biases in the prediction of institutional misconduct (Chenane et al., 2015) and re-arrest or parole revocation (Ostermann & Salerno, 2016) in prison inmates. In contrast, we found no moderating effects of race on the validity of START and LSI-R assessments in predicting arrests. Additionally, there was no evidence of racial bias in the predictive validity of START strength total scores across any follow-up period or recidivism measure. Because this is the first study to explore racial differences in the predictive validity of START assessments; future research is needed to replicate these findings.

Although we investigated racial bias across all follow-up periods, evidence of disparity in the predictive accuracy of START and LSI-R assessments emerged primarily during 3-month follow-up period. This may reflect that participants were recently enrolled diversion program clients who were referred to community-based treatment and supportive services. As a result, risk and need factors contributing to higher risk classifications among African American participants upon admission to the diversion program may have been effectively targeted throughout their participation, resulting in lower risk ratings. However, this is only speculation, and whether or not diversion programs can effectively target risks
and needs has received little attention in the existing literature (Campbell et al., 2015) and is an important direction for future research (Skeem et al., 2014).

Together, these findings add to the national dialogue on whether application of risk assessment instruments may result in discriminatory practices and contribute to harsher sentences for offenders of minority racial status (Harcourt, 2015; Holder, 2014; Starr, 2014). The lack of differences in the assessment scores and risk estimates, as well as in their validity in predicting arrests, provides little evidence of racial bias in the instruments themselves. Whether or not application of risk assessment instruments in sentencing and corrections results in racially biased decisions regarding level or length of supervision remains to be seen. There is evidence to suggest that White individuals are charged with lesser crimes and receive shorter sentences than non-White individuals, even though they are arrested as frequently and, potentially, for the same behavior(s) (Mauer, 2006; Pettit & Western, 2004). However, our findings suggest that protective factors are less susceptible to racial biases and more accurate than risk factors in predicting recidivism across groups of racially diverse offenders. Indeed, protective factors by definition do not include variables that have been cited as acting as proxies for race in risk assessment, such as history of criminal behavior (Harcourt, 2015). Thus, the consideration of protective factors during the process of assessing risk for re-offending may help reduce (rather than increase) discriminatory criminal justice practices, including decisions regarding supervision, treatment referral, and case management (Grove & Meehl, 1996).
Limitations and Future Directions

Our findings should be considered in light of limitations to the study design, which may inform directions of future research. First, participants represented a relatively small and unique sample of justice-involved adults with mental illnesses who were participating in mental health jail diversion programs in one U.S. county. Replication in other, larger samples of justice-involved adults with mental illnesses is needed. Second, recidivism data were limited to self-report and county-wide official records; we were unable to access state-wide records. As a result, our records may have underestimated true rates of recidivism; that is, participants may have been arrested and incarcerated outside of the county (or even the state). Relatedly, our recidivism measures were limited to arrests and jail days, both of which have the potential to be imperfect measures of criminal offending. In particular, the amount of time served in jail may be affected by a detainee’s financial resources, including ability to post bond and access a private attorney. Third, risk assessments in this study were completed by research assistants trained by the research team. Future research is needed to establish the field reliability and validity of START and LSI-R assessments completed by practitioners in mental health jail diversion programs. Fourth, the current investigation was part of a broader study evaluating the provision of behavioral health treatment services in the context of a successful mental health jail diversion program and, as such, lower predictive validity estimates could reflect true changes in recidivism risk attributable to intervention. Future research should examine associations between changes in START and LSI-R assessment scores attributable to treatment and recidivism risk in mental health jail diversion clients.
(Labrecque, Smith, Lovins, & Latessa, 2014; Vose et al., 2013). Findings of such a study may be particularly relevant to START with its inclusion of protective factors which may play a role in improving community supervision outcomes for offenders (Woldgabreel, Day, & Ward, 2014).

**Conclusions**

This study provides the first evidence supporting the reliability and validity of START and LSI-R assessments in predicting recidivism across multiple time frames among mental health jail diversion clients. It also provides some evidence supporting the use of START and LSI-R assessments—and of START strength total scores, in particular—with mental health jail diversion clients of diverse racial backgrounds, though questions remain regarding differences in the validity of START vulnerability and LSI-R total scores and risk estimates in predicting jail days in the short-term. Consistent with the Risk-Need-Responsivity model (Andrews, Bonta, et al., 1990), findings suggest that implementation of the START and LSI-R mental health jail diversion programs should assist in the identification of those at higher and lower risk of recidivism and, consequently, inform the development of case plans and risk management strategies. However, whether mental health jail diversion programs can employ the results of risk assessments using these (or other) instruments to successfully target risk and protective factors, improve rehabilitation, and, ultimately, decrease recidivism remains a critical avenue for future research.
Table 1. Descriptive Statistics for Arrests and Jail Days for All Follow-Up Periods

| Follow-up Period | Arrests       | |                | Jail Days     | |                |
|------------------|---------------|-----------------|---------------|-----------------|-----------------|
|                  | M  | SD  | Range        | M  | SD  | Range        |
| 3-Month          | 0.52 | 0.94 | 0-4         | 5.80 | 13.15 | 0-60         |
| 6-Month          | 0.82 | 1.31 | 0-7         | 10.51 | 23.56 | 0-153        |
| 9-Month          | 1.12 | 1.56 | 0-8         | 15.66 | 33.48 | 0-231        |
| 12-Month         | 1.38 | 1.77 | 0-8         | 19.78 | 37.78 | 0-245        |
| 18-Month         | 1.89 | 2.48 | 0-15        | 29.02 | 52.61 | 0-259        |

*Note. N = 94.*
Table 2. Distribution of and Concordance between LSI-R and START Risk Estimates

<table>
<thead>
<tr>
<th>LSI-R</th>
<th>START</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>n</td>
<td>%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td></td>
<td></td>
<td>n</td>
<td>%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n</td>
<td>%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>3</td>
<td>3.2</td>
<td></td>
<td>10</td>
<td>10.8</td>
<td>6</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>19</td>
<td>20.4</td>
<td></td>
<td>31</td>
<td>33.3</td>
<td>7</td>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>1.1</td>
<td></td>
<td>9</td>
<td>9.7</td>
<td>7</td>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. N = 93. κ = .04 (SE = .08), p = .623
Table 3. Validity of LSI-R and START Total Scores in Predicting Arrests and Jail Days across Follow-up Periods

<table>
<thead>
<tr>
<th>Follow-up period by predictor</th>
<th>Arrests</th>
<th></th>
<th></th>
<th></th>
<th>Jail days</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wald X²</td>
<td>IRR</td>
<td>95% CI</td>
<td></td>
<td>Wald X²</td>
<td>IRR</td>
<td>95% CI</td>
<td></td>
</tr>
<tr>
<td>START vulnerability total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Month</td>
<td>2.38</td>
<td>1.05</td>
<td>[0.99, 1.12]</td>
<td></td>
<td>1.05</td>
<td>1.02</td>
<td>[0.98, 1.05]</td>
<td></td>
</tr>
<tr>
<td>6-Month</td>
<td>3.38‡</td>
<td>1.05</td>
<td>[1.00, 1.11]</td>
<td>14.26***</td>
<td>1.07</td>
<td>[1.03, 1.10]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-Month</td>
<td>4.78*</td>
<td>1.06</td>
<td>[1.01, 1.11]</td>
<td>17.05***</td>
<td>1.07</td>
<td>[1.04, 1.11]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-Month</td>
<td>5.42*</td>
<td>1.06</td>
<td>[1.01, 1.11]</td>
<td>6.98**</td>
<td>1.04</td>
<td>[1.01, 1.07]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-Month</td>
<td>4.17*</td>
<td>1.04</td>
<td>[1.00, 1.09]</td>
<td>4.17*</td>
<td>1.03</td>
<td>[1.00, 1.06]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>START strength total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Month</td>
<td>16.00***</td>
<td>0.88</td>
<td>[0.82, 0.93]</td>
<td>58.59***</td>
<td>0.84</td>
<td>[0.80, 0.88]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-Month</td>
<td>17.27***</td>
<td>0.90</td>
<td>[0.85, 0.94]</td>
<td>68.75***</td>
<td>0.83</td>
<td>[0.80, 0.87]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-Month</td>
<td>13.74***</td>
<td>0.92</td>
<td>[0.88, 0.96]</td>
<td>21.27***</td>
<td>0.92</td>
<td>[0.89, 0.95]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-Month</td>
<td>12.96***</td>
<td>0.93</td>
<td>[0.89, 0.97]</td>
<td>6.72*</td>
<td>0.96</td>
<td>[0.93, 0.99]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-Month</td>
<td>7.30**</td>
<td>0.95</td>
<td>[0.91, 0.99]</td>
<td>4.56*</td>
<td>0.96</td>
<td>[0.93, 1.00]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSI-R total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Month</td>
<td>4.81*</td>
<td>1.07</td>
<td>[1.01, 1.14]</td>
<td>24.32***</td>
<td>1.11</td>
<td>[1.07, 1.16]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-Month</td>
<td>5.13*</td>
<td>1.06</td>
<td>[1.01, 1.12]</td>
<td>31.25***</td>
<td>1.13</td>
<td>[1.08, 1.18]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-Month</td>
<td>3.75‡</td>
<td>1.05</td>
<td>[1.00, 1.10]</td>
<td>20.49***</td>
<td>1.10</td>
<td>[1.05, 1.14]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-Month</td>
<td>4.76*</td>
<td>1.05</td>
<td>[1.00, 1.10]</td>
<td>8.38**</td>
<td>1.06</td>
<td>[1.02, 1.10]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-Month</td>
<td>0.27</td>
<td>1.01</td>
<td>[0.97, 1.05]</td>
<td>1.30</td>
<td>1.02</td>
<td>[0.98, 1.06]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


‡p < .10. *p < .05. **p < .01. ***p < .001.
Table 4. Validity of LSI-R and START Risk Estimates in Predicting Arrests and Jail Days across Follow-up Periods

| Follow-up period by predictor | Arrests | | | Jail days | | |
|-------------------------------|---------|---------|---------|---------|---------|
|                               | Low vs. high | Moderate vs. high | Low vs. high | Moderate vs. high |
|                               | Wald X² | IRR | 95% CI | Wald X² | IRR | 95% CI | Wald X² | IRR | 95% CI | Wald X² | IRR | 95% CI |
| START risk estimate           |         |       |        |         |       |        |         |       |        |         |       |        |
| 3-Month                       | 2.94‡ | 0.37   | [0.12, 1.15] | 0.32 | 0.79 | [0.34, 1.80] | 27.58*** | 0.15 | [0.07, 0.30] | 2.48 | 0.64 | [0.37, 1.11] |
| 6-Month                       | 2.85   | 0.43   | [0.16, 1.14] | 0.12 | 0.88 | [0.42, 1.84] | 39.40*** | 0.12 | [0.06, 0.23] | 7.13* | 0.48 | [0.28, 0.82] |
| 9-Month                       | 2.53   | 0.49   | [0.20, 1.18] | <0.01 | 1.00 | [0.50, 2.00] | 45.18*** | 0.11 | [0.06, 0.21] | 5.92* | 0.52 | [0.30, 0.88] |
| 12-Month                      | 4.99*  | 0.39   | [0.17, 0.89] | 0.81 | 0.74 | [0.39, 1.42] | 26.89*** | 0.19 | [0.10, 0.36] | 5.86* | 0.52 | [0.31, 0.88] |
| 18-Month                      | 1.32   | 0.62   | [0.28, 1.39] | 1.09 | 0.71 | [0.37, 1.35] | 3.03‡   | 0.55 | [0.28, 1.08] | 1.31 | 0.72 | [0.42, 1.26] |
| LSI-R risk estimate           |         |       |        |         |       |        |         |       |        |         |       |        |
| 3-Month                       | 4.80*  | 0.24   | [0.07, 0.86] | 1.92 | 0.56 | [0.24, 1.27] | 27.68*** | 0.13 | [0.06, 0.27] | 6.37* | 0.48 | [0.27, 0.85] |
| 6-Month                       | 4.85*  | 0.30   | [0.10, 0.88] | 1.24 | 0.65 | [0.31, 1.38] | 47.93*** | 0.07 | [0.03, 0.15] | 10.99** | 0.39 | [0.22, 0.68] |
| 9-Month                       | 5.27*  | 0.32   | [0.12, 0.85] | 1.15 | 0.68 | [0.34, 1.37] | 38.32*** | 0.11 | [0.05, 0.22] | 9.32** | 0.42 | [0.24, 0.73] |
| 12-Month                      | 5.89*  | 0.33   | [0.13, 0.81] | 2.42 | 0.59 | [0.30, 1.15] | 13.57*** | 0.28 | [0.14, 0.55] | 6.98** | 0.48 | [0.27, 0.83] |
| 18-Month                      | 2.78‡  | 0.52   | [0.25, 1.12] | 0.45 | 0.81 | [0.43, 1.51] | 14.20*** | 0.31 | [0.17, 0.57] | 2.28 | 0.67 | [0.39, 1.13] |

### Table 4. Moderation of Total Scores and Risk Estimates by Race on Recidivism at 3-Month Follow-up

<table>
<thead>
<tr>
<th>Interaction term</th>
<th>Arrests at 3-month</th>
<th>Jail days at 3-month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wald X²</td>
<td>IRR</td>
</tr>
<tr>
<td>START strength x race</td>
<td>0.24</td>
<td>0.97</td>
</tr>
<tr>
<td>START vulnerability x race</td>
<td>1.75</td>
<td>1.09</td>
</tr>
<tr>
<td>LSI-R total x race</td>
<td>0.18</td>
<td>1.03</td>
</tr>
<tr>
<td>START general offending (low) x race</td>
<td>1.23</td>
<td>4.61</td>
</tr>
<tr>
<td>START general offending (moderate) x race</td>
<td>0.17</td>
<td>0.70</td>
</tr>
<tr>
<td>LSI-R risk estimate (low) x race</td>
<td>0.16</td>
<td>1.78</td>
</tr>
<tr>
<td>LSI-R risk estimate (moderate) x race</td>
<td>0.24</td>
<td>0.66</td>
</tr>
</tbody>
</table>

**Notes.** N = 91-92. IRR = incidence rate ratio. CI = confidence interval for IRR. For risk estimates, High = 0. Interaction estimates produced in separate negative binomial regression models. Δ -2LL reflects improvement in model fit upon addition of the interaction term(s) in block 2. Block 1, not shown here, included main effects of race and risk assessment measures. ‡p < .10. *p < .05. **p < .01. ***p < .001.
Figure 1. Jail Days at 3-Month Follow-up by START General Offending Risk Estimate and Race.
Figure 2. Jail Days at 3-Month Follow-up by LSI-R Risk Estimate and Race
CHAPTER 3

Models of Protection against Recidivism in Justice-Involved Adults with Mental Illnesses

Protective factors may play an important role in the assessment, treatment, and supervision of adult offenders with mental illnesses. However, little is known regarding associations between protective factors, risk factors, and recidivism in this population. In research with adolescents, five models describing how protective factors and risk factors together are associated with adverse outcomes have been proposed and tested: 1) Compensatory, 2) Buffer, 3) Challenge, 4) Protective-Protective, and 5) Mediation. No study to date has tested and compared these models in adult offenders. To that end, this study evaluated these five models of protection against recidivism in a large sample of justice-involved adults with mental illnesses ($N = 550$). Risk and protective factors were operationalized using Short-Term Assessment of Risk and Treatability assessments (START; Webster et al., 2009) that were completed by case managers in the context of routine practice. Recidivism was measured by number of arrests and jail days one year following the date of the index START assessment. For both arrests and jail days, results showed consistent evidence for the Compensatory model, in which risk and protective factors were directly and independently associated with recidivism. Further, protective factors emerged as more robust predictors of recidivism relative to risk factors. In contrast, results failed to provide evidence supporting the other four models of protection. Overall, findings underscore the importance of considering both risk and protective factors in the assessment,
treatment, and supervision of adult offenders. Future efforts are needed to develop and test intervention strategies for promoting protective factors.

*Criminal Justice and Behavior*, in revision.
Introduction

Over the past two decades, there has been tremendous growth in the science and practice of conducting structured assessments of risk for violence and general offending (Desmarais et al., 2016; Skeem & Monahan, 2011). Across the globe, mental health practitioners report use of violence risk assessment and other instruments to identify individuals at heightened risk of violence and to inform risk management and treatment in the context of routine practice (e.g., Singh et al., 2014; Taxman, Cropsey, Young, & Wexler, 2007). Further, proper implementation of structured risk assessment approaches to inform intervention may reduce likelihood of negative outcomes in adult correctional and inpatient populations (Craissati, South, & Bierer, 2009; Doyle & Dolan, 2002; Singh et al., 2014; Vieira, Skilling, & Peterson-Badali, 2009; Vincent, Guy, Gershenson, & McCabe, 2012). Consequently, structured risk assessments may improve the risk management and treatment of adult offenders, particularly those with mental illnesses (i.e., major mental disorders including psychotic, mood, anxiety, and substance use disorders), who are more likely to reoffend relative to adult offenders generally (Baillargeon et al., 2009); yet, some limitations remain.

In particular, most of the work with adult offenders—with mental illness or generally—has focused on factors that increase risk of recidivism, termed “risk factors” (e.g., prior criminal history substance use), to the exclusion of factors that might protect against, buffer, or otherwise reduce the likelihood of such outcomes, termed “protective factors” (e.g., social support, employment; Rogers, 2000; Ryba, 2008). Indeed, a recent review of the
reliability and validity of instruments used to assess recidivism risk in US correctional settings found only two of 19 instruments—namely, the Inventory of Offender Risk, Needs, and Strengths (IORNS; Miller, 2006b, 2006a) and the Service Planning Instrument (SPIn; Orbis Partners, 2003)—included protective factors (Desmarais et al., 2016). Moreover, an international survey of forensic clinicians \((N = 199)\) revealed that participants were less likely to consider protective factors in the processes of assessing risk and planning treatment for adult offenders compared to juvenile offenders (Viljoen, McLachlan, & Vincent, 2010).

Though there is some debate in the field regarding the utility of protective factors, there are five reasons that protective factors may play a key role in the assessment, treatment, and supervision of risk for violence and other forms of recidivism among adult offenders. First, there is some empirical evidence demonstrating associations of protective factors with violent and general offending. For example, in a study of 1,311 offenders and non-offenders, Gilgun, Klein, and Pranis (2000) found that consideration of participants’ resources—specifically, feeling cared for and discussing problems with others—improved the accuracy over risk factors (i.e., substance use and delinquent behavior) in predicting offender classification. In a more recent investigation of protective factors associated with violent recidivism risk in 1,396 inmates released from prison in the UK, researchers found social support, emotional support, free time spent with family and friends, closeness to others, actively seeking work, and private (as opposed to public) housing were associated with lower risk of recidivism (Ullrich & Coid, 2011).

Second, protective factors have been shown to provide incremental validity over risk
factors in the assessment of violence risk. For instance, in a sample of 224 adolescent offenders, Lodewijks, Ruiter, and Doreleijers (2010) found protective factors, measured by the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2002) provided incremental validity in predicting violent offending above and beyond risk factors. Similarly, in a study of 120 male inpatients in a forensic psychiatric hospital, Short-Term Assessment of Risk and Treatability (START; Webster et al., 2009) strength total scores—which measure protective factors—added incremental validity to the Historical subscale scores of the Historical-Clinical-Risk-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997)—which measure risk factors—in predicting violent aggression toward others (Desmarais, Nicholls, et al., 2012). In another sample of 188 male forensic psychiatric inpatients, de Vries Robbé, de Vogel, and Douglas (2013) found Structured Assessment of Protective Factors (SAPROF; de Vogel, de Vries Robbé, de Ruiter, & Bouman, 2011) scores to increase the accuracy of HCR-20 total scores in predicting violent reconviction at 3-year and an average of 11.1 years follow-up.

Third, consideration of protective factors can enhance the therapeutic relationship between a client and therapist (Duckworth, Steen, & Seligman, 2005; Saleebey, 1996), contribute to a more neutral and less problem-oriented view of clients (Rogers, 2000; Webster et al., 2006), and orient the therapeutic session toward positive change (de Ruiter & Nicholls, 2011). In the context of offender supervision, specifically, promotion of offenders’ psychological flexibility, self-efficacy, optimism, and sense of hope may contribute to better community supervision outcomes, including fewer violations and lower reconviction rates.
Fourth, in many disciplines, mental health practitioners have a professional and ethical mandate to consider protective factors in assessment and intervention. For example, the American Psychological Association Presidential Task Force on Evidence-Based Practice (2006) recommends that practitioners “assess patient pathology as well as clinically relevant strengths” (p. 276). Similarly, the National Association for Social Workers “Standards for Social Work Practice in Health Care Settings” (NASW, 2005) describes consideration of client strengths as foundational to social work practice. Finally, consistent with the Risk-Need-Responsivity (RNR) and Good Lives models, protective factors may inform intervention by identifying factors that promote desistance (Andrews & Bonta, 2010; Ward, Mann, & Gannon, 2007).

It is not surprising, then, that several risk assessment instruments have been developed in recent years that include protective factors, such as the START (Webster, Martin, Brink, Nicholls, & Middleton, 2004; Webster et al., 2009), the SAPROF (de Vogel et al., 2011), the IORNS (Miller, 2006b, 2006a); the SPIn (Orbis Partners, 2003), the Static Risk and Offender Needs Guide (STRONG; Barnoski & Drake, 2007), and the Dynamic Risk Assessment for Offender Re-Entry (DRAOR; Serin, 2007). Research to date has supported the reliability and validity of assessments completed using these instruments in assessing risk of violence, violent offending, and general offending in diverse samples (Abidin et al., 2013; de Vries Robbê, de Vogel, & de Spa, 2011; Jones, Brown, Robinson, & Frey, 2015; Miller, 2015; O’Shea & Dickens, 2014; Yesberg & Polaschek, 2015).
Models of Protection

Despite the increasing presence of protective factors in violence risk assessment research and practice, there is no common definition or operationalization (for reviews, see de Ruiter & Nicholls, 2011; de Vries Robbé, Mann, Maruna, & Thornton, 2015). Some discuss protective factors as being the absence of risk factors (Jaffee, Moffitt, Caspi, & Taylor, 2003). Others argue that protective factors are independent of and qualitatively different from risk factors (Pandina, 1996; Webster et al., 2006) and have effects on adverse outcomes that are above and beyond the effects of risk factors (Farrington & Loeber, 2000). Protective factors also have been conceptualized as independent predictors, producing direct effects on negative outcomes; as moderator variables, interacting with the effect of risk factors on negative outcomes; and as dynamic predictors of negative behavioral outcomes over time (Costa, Jessor, & Turbin, 1999; Jessor, Van Den Bos, Vanderryn, Costa, & Turbin, 1995; van der Laan, Veenstra, Bogaerts, Verhulst, & Ormel, 2010). Most recently, Monahan and Skeem (2016) differentiated protective factors, which they describe as interacting with risk factors to decrease risk of recidivism, from promotive factors, which produce negative associations with recidivism. As a result of these inconsistent conceptualizations, our understanding of how protective factors are associated with adverse outcomes among adult offenders remains limited.

Five models of protection against adverse outcomes have been proposed and evaluated in the adolescent literature. The first is the Compensatory model, which describes protective factors and risk factors as having direct and independent effects on negative
outcomes (Zimmerman & Arunkumar, 1994). These direct and independent effects have been demonstrated in the prediction of violent offending (Salekin, Lee, Schrum Dillard, & Kubak, 2010) and, to some extent, aggressive behavior (Ladd & Burgess, 2001) among adolescent samples. These associations have also been demonstrated with respect to violent recidivism (de Vries Robbé et al., 2013; Miller, 2015; Ullrich & Coid, 2011) and general offending (Jones et al., 2015; Miller, 2015) among adult samples. However, in research in adult samples, the Compensatory model has not been explicitly identified, but rather inferred from the associations tested in the analytic models.

The second and third models are both moderation models that build on the direct effects described in the Compensatory model. The second model, or the Buffer model, proposes that protective factors weaken the association between risk factors and negative outcomes (Rutter, 1985). This model has been explicitly tested and supported in the adolescent research literature across multiple adverse outcomes, including offending, aggression, violent behavior, and other problem behaviors (Aceves & Cookston, 2006; Caldwell, Kohn-Wood, Schmeelk-Cone, Chavous, & Zimmerman, 2004; Farrell, Henry, Schoeny, Bettencourt, & Tolan, 2010; Farruggia & Germo, 2015; Jessor et al., 1995; Jessor, Turbin, & Costa, 1998; Lee, 2005; Lodewijks et al., 2010; Rennie & Dolan, 2010). Associations proposed by the Buffer model also have been tested in the prediction of violent (de Vries Robbé et al., 2013) and general recidivism (Jones et al., 2015) among adult offenders, but findings provide very limited support for the model.

The third model, the Protective-Protective model, proposes that the association
between risk factors and negative outcomes is weakened by the addition of each protective factor (Hollister-Wagner, Foshee, & Jackson, 2001). Thus, the Protective-Protective model is distinct from the previous two in that it considers the number of protective factors instead of a composite measure representing the degree or total amount of protection. Overall, this model has been infrequently evaluated. Though there is some empirical evidence supporting the associations described by the Protective-Protective model with respect to the prediction of aggression in adolescents (Donnon, 2010), the model has not been tested in adults, to our knowledge.

The fourth model, the Challenge model, builds upon the first two models. This model describes the association between risk factors and negative outcomes as depending on the degree of exposure to risk factors, such that moderate exposure to risk factors may decrease negative outcomes (Garmezy, Masten, & Tellegen, 1984). Research conducted in adolescent samples has provided support for the Challenge model in the prediction of adverse outcomes (e.g., Christiansen & Evans, 2005; Erdem & Slesnick, 2010). In contrast, the Challenge model has not been tested in adult samples and actually may be less relevant to adult compared to adolescent offenders. Specifically, adolescence is generally viewed as a period of plasticity, marked by growth and change (e.g., Holmbeck, Devine, Wasserman, Schellinger, & Tuminello, 2012; Toth & Cicchetti, 1999), whereas adolescents who persist in antisocial behavior into adulthood are more entrenched in negative behavior and less amenable to change (Moffitt, 1993). Thus, a moderate amount of risk factors in adolescents may reflect a transitional point on a developmental trajectory, but a moderate amount of risk
factors in adult offenders may be more likely to represent actual risk for negative outcomes.

The fifth and final model is the Mediation model, which has been tested extensively with adolescent samples, though it has not been explicitly labeled as such. This model suggests that risk factors operate through (i.e., are mediated by) the presence of protective factors. This pattern of associations has been supported in adolescent samples with respect to a variety of protective factors and across several outcomes, including suicide and suicidal ideation (Bagge, Lamis, Nadorff, & Osman, 2014), substance use (Madson, Moorer, Zeigler-Hill, Bonnell, & Villarosa, 2013; Thai, Connell, & Tebes, 2010), and violence (Reingle, Jennings, & Maldonado-Molina, 2012). To date, however, this model has not been tested in adult samples.

**Study Purpose**

For more than a decade, there have been calls for research examining how protective factors relate to risk for violence and recidivism in adult offenders. In contrast with the theoretical and empirical advances in our understanding of protective factors in the field of adolescent risk assessment, the five conceptual models of protection reviewed have been applied inconsistently and with limited evaluation in adult samples. However, there may exist meaningful differences in how protective factors contribute to risk in adults compared to adolescents (Viljoen, Cruise, Nicholls, Desmarais, & Webster, 2012). As such, there is a need for a systematic evaluation of these models of protection in adults. To that end, this study examined and compared five models describing associations between protective factors, risk factors, and recidivism in a sample of justice-involved adults with mental
illnesses.

Method

Study Setting and Design

This study involved a secondary analysis of administrative records on clients participating in mental health jail diversion programs in a large, metropolitan county in the southern United States. Data for this study were drawn from primarily misdemeanor and felony post-booking, pre-trial jail diversion programs. As part of the diversion program, participants received optional case management services (e.g., referral to community treatment and housing services) for one year. However, participants were not under judicial supervision as a function of their diversion participation. Diversion program participation was only revoked in the case of a new offense. In 2011, the START was implemented throughout the jail diversion programs as part of the intake process to enhance the uniformity, utility, and accuracy of risk assessments and case management strategies in this setting. Data for this study were drawn from all clients who had a START assessment completed between 2011 and 2013, which represented the 3-year evaluation period for this pilot implementation.

Prior to the implementation of START assessments, staff participated in a 1-day, in-person training on the START assessment provided by a START author (S.L.D.). The training included an introduction to the structured professional judgment approach to risk assessment, an introduction to the START, an overview of START coding, and a hands-on coding session where each staff person coded two cases, which were then reviewed as a
group to achieve consensus. Staff also participated in three separate half-day booster sessions that were provided at three to six-month intervals. During these trainings, staff received an overview of the START, feedback on completed assessments, and additional practice coding and receiving feedback on START assessments (Desmarais, Van Dorn, et al., 2012). When multiple START assessments were conducted per client, we included the START assessment completed at program intake.

**Participants**

The study sample consisted of 550 mental health jail diversion participants. The average age of the sample was 37.45 years (SD = 13.24). The majority of participants were male (n = 428, 77.8%) and of both White (n = 298, 54.2%) and non-White (n = 252, 45.8%) racial backgrounds. The most frequent primary diagnosis was schizophrenia spectrum (n = 340, 61.8%), followed by bipolar (n = 125, 22.7%), and depressive (n = 61, 11.1%) disorders. Other disorders and missing data accounted for an additional 4.3% (n = 24). Participants were primarily in the misdemeanor (n = 282, 51.3%) and felony (n = 223, 40.5%) jail diversion programs.

**Measures**

Variables of interest in this study included: 1) risk factors, operationalized via START vulnerability total scores; 2) protective factors, operationalized using START strength total scores; 3) number of protective factors, operationalized via START strength items; 4) recidivism, including arrests and jail days; and 5) participant characteristics, including sex, age, race, and mental health diagnosis. All data were drawn from
administrative records.

**START assessments.** The START is a structured professional judgment instrument designed to assess risk of violence to others, suicide, self-harm, victimization, substance use, unauthorized leave, and self-neglect (Webster et al., 2004, 2009). START assessments have demonstrated reliability and validity in predicting violence among adults with mental illnesses (for a review, see O’Shea & Dickens, 2014). Recently, the START was piloted in a jail diversion setting, demonstrating good psychometric properties for assessments completed in this setting (Desmarais, Van Dorn, et al., 2012). Results from this pilot suggested inclusion of an eighth risk estimate to predict general offending, which was evaluated vis-à-vis recidivism with promising results in a recent investigation (Lowder, Desmarais, Rade, Johnson, & Van Dorn, in press). Indeed, justice-involved adults with mental illnesses are not only at high risk of reoffending (Baillargeon et al., 2009), they are also experience elevated risk for multiple outcomes measured by the START instrument (McNiel, Binder, & Robinson, 2005; Ogloff, 2002; Van Dorn et al., 2011; Wallace, Mullen, & Burgess, 2004), and there is substantial overlap in the factors that have been shown to increase risk for these outcomes (e.g., substance use, antisocial attitudes). As such, a comprehensive strategy assessing these shared risk and protective factors may inform management strategies that are likely to be more effective and efficient.

The START includes 20 items that are assessed as both risk and protective factors. Items cover domains related to psychosocial functioning (e.g., occupational, recreational, self-care, social support), mental illness and substance use (e.g., mental state, emotional state,
substance use), and treatment (e.g., medication adherence, treatability), among others. For each item, assessors provide separate 0 (minimally present), 1 (moderately present), or 2 (maximally present) ratings for strength and vulnerability in the past three months. That is, strength and vulnerability ratings are completed independently for each item, and a client may demonstrate strength and vulnerability with respect to the same risk domain (Webster et al., 2009). Item ratings can be summed in research (but not in practice) to create strength and vulnerability total scores, which were used to operationalize protective and risk factors, respectively, in the present study. The possible range of these total scores is 0 to 40, with higher scores representing greater risk for vulnerability ratings and greater protection for strength ratings. When assessments have five or fewer missing item ratings, total scores are prorated by dividing the summed item ratings by the total possible scale score (i.e., 40), multiplying the value by the total number of omitted items (i.e., 1-5), and adding the original score to this obtained value (see Webster et al., 2009, p. 34). The number of protective factors variable was created by counting the number of strength items rated as 1 (moderately present) or 2 (maximally present).

**Recidivism.** Raw jail booking records were queried by the jail diversion coordinator and provided to the research team for tabulation. Number of arrests were calculated using date of jail booking as an indication of arrest and summing the number of jail bookings for each participant for 12 months following the date of the index START assessment. Arrests for both new offenses as well as other reasons i.e., technical violations, outstanding warrants, etc.) were included in this tabulation. Jail days were calculated from date of booking to date
of release (inclusive of end points) for 12 months following the date of the START assessment. Time at risk was included in all models of arrest and measured the number of days (out of 365) a participant was in the community (i.e., not in jail).

**Participant characteristics.** Participant characteristics included age at the time of referral to the program (continuous), sex (0=male, 1=female), race (0=White, 1=non-White), and dummy-coded mental health diagnosis, with schizophrenia spectrum disorders (i.e., schizophrenia, schizophreniform, or schizoaffective diagnoses) as the reference group, resulting in two variables: bipolar disorder (0=schizophrenia spectrum, 1=bipolar) and major depressive disorder (0=schizophrenia spectrum, 1=depressive). Information on psychiatric diagnoses was pulled from existing clinical records maintained by the jail diversion staff.

**Data Cleaning**

An outlier analysis was conducted on START strength and vulnerability total scores by diversion program; three cases were identified as having outlying strength scores. These three cases were excluded from analysis via listwise deletion. Recidivism data were unavailable for an additional 15 participants. There also were some START assessments for which a strength \( n = 17 \) or vulnerability \( n = 15 \) total scores could not be computed due to missing item-level data.

**Analytic Plan**

**Preliminary analyses.** Prior to hypothesis testing, we computed descriptive statistics for all study variables, examined the internal consistency (i.e., Cronbach’s alpha) of the START assessments, and examined multicollinearity between START strength and
vulnerability items and total scores. We additionally conducted bivariate comparisons between participant characteristics, strength and vulnerability total scores, and recidivism measures to identify potential covariates. A cutoff value of \( p < .10 \) determined inclusion of covariates.

**Models of protection.** To test the first four models of protection, we conducted two sets of hierarchical regression analyses (i.e., one for each recidivism variable), consistent with analytic approaches used in adolescent samples (Garmezy et al., 1984; Hollister-Wagner et al., 2001; Pollard, Hawkins, & Arthur, 1999). The first set of hierarchical regression analyses tested the Compensatory, Buffer, and Challenge models. In the first block, covariates were added. For the arrest model, time at risk was additionally included in the first block. In the second block, vulnerability total scores and strength total scores were added to test the Compensatory model. In the third block, a vulnerability x strength product term was added to evaluate the Buffer model. In the fourth and final block, a quadratic vulnerability x vulnerability term was added to assess for evidence of the Challenge model. The second set of hierarchical regression analyses tested the Protective-Protective model. Block one was identical to that in the first set of analyses. In the second block, strength total scores were substituted with number of protective factors. In the third block, the vulnerability x strength product term was substituted with a number of protective factors x vulnerability total scores product term to test the Protective-Protective model. Improvement in model fit was determined by calculating the difference in \(-2 \log(\text{Likelihood})\) statistics between blocks and computing a \( p \) value for chi-square statistics, as well as examination of model fit indices (i.e.,
AIC, BIC). Because recidivism variables were positively skewed, we employed negative binomial regression analyses in all hierarchical models (Walters, 2007). Though precise statistical power calculations are available for discrete predictors in negative binomial models (Cundill & Alexander, 2015; Zhu & Lakkis, 2014), few guidelines exist for continuous predictors. However, given the high response rates in the present study (e.g., 45.3% of participants were arrested during the follow-up period), we deemed our sample size sufficient based on previously published estimates for generalized linear models (Shieh, 2000).

Finally, the Mediation model was tested in separate regression analyses using the conditional process analysis approach of Hayes (2013). Effects were estimated using OLS regression coefficients and quantified using bias-corrected bootstrapped confidence intervals produced by the PROCESS macro for SPSS.

Results

Descriptive Statistics

The mean START vulnerability total scores was 18.02 out of 40 ($SD = 7.14$, range 0 to 35), which was higher than the mean START strength total score of 11.07 out of 40 ($SD = 7.72$, range 0 to 35). START vulnerability scores were more normally distributed, with a skewness of -0.25 ($SE = 0.11$) and kurtosis of -0.37 ($SE = 0.21$), relative to START strength scores, which had a skewness of 0.46 ($SE = 0.11$) and kurtosis of -0.71 ($SE = 0.21$). Participants had an average of 9.83 ($SD = 6.28$) protective factors, reflecting START strength item ratings of moderately present (1) or maximally present (2). Assessments of both START
strength and vulnerability scores showed good levels of internal consistency: Cronbach’s coefficients of .93 and .90, respectively. During the 12-month follow-up period, participants were arrested, on average, 1.17 (SD = 2.25, range 0 to 23) times, which corresponded to an average of 19.41 days in jail (SD = 45.59, range 0 to 324).

**Bivariate Associations**

We first examined associations between item-level ratings, and second, between the total scores. Strength and vulnerability ratings were significantly correlated for all items except Conduct, Attitudes, External Triggers, or Recreational (ps ≥ .253). The strongest correlation was observed between strength and vulnerability ratings on Material Resources, $r (531) = -.49, p < .001$. The correlation between strength total scores and vulnerability total scores was small, $r (530) = -.09, p = .032$. Variance inflation factors (VIFs) were only 1.00 when each total score was regressed on the other. Taken together, these results failed to show multicollinearity and support the inclusion of strength and vulnerability ratings in the same models.

We conducted bivariate comparisons to identify covariates. Diagnosis was associated with strength total scores ($F [2, 507] = 9.89, p < .001$), arrests ($\chi^2 [2] = 10.49, p = .005$), and jail days ($\chi^2 [2] = 26.89, p < .001$). Race was associated with strength total scores ($t [531] = 4.08, p < .001$), arrests ($\chi^2 [1] = 9.02, p = .003$), and jail days ($\chi^2 [1] = 7.41, p = .006$). Sex was related to vulnerability total scores ($t [221.93] = -2.24, p = .026$), arrests ($\chi^2 [1] = 8.26, p = .004$), and jail days ($\chi^2 [1] = 15.50, p < .001$). Age was not related to total scores or recidivism measures (ps ≥ .292). Thus, diagnosis, race, and sex were included as covariates.
in subsequent analyses.

**Model Testing**

Findings for hierarchical negative binomial regression analyses testing the Compensatory, Buffer, Challenge and Protective-Protective models are presented in Table 1 and Table 2. Table 3 presents results of analyses testing the Mediation model.

**Compensatory model.** First, we tested whether risk and protective factors showed direct and independent effects on recidivism (i.e., theCompensatory model). For both measures of recidivism, strength total scores were significantly and inversely associated with risk of reoffending, such that each 1-point increase in START strength scores was associated with 1.09 times fewer jail days and 1.06 times fewer arrests ($p < .001$). In contrast, START vulnerability scores were only significantly associated with jail days, such that a 1-point increase in vulnerability total scores was associated with 1.05 times more jail days, $p < .001$. No significant association was observed between START vulnerability scores and arrests, $p = .356$ (see Table 1). Model fit statistics (presented in Table 2) showed significant improvements in fit for the Compensatory model relative to the covariate-only model, $p < .001$.

**Buffer model.** Second, we tested whether protective factors moderated the association between risk factors and recidivism (i.e., the Buffer model). For both outcomes, results did not show significant interactions between strength and vulnerability total scores (see strength x vulnerability interaction terms in Table 1), $ps \geq .227$. Moreover, the addition of an interaction term did not improve model fit relative to the Compensatory Model (see
Overall, these findings fail to provide support for the Buffer model.

**Challenge model.** Third, we tested whether risk factors demonstrated a curvilinear association with recidivism (i.e., the Challenge model). Results provide little support for this model. Across outcomes, an exponential increase in vulnerability total scores was not associated with increased risk of recidivism, $ps \geq .575$ (see vulnerability x vulnerability interaction terms in Table 1). Further, the addition of the vulnerability x vulnerability interaction term did not improve model fit relative to the Buffer Model, $ps \geq .718$ (see Table 2).

**Protective-Protective model.** Fourth, we tested whether the association between risk factors and recidivism was moderated by number of protective factors (i.e., the Protective-Protective model). Results fail to provide evidence supporting this model. There was no interaction between number of protective factors and vulnerability total scores on arrests or jail days, $ps \geq .204$ (see vulnerability x protective terms in Table 1), and the Protective-Protective model did not improve upon the fit of the Compensatory Model, $ps \geq .202$ (see Table 2).

**Mediation model.** Finally, we tested whether the association between risk factors and recidivism operated through protective factors (i.e., the Mediation model). As shown in Table 3, significant direct effects emerged between risk factors and jail days ($p = .004$), but not between risk factors and arrests ($p = .265$). However, results did not provide evidence of a mediating role of protective factors in the association between risk factors and recidivism for both outcomes, $ps > .05$ (see confidence intervals for indirect effect terms).
Model Comparison

Finally, we compared models in terms of their fit and strength of associations between risk and protective factors with recidivism. The Compensatory model was the only model that produced both: a) significant associations between protective factors, risk factors, and recidivism, and b) improvement in model fit over the covariate-only model. The Buffer, Challenge, and Protective-Protective models evinced better fit statistics, but did not yield significant predictors (i.e., risk or protective factor terms).

Discussion

Protective factors have been identified as important considerations in the assessment (Ullrich & Coid, 2011), treatment (de Ruiter & Nicholls, 2011; Duckworth et al., 2005; Saleebey, 1996), and supervision of adult offenders (de Ruiter & Nicholls, 2011; Woldgabreal et al., 2014). However, there has been limited research on how to incorporate protective factors in the risk assessment process in this population (de Ruiter & Nicholls, 2011; Ullrich & Coid, 2011). In contrast, five models of protection that describe how risk and protective factors relate to each other to affect risk of adverse outcomes have been proposed, evaluated, and compared in adolescent samples (Garmezy et al., 1984; Hollister-Wagner et al., 2001; Reingle et al., 2012; Rutter, 1985; Zimmerman & Arunkumar, 1994). To date, these models have not been systematically tested in adult offenders, which was the goal of the present study.

Specifically, we tested the Compensatory, Buffer, Challenge, Protective-Protective, and Mediation models. The Compensatory model describes independent and direct effects of
risk and protective factors on recidivism. In particular, protective factors, operationalized as START strength total scores, demonstrated consistent negative associations with recidivism across both recidivism measures (i.e., arrests and jail days). Similarly, risk factors, operationalized as START vulnerability total scores, showed a direct and positive association with jail days. These findings are consistent with previous research conducted among adolescents with respect to the prediction of violent offending (Salekin et al., 2010) and aggressive behavior (Ladd & Burgess, 2001). These findings are also consistent with the extant, but relatively small, body of work examining risk and protective factors in adults. Specifically, prior research has established direct and independent effects of risk and protective factors on violent reconviction (de Vries Robbé et al., 2013; Ullrich & Coid, 2011) and general recidivism (Jones et al., 2015; Miller, 2015). In other adult populations (e.g., forensic and civil psychiatric inpatients), START strength and vulnerability total scores similarly have shown direct and independent effects on violence and aggression (Braithwaite et al., 2010; Chu et al., 2011a, 2011b; Desmarais, Nicholls, et al., 2012; O’Shea & Dickens, 2014).

Findings of the present study failed to provide support for an interaction effect between risk and protective factors on recidivism, as described by the Buffer model. These findings are in contrast to those reported in the adolescent research, which has found overwhelming support for the Buffer model across many samples and adverse outcomes (e.g., Aceves & Cookston, 2006; Caldwell et al., 2004; Jessor et al., 1998; Lodewijks et al., 2010). Yet, our findings are not altogether inconsistent with findings of research on
protective factors conducted in other adult offender samples. In particular, Ullrich and Coid (2011) failed to find a significant interaction effect between protective factors and risk factors on violent reconviction risk five years following assessment among male prisoners. Although de Vries Robbé and colleagues (2013) reported a significant interaction between protective factors and risk factors, measured by the SAPROF and HCR-20, respectively, on reconviction risk over a 3-year period among male offenders, they observed multicollinearity between their model terms (VIF = 3.06), which should have precluded inclusion of risk and protective factors in the same analytic model. As one last example, a study of adult offenders revealed a significant interaction between protective factors and risk factors, measured by risk and strength scores on the SPIIn, on risk for new offenses over an 18-month follow-up period (Jones et al., 2015); however, the size of the effect was very small (i.e., OR = 1.001). Thus, a critical review of the empirical evidence to date suggests that protective factors may not buffer against risk of recidivism in adults as they do in adolescents.

Like the Buffer model, the associations between risk and protective factors described by the Challenge, Protective-Protective, and Mediation models were not supported in the current study. These models demonstrated reasonable fit with the study data; yet, they did not add incrementally to the prediction of recidivism relative to the Compensatory model. These findings are in contrast to those reported in studies of adolescents, which have shown support for the Challenge (e.g., Christiansen & Evans, 2005; Erdem & Slesnick, 2010), Protective-Protective (Donnon, 2010) and Mediation (Bagge et al., 2014; Reingle et al., 2012; Thai et al., 2010) models. To our knowledge, these models of protection have not been tested with
respect to the prediction of violence or offending in adult samples.

Overall, the lack of support for the Buffer, Challenge, Protective-Protective, and Mediation models suggests that protective factors may be associated with risk factors and adverse outcomes in meaningfully different ways in adults and adolescents. Specifically, risk and protective factors may be more potent for adolescents relative to adults because adolescents are already at heightened risk for very specific adverse outcomes (e.g., self-harm and substance use; Boyer, 2006; Steinberg, 2004). Additionally, and related to the previous point, protective factors may be more directly responsive to risk factors for this group because adolescence is a period of developmental change and skill acquisition (Steinberg & Cauffman, 1996; Steinberg & Monahan, 2007; Steinberg et al., 2007). Thus, the capacity for skill development in adolescents may better allow for the buffering or attenuating of risk factors seen in this population. Further, adolescents experience developmental transitions where limited exposure to risk factors can be considered normal and part of a developmental trajectory (Jessor, 1991; Moffitt, 1993). After an initial exposure to risk factors, adolescents may develop the skills to become less susceptible to peer influence and able to avoid problem behaviors (Modecki, 2008; Moffitt, 1993; Monahan, Steinberg, Cauffman, & Mulvey, 2009). In comparison, adolescents who persist in their antisocial behavior to become adult offenders are likely less amenable to change (Moffitt, 1993).

Limitations

Findings should be considered in light of several limitations to the study design. First, we used administrative data derived from assessments completed using one instrument, the
START, in a mental health jail diversion sample. Generalizability of findings to assessments completed with other instruments, with respect to other outcomes, and in other samples of adult offenders will need to be established. Second, we were unable to establish inter-rater reliability because START assessments were completed during routine practice. That said, case managers completed rigorous training using protocols similar to those shown to produce good inter-rater agreement in other studies (Desmarais, Nicholls, et al., 2012). Moreover, the lack of data on inter-rater reliability, though not ideal, is not uncommon in the risk assessment literature, particularly in field studies (Desmarais et al., 2016). Third, we did not evaluate associations between risk factors, protective factors, and recidivism over multiple time points, which would have provided a stronger test of the mediation model (Douglas & Skeem, 2005; MacKinnon, Fairchild, & Fritz, 2007). Likewise, OLS regression was used to test the mediation model, but our recidivism data were not normally distributed. Though a nonparametric approach to mediation analyses has been developed (e.g., Imai, Keele, & Tingley, 2010), the present data did not meet independence requirements between predictors and mediators. Fourth, START strength and vulnerability item ratings were summed to create total scores, which should not be done in practice (Webster et al., 2009). As such, there is a need for research on associations between risk factors and protective factors as they are used in practice (Guy, Douglas, & Hart, 2015).

**Implications**

The limitations notwithstanding, the current study represents the first comprehensive test of models of protection found in the adolescent literature in an adult offending sample.
Only two of the models had previously been tested in the adult samples and only in a piecemeal fashion (de Vries Robbé et al., 2013; Gagliardi, Lovell, Peterson, & Jemelka, 2004; Jones et al., 2015; Miller, 2015; Ullrich & Coid, 2011). Our analyses suggest that risk factors and protective factors are distinct constructs that exert direct and independent effects on recidivism risk. As such, the Compensatory model (Zimmerman & Arunkumar, 1994) may be the most appropriate conceptualization of protective factors in adult offenders. Consequently, practitioners should consider both risk and protective factors as unique considerations and targets in the assessment, treatment, and supervision of adult offenders. Though we examined protective factors measured using the START, our findings lend support for the use of instruments that include protective factors in the assessment and management of recidivism risk among adult offenders.

Findings of the current study also provide several directions for future research. First, this is only the first empirical evaluation and comparison of multiple models of protection in an adult offender sample; replication is needed. Associations between risk factors, protective factors, and recidivism may differ as a function of the risk assessment instrument, sample, and outcome. Second, conceptualizations of protective factors as buffering against or interacting with risk factors and distinctions between promotive and protective factors (Monahan & Skeem, 2016), as described earlier, were not supported herein. However, research in other samples and with assessments completed using other instruments is needed to establish the Compensatory model as the most appropriate conceptualization of risk and protective factors in adult offenders. Third, future research should investigate models of
protection over time. Though the current study provides evidence supporting risk factors and protective factors as distinct constructs, we do not know how the relationship between risk and protective factors may change over time, as risk and protective factors themselves change over time. Fourth, and finally, research is needed to identify strategies that may be effective in developing or enhancing protective factors and reducing risk of adverse outcomes over time (de Ruiter & Nicholls, 2011).

**Conclusions**

Although many have argued that protective factors are important considerations in risk assessment, treatment, and supervision of adult offenders, our conceptual understanding of the relationship between risk and protective factors, and their impact on recidivism risk, is limited. Though much work has been done on protective factors in adolescent offenders, the present study represents the first comparative evaluation of five models of protection against recidivism in an adult offender sample. Findings indicate that risk factors and protective factors are distinct constructs that have direct and independent effects on recidivism, and suggest important differences between adolescent and adult offenders with regard to the associations of risk and protective factors with recidivism risk. They also provide empirical support for the role of protective factors in risk assessment and management of adult offenders. However, further research is needed to test the generalizability of the current study’s findings to other assessment instruments and other samples of adult offenders and to inform how protective factors may be targeted in intervention to reduce recidivism in this population.
### Table 1. Summary of Negative Binomial Regression Results by Model of Protection

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Recidivism outcomes</th>
<th>Number of arrests</th>
<th></th>
<th></th>
<th>Number of jail days</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Wald</td>
<td>IRR</td>
<td>IRR 95% CI</td>
<td>Wald</td>
<td>IRR</td>
</tr>
<tr>
<td>Step 1: Covariate-only model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td>1.35</td>
<td>1.22</td>
<td>[0.87, 1.71]</td>
<td>11.55**</td>
<td>1.46</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td>1.93</td>
<td>0.83</td>
<td>[0.63, 1.08]</td>
<td>15.13***</td>
<td>0.70</td>
</tr>
<tr>
<td>Bipolar</td>
<td></td>
<td>5.30*</td>
<td>0.67</td>
<td>[0.47, 0.94]</td>
<td>18.47***</td>
<td>1.61</td>
</tr>
<tr>
<td>Depressive</td>
<td></td>
<td>5.04*</td>
<td>0.59</td>
<td>[0.37, 0.93]</td>
<td>18.01***</td>
<td>1.85</td>
</tr>
<tr>
<td>Time at risk</td>
<td></td>
<td>102.23***</td>
<td>0.98</td>
<td>[0.98, 0.98]</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Step 2: Compensatory model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vulnerability total score</td>
<td></td>
<td>0.85</td>
<td>1.01</td>
<td>[0.99, 1.03]</td>
<td>56.11***</td>
<td>1.05</td>
</tr>
<tr>
<td>Strength total score</td>
<td></td>
<td>32.67***</td>
<td>0.94</td>
<td>[0.92, 0.96]</td>
<td>107.87***</td>
<td>0.93</td>
</tr>
<tr>
<td>Step 3: Buffer model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strength x vulnerability</td>
<td></td>
<td>1.46</td>
<td>1.00</td>
<td>[1.00, 1.00]</td>
<td>0.23</td>
<td>1.00</td>
</tr>
<tr>
<td>Step 4: Challenge model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vulnerability x vulnerability</td>
<td></td>
<td>0.13</td>
<td>1.00</td>
<td>[1.00, 1.00]</td>
<td>0.31</td>
<td>1.00</td>
</tr>
<tr>
<td>Step 2: Compensatory model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vulnerability total score</td>
<td></td>
<td>0.77</td>
<td>1.01</td>
<td>[0.99, 1.03]</td>
<td>54.84***</td>
<td>1.05</td>
</tr>
<tr>
<td>Number protective factors</td>
<td></td>
<td>28.82***</td>
<td>0.94</td>
<td>[0.92, 0.96]</td>
<td>88.27**</td>
<td>0.93</td>
</tr>
<tr>
<td>Step 3: Protect-protect model</td>
<td></td>
<td>1.61</td>
<td>1.00</td>
<td>[1.00, 1.00]</td>
<td>0.28</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Notes. $N = 496-511$. IRR = Incidence rate ratio. CI = confidence interval. For Sex, female participants represent the reference group. For Race, non-White participants represent the reference group. Bipolar and Depressive variables are dummy coded whereby participants with schizophrenia spectrum diagnoses represent the reference group. Each model shows only unique terms added; terms in previous steps were included in subsequent steps. Compensatory model replaced strength total score with number of protective factors; Step 1 was the same in this model.

‡$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$. 
Table 2. Model Fit Statistics for Hierarchical Negative Binomial Regression Analyses by Model of Protection

<table>
<thead>
<tr>
<th>Models</th>
<th>Number of arrests</th>
<th>Recidivism outcomes</th>
<th>Number of jail days</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\chi^2(df))</td>
<td>-2LL</td>
<td>AIC</td>
</tr>
<tr>
<td>Step 1: Covariate-only</td>
<td>191.12*** (5)</td>
<td>456.61</td>
<td>1,352.15</td>
</tr>
<tr>
<td>Step 2: Compensatory</td>
<td>226.04*** (7)</td>
<td>406.32***</td>
<td>1,282.44</td>
</tr>
<tr>
<td>Step 3: Buffer</td>
<td>227.52*** (8)</td>
<td>404.84</td>
<td>1,282.96</td>
</tr>
<tr>
<td>Step 4: Challenge</td>
<td>227.65*** (9)</td>
<td>404.71</td>
<td>1,284.83</td>
</tr>
<tr>
<td>Step 2: Compensatory^</td>
<td>221.44*** (7)</td>
<td>410.92***</td>
<td>1,287.04</td>
</tr>
<tr>
<td>Step 3: Protect-protect</td>
<td>223.07*** (8)</td>
<td>409.29</td>
<td>1,287.40</td>
</tr>
</tbody>
</table>

Notes. Ns = 496-511. \(\hat{\chi} = \) chi-square likelihood ratio statistic for model fit. -2LL = -2 Log(Likelihood) or deviance statistic. Model improvement from previous step to next step represented by significance indicators on -2LL, calculated using difference in -2LL and chi-square statistic with 1-2 degrees of freedom. Compensatory^ = Direct effects of risk factors and number of protective factors on outcome; significance determined by comparison with covariate-only model. Protect-protect = Protective-Protective model; significance determined by comparison with Compensatory^ model. ***p < .001
Table 3. Mediation of START Vulnerability Scores and Recidivism by START Strength Scores

<table>
<thead>
<tr>
<th>Effects</th>
<th>Number of arrests</th>
<th></th>
<th></th>
<th>Number of jail days</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>95% CI</td>
<td>Estimate</td>
<td>SE</td>
<td>95% CI</td>
</tr>
<tr>
<td>Count outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct effect</td>
<td>0.02</td>
<td>0.01</td>
<td>[-0.01, 0.04]</td>
<td>0.83**</td>
<td>0.28</td>
<td>[0.27, 1.39]</td>
</tr>
<tr>
<td>Indirect effect</td>
<td>0.003</td>
<td>0.003</td>
<td>[-0.002, 0.01]</td>
<td>0.07</td>
<td>0.05</td>
<td>[-0.01, 0.18]</td>
</tr>
<tr>
<td>Total effect</td>
<td>0.02</td>
<td>0.01</td>
<td>[-0.01, 0.04]</td>
<td>0.90**</td>
<td>0.29</td>
<td>[0.33, 1.46]</td>
</tr>
</tbody>
</table>

Notes. N = 496. **p < .01
CHAPTER 4

Recidivism Following Mental Health Court Exit: Between and Within-Group Comparisons

Over the past decade, Mental Health Courts (MHCs) have spread rapidly across the U.S. These courts aim to reduce recidivism among adults with mental illnesses through diversion into community-based treatment. Extant research suggests that MHCs can be effective in reducing recidivism, but also demonstrates that effectiveness varies as a function of characteristics of the participants (e.g., criminal history) and the program (e.g., coercion). Less is known regarding the extent to which process-related factors (e.g., length of participation, time between referral and receipt of services) impact effectiveness. Prior research also is limited by a focus on recidivism during MHC as opposed to post-exit. To address these knowledge gaps, we examined recidivism one year post-exit for a group of MHC participants \((n = 57)\) and offenders receiving treatment as usual (TAU; \(n = 40)\), total \(N = 97\). We also investigated the influence of individual characteristics and process factors on changes in jail days one year pre-entry to one year post-exit for MHC participants. Overall, results provide some evidence supporting the effectiveness of MHCs. MHC participants had significantly fewer jail days, but not charges or convictions, relative to TAU participants. Among MHC participants, graduation from the MHC, presence of co-occurring substance use, and longer length of MHC participation were associated with greater reductions in jail days. Other process factors were unrelated to reductions in recidivism. Findings suggest that
MHCs may be particularly effective for high-risk participants and that time spent in a MHC has positive effects on recidivism, regardless of graduation status.

*Law and Human Behavior, 40(2), 118-27.*
Introduction

In the mid-1990s, Mental Health Courts (MHCs) were developed in response to the rising number of adults with mental illnesses in the criminal justice system (Almquist & Dodd, 2009). Briefly, MHCs are court-based diversion programs that aim to connect adults with mental illnesses to community-based mental health treatment services as an alternative to incarceration (Keator, Callahan, Steadman, & Vesselino, 2013; Steadman, Redlich, Callahan, Robbins, & Vesselino, 2011). Through court-based monitoring, MHCs aim to improve access to treatment for offenders with mental illness, facilitate better criminal justice outcomes, and improve use of resources by the criminal justice system (Almquist & Dodd, 2009). Although the specific structure and components differ across MHCs, the earliest MHCs had several defining features, including specialized dockets devoted to processing criminal cases for adults with mental illnesses, diversion of participants into community-based treatment services, and monitoring of adherence to treatment (Steadman, Davidson, & Brown, 2001). Within the last decade, MHCs have transitioned toward accepting more felony cases in addition to misdemeanor cases, admitting participants post-adjudication, using jail as a sanction for noncompliance, and increasing community supervision of participants (Redlich, Steadman, Monahan, Petrila, & Griffin, 2005). Since the first MHC was established in 1997, these courts have grown in prevalence across the U.S. and particularly in recent years. To demonstrate, as of 2009, there were over 250 MHCs in the U.S. (Almquist & Dodd, 2009), and as of 2013, that number increased by almost 40% to nearly 350 MHCs in 43 states (SAMHSA’s GAINS Center, 2015).
**Effectiveness of Mental Health Courts**

Many studies have investigated the effectiveness of MHCs in decreasing recidivism. Two early efforts with random assignment of participants to study conditions yielded mixed results. Specifically, Cosden, Ellens, Schnell, Yamini-Diouf, and Wolfe (2003) found no differences in time spent in jail and arrests 12 months following diversion between an MHC group \((n = 137)\) and Treatment As Usual (TAU) group \((n = 98)\). Compared to TAU participants, MHC participants were less likely to face new convictions, and a greater percentage of their convictions were for probation violations, likely due to increased monitoring by probation staff. A follow-up study of the same MHC looked at outcomes 24 months following entry into the Court (Cosden, Ellens, Schnell, & Yamini-Diouf, 2005). Results failed to show differences in jail days, arrests, and convictions between MHC participants \((n = 235)\) and TAU participants \((n = 98)\); however, both MHC and TAU participants received case management services, possibly increasing the likelihood of favorable outcomes for TAU participants.

Several observational and quasi-experimental studies also have examined the effectiveness of MHCs in reducing recidivism with favorable results. For instance, within-subjects evaluations show decreases in arrests in the 12 months pre- and post-entry (Herinckx, Swart, Ama, Dolezal, & King, 2005), as well as in the two years prior to entering an MHC and two years after exit (Hiday & Ray, 2010). Studies conducting between-subjects comparisons similarly show generally positive effects on recidivism, including reduced arrests (Hiday, Wales, & Ray, 2013; Moore & Hiday, 2006), decreased risk of new charges
(McNiel & Binder, 2007), and longer periods without new charges (McNiel & Binder, 2007). Finally, Steadman and colleagues (2011) conducted a multi-site, longitudinal study of four MHCs that found MHC participants \( n = 447 \) were less likely to be rearrested in the 18 months following entry into court than TAU participants \( n = 600 \). Together, findings of the extant research suggest that participation in MHCs decreases recidivism for justice-involved adults with mental illnesses, though effectiveness can vary as a function of characteristics of the participants and the program.

**Factors Associated with Effectiveness**

MHCs have limited capacity and can only adjudicate a select number of cases at any given time. As a result, MHCs rely on varying selection criteria, such as qualifying diagnosis or level and type of offense (Erickson, Campbell, & Lamberti, 2006; Wolff, Fabrikant, & Belenko, 2011), to target individuals who are most likely to benefit from MHC participation. Toward that end, recent research has examined characteristics of the participants that may be associated with effectiveness. Indeed, several studies have shown that MHC participants who fail to graduate from the MHC (Broner, Lang, & Behler, 2009; Herinckx et al., 2005; Moore & Hiday, 2006; Ray, 2014; Sarteschi, Vaughn, & Kim, 2011; Steadman et al., 2011), have a prior criminal history (Hiday et al., 2013; Reich, Picard-Fritsche, Lebrón, & Hahn, 2015; Rossman et al., 2012; Steadman et al., 2011), are younger (Dirks-Linhorst & Linhorst, 2012; Reich et al., 2015), and have co-occurring substance use (Burns et al., 2013; Cosden et al., 2005; Hiday et al., 2013; Reich et al., 2015; Steadman et al., 2011) may be less likely to experience reductions in recidivism associated with MHC participation. Homelessness
(Burns et al., 2013) has also been associated with increased recidivism following MHC participation. Finally, although mental health diagnoses have been investigated in relationship to recidivism among MHC participants, findings have been inconsistent (Comartin, Kubiak, Ray, Tillander, & Hanna, 2015; Steadman et al., 2011).

In addition to client characteristics, there may be important differences in programmatic elements of MHCs across settings that contribute to effectiveness (Wolff & Pogorzelski, 2005). To demonstrate, an MHC’s ability to exercise therapeutic jurisprudence has been operationalized and investigated with respect to recidivism in recent research. Redlich and Han (2014) found baseline measures of therapeutic jurisprudence (e.g., perceptions of the voluntariness of the court, perceptions of procedural justice, and knowledge of the MHC) to be associated with decreased arrests and jail time and fewer bench warrants 12 months following entry into an MHC. Another study found perceptions of procedural justice independently predicted likelihood of criminal activity; however, perception of negative pressures to coerce MHC participation was a stronger predictor of any criminal activity in the 12-month post-entry follow-up period (Pratt, Koerner, Alexander, Yanos, & Kopelovich, 2013). In contrast, Canada and Hiday (2014) found perceptions of procedural justice based on interactions with staff to be unrelated to short-term (i.e., 6 months) recidivism during MHC participation. Though some have suggested that quality of mental health treatment services may play a role in the effectiveness of MHCs (e.g., Sarteschi et al., 2011), existing evidence questions the association between increased use of mental health treatment services and reductions in recidivism among MHC participants.
(Keator et al., 2013). However, current research is limited by a dearth of in-depth investigations into associations between mental health service use and recidivism in the context of MHCs. Other programmatic elements, such as frequency of judicial supervision or use of sanctions for noncompliance (Callahan, Steadman, Tillman, & Vesselino, 2013; Erickson et al., 2006; Redlich et al., 2010), also may be associated with effectiveness but have yet to be investigated with respect to recidivism.

Beyond client and programmatic elements, characteristics related to the process of participating in the MHC, such as tenure in the MHC, time between referral and entry into the MHC, or time between referral and first receipt of services, may be associated with effectiveness. Process factors have not been defined explicitly in existing literature; however, for the purposes of this discussion, we define them as measures that describe the movement of participants through an MHC program. Such factors may be temporal in nature (e.g., time between two events), but not necessarily (e.g., source of referral to an MHC). In one study, Redlich, Liu, Steadman, Callahan, and Robbins (2012) examined differences in processing time (i.e., time between initial arrest and MHC entry or case disposition) between MHC participants and defendants experiencing traditional case processing, finding that MHC participants experienced faster case processing relative to defendants without mental illnesses but slower processing time relative to defendants with mental illnesses. However, to our knowledge, only one study has examined associations between such process measures and recidivism in the context of MHCs. Burns and colleagues (2013) found that longer tenure in the MHC was associated with decreased recidivism, though this association did not persist.
after accounting for graduation status. Research conducted in other criminal justice treatment settings has established associations between factors such as length of treatment—or treatment dosage—and recidivism. For example, in drug courts, longer participation has been linked to decreased rearrest one year following exit from the court (Peters, Haas, & Hunt, 2001). Longer treatment provided in correctional settings more generally, such as prisons, similarly has been associated with decrease risk of reincarceration, with greater reductions seen after each additional week of treatment (Bourgon & Armstrong, 2005).

**The Present Study**

More research is needed to both replicate findings of prior research on the effectiveness of MHCs and to further identify the ways in which MHCs can contribute to reductions in recidivism. Although several studies have investigated associations between individual characteristics and recidivism, the majority have examined recidivism post-entry, when participants are receiving treatment and support services through the MHC during a portion of the follow-up period. Only a handful of studies (Burns et al., 2013; Hiday & Ray, 2010; Hiday et al., 2013) have investigated these associations exclusively following exit from the MHC, arguably a higher risk period for recidivism. Moreover, the effects of certain characteristics, such as co-occurring substance use, on recidivism have been inconsistent. Additionally, as noted above, few studies have examined associations between process factors (e.g., time in the MHC program) and recidivism following exit from the MHC.

The present study will build upon previous research first, by comparing recidivism outcomes between MHC participants and a group of offenders with mental illnesses
experiencing treatment as usual (TAU), and second, by examining the impact of individual characteristics and process factors on changes in recidivism one year following exit from the MHC. Our specific aims are to determine whether: 1) there are significant differences in post-program charges, convictions, and jail days between MHC participants and a comparison group of offenders with mental illnesses; 2) individual factors (i.e., graduation status, presence of co-occurring substance use, and level of charge) are associated with changes in pre- and post-program jail days among MHC participants; and 3) process factors (i.e., time in program, time between referral and program entry, and time between referral and first receipt of mental health services) are associated with changes in pre- and post-program jail days among MHC participants. Consistent with previous investigations, we hypothesize that MHC participants will have fewer instances of recidivism relative to the comparison group. Additionally, individual factors of graduation from the MHC, absence of co-occurring substance use, and lower-level charge are hypothesized to predict greater reductions in recidivism from pre-program to post-program. Finally, we expect more efficient processing of MHC participants to be associated with greater reductions in recidivism from pre-program to post-program.

**Method**

**Study Setting**

Study data were collected from the Ramsey County Mental Health Court, established in 2005 in St. Paul, MN (see Guthmann, 2015). At the time of data collection, this MHC accepted participants with misdemeanor and gross-misdemeanor-level charges both pre-
adjudication and post-adjudication. To be eligible for participation in the MHC, defendants must have: been over 18 years of age, been charged with a misdemeanor or gross misdemeanor, had no history of violent offenses, and been diagnosed with a mental illness by a licensed mental health professional. All diagnostic information on MHC participants came from a state-level database of community mental health provider records. If a participant did not have a standing diagnosis, a diagnostic assessment was completed by a psychiatrist at the Ramsey County Mental Health Clinic prior to acceptance into the MHC. Eligible defendants were required to participate in the MHC for varying lengths of time that depended on their level of criminal charge at entry (one year for misdemeanors and two years for gross misdemeanors) as well as their progress throughout the program, which was based on compliance with terms of participation. Once admitted to the MHC, participants were assigned mental health case managers and prescribed individualized treatment plans. The mental health case manager was responsible for facilitating access to community-based treatment services, such as outpatient mental health treatment, mental health support groups, inpatient mental health treatment, and substance use or chemical dependency treatment. Mental health case managers additionally were responsible for overseeing compliance with conditions of participation. Throughout their participation in the MHC, participants were required to abstain from drug use, complete community work service, identify and maintain housing, and attend bimonthly court appearances. Failure to abide by these conditions resulted in graduated sanctions, including additional court appearances, increased supervision, and jail time during the program. Participants in the TAU condition whose cases
were adjudicated through traditional case processing would not have received referral to community-based services through the court.

**Data Collection**

The Institutional Review Board of the university and the Second Judicial District of the State of Minnesota approved this study. This study uses secondary, administrative data on all participants who exited the MHC during the first three years following its inception and a comparison group of offenders not participating in the MHC but who were adjudicated during the same timeframe and receiving TAU. To create the TAU group, court staff first reviewed all criminal cases within the last six months of the aforementioned 3-year period. This list was then narrowed to offenders who had misdemeanor or gross misdemeanor-level charges and who were convicted. From this pool, court staff randomly selected 400 cases and screened these cases to include offenders who self-reported a mental illness at intake and to exclude offenders who had previously been referred to or accepted into the MHC. This process resulted in a group of 56 offenders, of which 40 were randomly selected to form the TAU group.

**Sample**

The original sample comprised 98 offenders, including 58 participants of the MHC (30 graduates, 28 non-completers) who exited the program at least one year prior to the end of 2008 and 40 TAU participants who had committed offenses within the last six months of 2008. One MHC participant was dropped from analyses because no recidivism data were available (original $n = 58$; final $n = 57$). See Table 1 for demographic characteristics of each
group. Roughly half of participants were male and a greater proportion of participants identified as a non-white race. MHC participants were slightly, but not significantly, younger ($M = 34.48$, $SD = 9.62$) than participants in the TAU group ($M = 36.05$, $SD = 9.55$). Mental health diagnoses included primarily schizophrenia spectrum (schizophrenia, schizoaffective, and schizophreniform), mood (bipolar and major depressive), and anxiety (generalized anxiety, post-traumatic stress) disorders.

**Variables**

**Demographic characteristics and covariates.** Demographic variables included age (measured continuously), race (white, non-white), and gender (male, female). Mental health diagnosis could not be used as a covariate because individual-level data were not available for the comparison group. Bivariate comparisons conducted between MHC participants and comparison group participants on demographic characteristics were non-significant (all $ps \geq .439$). We conducted preliminary multivariate, between-groups (MHC vs. comparison) analyses using a cutoff value of $p < .10$ to identify potential covariates. Age ($B = -0.03$, $SE = 0.01$, $p = .052$, 95% CI [-0.06, 0.00]) and race ($B = -1.03$, $SE = 0.32$, $p = .001$, 95% CI [-1.67, -0.40]) predicted post-program charges; age ($B = -0.03$, $SE = 0.02$, $p = .083$, 95% CI [-0.07, 0.00]) predicted post-program convictions; and both age ($B = -0.04$, $SE = 0.01$, $p = .002$, 95% CI [-0.06, -0.01]) and gender ($B = 0.51$, $SE = 0.28$, $p = .072$, 95% CI [-0.05, 1.06]) predicted post-program jail days. Thus, all three variables were included as covariates in our propensity score analysis, as described below.

**Recidivism.** For comparisons between MHC participants and TAU group
participants, three count variables measured post-program recidivism: charges, convictions, and jail days. For MHC participants, these variables were measured one year after program exit, and for the TAU group, they were measured one year after case disposition. For within-subjects analyses, we examined number of pre- and post-program jail days served by MHC participants. Pre-program jail days were measured one year prior to acceptance into the MHC and post-program jail days were measured one year following exit from the MHC.

**Individual factors.** Graduation status (no, yes) was measured by whether or not a participant successfully completed the conditions of the MHC. Reasons for not graduating included opting out of the MHC ($n = 10$), being terminated from the MHC ($n = 14$) due to noncompliance with conditions, or being denied participation in the MHC after a provisional acceptance ($n = 3$). Co-occurring substance use (no, yes) was measured by whether or not a participant had a DSM-IV substance use disorder diagnosis at program entry or was currently using at the time of program entry (as indicated through self-report, clinician report, and/or urine drug screen); if either condition was met, participants were coded as positive for co-occurring substance use. Level of referral charge (misdemeanor, gross misdemeanor) was measured by the highest level of charge for which a participant was referred to the MHC.

**Process factors.** Time in program was measured as number of days from date of entry into the MHC to date of exit from the MHC. Time between referral and first receipt of mental health services was measured by number of days between first date of referral to the MHC and date of first receipt of any type of mental health services (e.g., appointment with community provider, appointment with case manager, etc.). Participants who were referred to
the MHC after their first receipt of mental health services were coded as having zero days between referral and first receipt of services. Finally, *time between referral and entry* into the MHC was measured by the number of days from first date of referral to the MHC to date of entry into the MHC.

**Analyses**

All analyses were conducted in SPSS 21. For Aim 1 between-groups comparisons, we first conducted a propensity score analysis using available covariates (i.e., age, gender, and race) to match the comparison group to the MHC group. Prior to the matching procedure, multiple imputation was performed using the SPSS Missing Values package to address missing values on age for *N* = 5 participants. Preliminary chi-squared analyses showed missingness on age to be unrelated to gender (*p* = .191) and race (*p* = .851), justifying the multiple imputation procedure, which was subsequently performed for *N* = 10 imputations. To conduct the propensity score analysis, we employed the MatchIt package in R (Ho, Imai, King, & Stuart, 2011) using full matching for each of the ten imputed datasets. Balance summary statistics for each propensity score analysis showed a reduction in mean differences between the two groups across all three covariates. The resulting ten sets of propensity scores were then averaged to create the final set of scores, per the method advocated for by Mitra and Reiter (2012).

Aim 1 analyses were conducted using generalized linear models and specifying a negative binomial distribution. Exploratory analyses to determine the appropriate model fit using negative binomial regression and maximum likelihood estimation for the dispersion
parameter. These analyses failed to yield a significant model fit \((p > .05\) for all likelihood ratios). We then compared Poisson regression models (dispersion parameter at 0) to negative binomial models with the dispersion parameter fixed at 1. For the majority of analyses, both models were significant; however, the negative binomial regression analyses produced models that were significant and had more conservative \(p\) values for coefficient estimates. Due to the sample size, we chose to proceed with the more conservative approach, using negative binomial regressions with the dispersion parameter fixed at 1.

For Aims 2 and 3 within-subjects analyses, we used the generalized estimating equation (GEE) with a negative binomial distribution to model changes in jail days (i.e., time, defined as pre-program to post-program) as a function of individual and process factors, as defined above (Liang & Zeger, 1986). Importantly, the GEE approach allows for correlations between repeated measurements of a dependent variable (Diggle, Heagerty, Liang, & Zeger, 2002), which is appropriate in the current study because prior criminal history is a robust predictor of future recidivism among adults with mental illnesses (e.g., Lovell, Gagliardi, & Peterson, 2002). A repeated subcommand was specified to allow for correlations between pre- and post-program jail days.

**Results**

**Descriptive Statistics**

MHC participants served an average of 9.16 \((SD = 16.06,\) range 0 to 88) days in jail in the year prior to entry into the MHC. Around half of MHC participants completed the MHC program (52.6\%, \(n = 30\)). Over two-thirds of MHC participants (67.3\%, \(n = 35\)) had a
current substance use disorder diagnosis or reported substance use at program entry. Time between referral to the MHC and first receipt of mental health services averaged 11.38 days ($SD = 28.15$, range 0 to 199). Time between referral to the MHC and entry into the MHC averaged 40.95 days ($SD = 36.08$, range 0 to 167). Time in program, or time between program entry and program exit, averaged 245.44 days ($SD = 122.39$, range 28 to 455).

Table 1 provides descriptive statistics of all post-program recidivism variables for MHC and TAU participants. Briefly, participants in the TAU group had greater numbers of post-program charges, convictions, and jail days relative to MHC participants, though non-graduates of the MHC reported greater numbers of charges and convictions relative to TAU participants.

**Aim 1**

In Aim 1, we investigated whether there were differences in post-program charges, convictions, and jail days between MHC participants and TAU participants. Negative binomial regression results for these analyses are presented in Table 2.

**Charges.** Results showed no effect of group on post-program charges (model fit: $X^2 = 13.14, p = .001$, AIC = 377.28, BIC = 384.98) after controlling for propensity score. Examination of MHC graduates and TAU participants, however, revealed a trending effect on post-program charges (model fit: $X^2 = 19.95, p < .001$, AIC = 241.21, BIC = 247.91). TAU participants incurred 1.81 times more charges in the 1-year post-disposition period compared to MHC graduates in the 1-year post-program period.

**Convictions.** Controlling for the propensity score, findings also showed no
significant effect of MHC participation on post-program convictions (model fit: \( \chi^2 = 7.33, p = .026, \text{AIC} = 243.09, \text{BIC} = 250.75 \)). However, examination of MHC graduates and comparison group participants revealed a significant effect of MHC graduation on post-program convictions (model fit: \( \chi^2 = 13.17, p = .001, \text{AIC} = 158.17, \text{BIC} = 164.87 \)). Specifically, TAU participants had 2.57 times more post-program convictions than MHC graduates.

**Jail days.** Results showed a significant effect of MHC participation on post-program jail days served (model fit: \( \chi^2 = 13.50, p = .001, \text{AIC} = 574.01, \text{BIC} = 581.73 \)), adjusting for propensity score. TAU participants incurred 1.62 times more jail days in the year following case disposition relative to MHC participants post-program exit. Examination of MHC graduates and comparison group participants also revealed a significant effect of MHC graduation on post-program jail days (\( \chi^2 = 38.70, p < .001, \text{AIC} = 370.14, \text{BIC} = 376.88 \)), showing TAU participants incurred 4.77 times more jail days relative to MHC graduates.

**Aim 2**

In Aim 2, we investigated whether individual factors were associated with changes in pre- and post-program jail days. Results of negative binomial regressions for these analyses are presented in Table 3.

**Graduation status.** Significant main effects of time (i.e., pre-program to post-program) and graduation status on overall jail days were observed, qualified by a significant interaction between time and graduation status (graduated from the MHC or not) (see Table 3). Specifically, graduates experienced a 7.69 times greater decrease in jail days served in the
year prior to MHC entry ($M = 10.77$, $SD = 3.82$) to the year following exit from the MHC ($M = 1.77$, $SD = 1.03$) compared to non-graduates, who actually increased in jail days served from pre-program ($M = 7.37$, $SD = 1.50$) to post-program ($M = 9.04$, $SD = 3.17$).

**Co-occurring substance use.** Results showed a significant main effect of time, controlling for the influence of substance use on overall jail days served (see Table 3). A significant interaction between substance use at program entry and time was observed. Participants with co-occurring substance use experienced a 4.76 greater decline in jail days served from pre-program ($M = 13.31$, $SD = 3.26$) to post-program ($M = 4.57$, $SD = 1.95$) compared to participants without a substance use disorder diagnosis or reported substance use at program entry, who experienced a slight increase in jail days from pre-program ($M = 2.29$, $SD = 0.68$) to post-program (3.76, $SD = 2.11$).

**Level of referral charge.** Results again revealed a significant main effect of time, controlling for level of referral charge (see Table 3). However, the interaction between level of referral charge (misdemeanor or gross misdemeanor) and time was not significant, $p = .262$. Specifically, participants with misdemeanor charges had an average of 7.87 ($SD = 2.17$) jail days in the year prior to MHC entry and 5.24 ($SD = 1.98$) jail days in the year following exit. Participants with gross misdemeanor charges had an average of 14.00 ($SD = 5.99$) jail days in the year prior to entry and 5.08 ($SD = 2.74$) days in jail following exit.

**Aim 3**

In Aim 3, we investigated whether process factors were associated with changes in pre- and post-program jail days. Results of negative binomial regressions for these analyses
are presented in Table 3.

**Time between referral and receipt of mental health services.** There was a significant main effect of time between referral and receipt of mental health services on overall jail days served ($p = .033$). Each additional day between referral to the MHC and first receipt of mental health services was associated with 1.03 times fewer overall jail days served. No significant interaction was observed between time from referral to the MHC and first receipt of mental health services and time (i.e., pre-program to post-program), $p = .553$.

**Time between referral and program entry.** No main effects of time between referral to the MHC and program entry on post-program jail days ($p = .667$) and time ($p = .135$) were observed, nor was the interaction between time from referral to the MHC program and entry into the MHC and time (i.e., pre-program to post-program) significant ($p = .473$).

**Time between program entry and program exit.** Results indicated a small but significant interaction between time in program and time (IRR = 1.005, $p = .045$). This interaction remained significant even after controlling for graduation status (IRR = 1.01, $p = .039$), suggesting that longer participation in the MHC was associated with greater decreases in jail days from pre-program to post-program regardless of whether or not a participant graduated from the MHC. Main effects of time in program ($p = .070$) and time were not significant ($p = .407$).
Discussion

MHCs have proliferated in numbers in recent years and so too has research examining their effectiveness in reducing recidivism (Sarteschi et al., 2011; Steadman et al., 2011). Recent investigations have focused on factors that contribute to MHC effectiveness (Burns et al., 2013; Steadman et al., 2011). However, additional research is needed to replicate prior findings and to increase our understanding of how MHCs are effective at decreasing recidivism, with particular attention to individual, programmatic, and process-related factors. The current study investigated the effectiveness of an MHC by first examining associations between MHC participation and post-exit measures of recidivism relative to a comparison group of participants experiencing treatment as usual (TAU). Second, we investigated whether individual factors (i.e., graduation status, co-occurring substance use, and referral charge level) and process factors (i.e., time in program, time between referral and program entry, and time between referral and first receipt of mental health services) influenced the association between MHC participation and changes in pre- and post-exit jail days. In the sections that follow, we review study findings and discuss how they relate to the broader literature on MHCs.

Summary of Findings

First, we examined differences in recidivism outcomes between a group of MHC participants and a comparison group of offenders receiving treatment as usual (TAU). We found that MHC participants had fewer jail days, but not charges or convictions, in the 1-year period following exit from the MHC relative to TAU offenders. MHC graduates, specifically,
had even fewer jail days and significantly fewer convictions relative to TAU participants. These findings are consistent with other quasi-experimental studies demonstrating some positive effect of MHC participation on recidivism (Christy, Poythress, Boothroyd, Petrila, & Mehra, 2005; Herinckx et al., 2005; Hiday & Ray, 2010; McNiel & Binder, 2007; Moore & Hiday, 2006; Steadman et al., 2011), but add to prior research investigating the effectiveness of the MHC when participants are no longer receiving services from the court.

To our knowledge, only two studies have compared recidivism as a function of MHC participation solely post-exit, excluding recidivism occurring during MHC participation, similarly showing that MHC participation was associated with fewer overall charges and rearrests relative to comparison groups (Hiday et al., 2013; McNiel & Binder, 2007). Importantly, these studies did not examine days incarcerated, a measure that may be more relevant to understanding the cost-effectiveness of MHC programs over time (McCollister, French, & Fang, 2010). Although the growing body of research suggests MHCs are effective, less is known about the extent to which MHCs represent an efficient use of criminal justice resources or the broader impact they have on the numbers of adults with mental illnesses involved in the criminal justice system (Steadman et al., 2009). A few studies have investigated the cost-effectiveness of MHCs, finding largely inconsistent results (Kubiak et al., 2015; Ridgely et al., 2007; Steadman et al., 2014); more research is needed.

Second, we explored associations between individual factors and changes in pre-program and post-program jail days among MHC participants. We found MHC graduation status to be associated with significantly greater reductions in jail days served prior to
program entry to following program exit, which is consistent with findings of previous studies (Broner et al., 2009; Herinckx et al., 2005; Steadman et al., 2011). Again, however, the present study adds to this body of work with its focus on recidivism exclusively following exit from the MHC when participants were no longer receiving the treatment supports and monitoring inherent to the MHC that may buffer against recidivism. Our findings suggest that long-term engagement with the MHC has a positive effect on criminal justice outcomes and that receiving the “full dose” of the MHC program is especially critical to recidivism reduction. Length of the MHC program itself has not been investigated systematically in the literature, but is an important avenue for future research.

Additionally, we found that MHC participants with co-morbid substance use experienced greater decreases in jail days from pre-program to post-program compared to participants without co-morbid substance use. This finding contradicts previous research showing decreased effectiveness of MHCs among participants with co-occurring substance use (e.g., Cosden et al., 2005; Steadman et al., 2011). However, differences between the current and prior findings may reflect the higher number of pre-program jail days for participants with co-occurring substance use in the current sample and our inclusion of only misdemeanor, as opposed to both misdemeanor and felony, offenders. Notwithstanding the uniqueness of the present sample, substance use is a significant risk factor for recidivism among adults with mental illness (Baillargeon et al., 2010; Castillo & Alarid, 2011; Wilson, Draine, Hadley, Metraux, & Evans, 2011) and our finding provides evidence that MHC monitoring and associated treatment can reduce recidivism risk in this subgroup. Future
research is warranted to replicate this finding in other contexts and to explore MHC program elements (e.g., support services used by participants, substance use testing, length of participation, etc.) that may account for such an effect.

Third, we examined associations between process factors and changes in pre-program and post-program jail days among MHC participants. We failed to identify significant interactions between two of the three process measures and criminal justice contact, namely, time between referral and program entry, and time between referral and first receipt of mental health services. Upon closer examination, almost half (49.1%, \( n = 28 \)) of MHC participants were already receiving mental health treatment services at the time of referral, resulting in somewhat limited variability in the time between referral and receipt of services variable. In contrast, there was greater variability in the time between referral and program entry variable; however, our null findings could reflect the average length of MHC participation and our use of a post-exit measure of recidivism, attenuating the effect of a timely referral on changes in recidivism. We did find a significant main effect of time between referral to the MHC and first receipt of recidivism on overall jail days served (including pre-program and post-program). This finding may reflect that higher-risk participants with more extensive criminal histories prior to MHC entry were provided with expedited access to mental health services; however, we could not support this explanation with available data.

Although we expected to find shorter time between referral and receipt of mental health services and shorter time between referral and program entry to be associated with
reductions in recidivism, our findings are not altogether inconsistent with the existing literature. A recent investigation by (Hiday, Ray, & Wales, 2014), for example, found time between index arrest and MHC entry, a comparable process measure, to be unrelated to whether participants completed the terms of participation in a short-term MHC; effects on recidivism were not examined. More research on the association between timing of referral to mental health services, receipt of treatment, and recidivism is needed; however, our results, taken into consideration with those of prior research, suggest that timely referral to the MHC may be less important to MHC effectiveness than other factors.

Finally, we found a significant interaction between time in program and changes in pre- and post-program jail days. Specifically, each day spent in the MHC was associated with a 1.01 times greater decrease in jail days served from pre-program to post-program, even after controlling for whether or not a participant graduated from the MHC. This result suggests that participants may benefit from a longer tenure in an MHC program regardless of graduation status, in contrast to prior research showing the relationship between time in program and recidivism did not persist after controlling for graduation status (Burns et al., 2013). Consistent with prior research on MHCs and the broader mental health treatment literature, benefits to longer tenure in the MHC may include ongoing case management services (e.g., Ventura, Cassel, Jacoby, & Huang, 1998), increased access to mental health treatment (e.g., Keator et al., 2013), and social relationships (Canada & Epperson, 2014; Skeem, Eno Louden, Manchak, Vidal, & Haddad, 2009). Findings also provide evidence against the practice of many MHCs to accept only participants who are likely to graduate
from the program (Wolff & Pogorzelski, 2005).

**Limitations**

Results must be considered in light of some study limitations. First, the present study used a quasi-experimental design with between-groups and within-subjects comparisons. It is possible that between-group differences are attributable to other factors beyond our control in the current study. For example, defendants may have been selected into the MHC based on likelihood of success, thereby biasing results in favor of the MHC condition. Additionally, self-reported diagnoses in the TAU group could not be verified through clinical interview or otherwise and were not provided to us at an individual-level. As a result, we were unable to include diagnoses as a possible covariate. Additionally, pre-program criminal justice data were not available for the TAU group, limiting our ability to determine whether participants in the TAU group already were at an increased risk for recidivism relative to MHC participants. Relatedly, although our follow-up period was uniform across participants, we could not account for time spent in a secure setting other than jail (e.g., hospital) during the follow-up period. Second, because this particular MHC operates a limited caseload (40 participants a year), our sample size was relatively small, which may have limited our ability to detect statistical significance for small effects in negative binomial models. Third, we investigated a single-site MHC, which limits generalizability of our findings, especially to MHCs with felony-level cases. Fourth, due to our reliance on secondary data, we were limited in the number and types of variables we could investigate. Additional individual and programmatic factors (e.g., failure to appear for court hearings, arrests during participation,
and noncompliance with MHC conditions) and process factors (e.g., time from index arrest to MHC entry, time from entry to Deferred Prosecution Agreement, and number of MHC hearings) that have been shown to be predictive of MHC graduation (Hiday et al., 2014) also may be relevant to recidivism, but were not available in the administrative records.

**Future Directions**

The current study supports the growing literature on the effectiveness of MHC programs in decreasing recidivism and also suggests directions for future research. Specifically, future research should explore whether certain high-risk groups, such as adults with mental illnesses and co-occurring substance use, benefit from time spent in an MHC setting, regardless of whether they complete the MHC program. Indeed, the positive effect of length of MHC participation on recidivism—regardless of graduation status—is arguably the most important finding of the present study and replication of this effect would imply a reconceptualization of what constitutes “successful” participation in the context of an MHC. More broadly, there is need for continued investigation into factors influencing the effectiveness of MHCs, such as service utilization or interactions with the MHC team. Whether MHCs can effectively address criminogenic risks and needs has been a focus of recent investigation (Campbell et al., 2015) and should be a focus in future research, particularly in the context of risk assessment and treatment planning to target risk factors predictive of reoffending (Andrews & Bonta, 2010; Honegger, 2015). Moreover, as evidence on factors influencing the effectiveness of MHCs grows, there will be a need for systematic investigation of effects across MHC sites. Finally, although investigations of the
effectiveness of MHCs are available and growing in the U.S., fewer investigations have taken place in countries outside of the U.S. (Richardson & McSherry, 2010; Slinger & Roesch, 2010). Research should explore whether the effectiveness of MHCs varies further on an international basis.

**Conclusion**

This study adds to the growing body of literature suggesting MHCs are effective, but also that effectiveness varies as a function of individual and MHC-specific factors. Because every community boasts varied services and a distinct population, the factors emphasized in the development of an MHC program inevitably will be different (Raines & Laws, 2008; Watson, Hanrahan, Luchins, & Lurigio, 2001). Moreover, case processing from referral to acceptance or rejection of an applicant is not uniform across MHCs, in part, because each MHC has diverse resources to devote to the processing of criminal offenders (Steadman, Redlich, Griffín, Petrila, & Monahan, 2005). Nevertheless, there is a need for continued research to identify ways to improve MHC effectiveness, given their widespread implementation across the U.S. and internationally. To the extent that evidence-based factors associated with reduced recidivism can be implemented, MHCs may better achieve the ultimate goal of reducing criminal justice contact among adults with mental illnesses.
Table 1. Demographic Characteristics and Recidivism for MHC and TAU Participants

<table>
<thead>
<tr>
<th>Variables</th>
<th>MHC participants</th>
<th>TAU participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall (N = 57)</td>
<td>Graduates (N = 30)</td>
</tr>
<tr>
<td></td>
<td>n %</td>
<td>n %</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>26</td>
<td>45.6</td>
</tr>
<tr>
<td>Female</td>
<td>31</td>
<td>54.4</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-White</td>
<td>37</td>
<td>64.9</td>
</tr>
<tr>
<td>White</td>
<td>21</td>
<td>35.1</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia spectrum</td>
<td>20</td>
<td>35.7</td>
</tr>
<tr>
<td>Mood</td>
<td>37</td>
<td>66.1</td>
</tr>
<tr>
<td>Anxiety</td>
<td>12</td>
<td>21.4</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Age</td>
<td>34.48</td>
<td>9.62</td>
</tr>
<tr>
<td>Post-program recidivism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charges</td>
<td>2.09</td>
<td>5.87</td>
</tr>
<tr>
<td>Convictions</td>
<td>0.73</td>
<td>1.98</td>
</tr>
<tr>
<td>Jail days</td>
<td>5.09</td>
<td>12.61</td>
</tr>
</tbody>
</table>

*Note. All percentages represent valid percentages. Discrepancies in cell sizes reflect missing data.*
Table 2. Group Comparisons between MHC and TAU Participants for Charges, Convictions, and Jail Days

<table>
<thead>
<tr>
<th>Predictors</th>
<th>MHC participants vs. TAU participants</th>
<th>MHC graduates vs. TAU participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N B SE B Wald X² IRR 95% CI</td>
<td>N B SE B Wald X² IRR 95% CI</td>
</tr>
<tr>
<td>Charges</td>
<td>96 -0.11 0.27 0.16 0.90 [0.53, 1.51]</td>
<td>69 0.59 0.33 3.15‡ 1.81 [0.94, 3.49]</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convictions</td>
<td>95 0.20 0.32 0.37 1.22 [0.65, 2.29]</td>
<td>69 0.94 0.44 4.57* 2.57 [1.08, 6.10]</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jail Days</td>
<td>97 0.48 0.22 4.68* 1.62 [1.05, 2.50]</td>
<td>70 1.56 0.28 30.29*** 4.77 [2.73, 8.31]</td>
</tr>
</tbody>
</table>

Note. IRR = incidence rate ratio. CI = confidence interval for incidence rate ratio. For group, TAU participants are coded as 0. ‡p < .10 *p < .05 **p < .01 ***p < .001
Table 3. Individual and Process Factors Predicting Changes in Jail Days among MHC Participants

<table>
<thead>
<tr>
<th>Predictors</th>
<th>N</th>
<th>B</th>
<th>SE B</th>
<th>Wald X²</th>
<th>IRR</th>
<th>Wald 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduation status</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>1.81</td>
<td>0.48</td>
<td>14.36***</td>
<td>6.09</td>
<td>2.39, 15.52</td>
<td></td>
</tr>
<tr>
<td>Graduation status</td>
<td>1.63</td>
<td>0.67</td>
<td>5.92*</td>
<td>5.11</td>
<td>1.37, 19.04</td>
<td></td>
</tr>
<tr>
<td>Graduation status x time</td>
<td>-2.01</td>
<td>0.59</td>
<td>11.42**</td>
<td>0.13</td>
<td>0.04, 0.43</td>
<td></td>
</tr>
<tr>
<td>Substance use</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>1.07</td>
<td>0.43</td>
<td>6.21*</td>
<td>2.91</td>
<td>1.26, 6.75</td>
<td></td>
</tr>
<tr>
<td>Substance use</td>
<td>-0.19</td>
<td>0.69</td>
<td>0.08</td>
<td>0.82</td>
<td>0.21, 3.17</td>
<td></td>
</tr>
<tr>
<td>Substance use x time</td>
<td>-1.56</td>
<td>0.67</td>
<td>5.52*</td>
<td>0.21</td>
<td>0.06, 0.77</td>
<td></td>
</tr>
<tr>
<td>Level of charge</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>1.01</td>
<td>0.31</td>
<td>10.81**</td>
<td>2.75</td>
<td>1.51, 5.04</td>
<td></td>
</tr>
<tr>
<td>Level of charge</td>
<td>0.03</td>
<td>0.64</td>
<td>&lt;0.01</td>
<td>1.03</td>
<td>0.30, 3.59</td>
<td></td>
</tr>
<tr>
<td>Level of charge x time</td>
<td>-0.61</td>
<td>0.54</td>
<td>1.26</td>
<td>0.54</td>
<td>0.19, 1.58</td>
<td></td>
</tr>
<tr>
<td><strong>Process factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time referral to MH services</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.52</td>
<td>0.40</td>
<td>1.69</td>
<td>1.68</td>
<td>0.77, 3.68</td>
<td></td>
</tr>
<tr>
<td>Referral to MH services</td>
<td>-0.03</td>
<td>0.01</td>
<td>4.54*</td>
<td>0.97</td>
<td>0.94, 0.988</td>
<td></td>
</tr>
<tr>
<td>Referral to MH services x time</td>
<td>0.01</td>
<td>0.02</td>
<td>0.35</td>
<td>1.01</td>
<td>0.98, 1.04</td>
<td></td>
</tr>
<tr>
<td>Time referral to program entry</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.82</td>
<td>0.55</td>
<td>2.23</td>
<td>2.27</td>
<td>0.77, 6.66</td>
<td></td>
</tr>
<tr>
<td>Referral to entry</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>0.18</td>
<td>1.00</td>
<td>0.99, 1.01</td>
<td></td>
</tr>
<tr>
<td>Referral to entry x time</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.52</td>
<td>0.99</td>
<td>0.98, 1.01</td>
<td></td>
</tr>
<tr>
<td>Time program entry to exit</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-0.54</td>
<td>0.65</td>
<td>0.70</td>
<td>0.58</td>
<td>0.16, 2.07</td>
<td></td>
</tr>
<tr>
<td>Program entry to exit</td>
<td>-0.01</td>
<td>&lt;0.01</td>
<td>3.23†</td>
<td>0.99</td>
<td>0.99, 1.00</td>
<td></td>
</tr>
<tr>
<td>Program entry to exit x time</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>4.27*</td>
<td>1.01</td>
<td>1.00, 1.01</td>
<td></td>
</tr>
<tr>
<td>Graduation status</td>
<td>0.19</td>
<td>0.38</td>
<td>0.26</td>
<td>1.21</td>
<td>0.58, 2.54</td>
<td></td>
</tr>
</tbody>
</table>

Note. IRR = incidence rate ratio. CI = confidence interval for incidence rate ratio. MH = mental health. For time, pre-program time is coded as 0. For graduation status, participants who did not graduate represent 0. For substance use, participants without co-occurring substance use represent 0. †p < .10 *p < .05 **p < .01 ***p < .001
CHAPTER 5

Integrative Review

Developed in response to the overrepresentation of adults with mental illnesses in U.S. jails (James & Glaze, 2006; Steadman et al., 2009), mental health diversion programs are now a widely used strategy to connect justice-involved adults with mental illnesses to community-based treatment in lieu of incarceration (Case et al., 2009; SAMHSA’s GAINS Center, 2015). However, prior research has not provided consistent support of their effectiveness in improving criminal justice outcomes and even suggests that effectiveness may vary widely across sites and studies. Alternative research strategies have the potential to improve the effectiveness of diversion strategies; however, the current literature is limited in several respects. Primarily, although modern theories of offender rehabilitation suggest assessing and targeting both risk and protective factors in treatment may improve criminal justice outcomes (Andrews & Bonta, 2010; Ward, 2002), few studies to date have examined the use of risk assessments in mental health diversion settings. Consequently, we know little regarding the relative importance of risk and protective factors in predicting recidivism in diversion participants. Secondarily, few studies have identified factors (e.g., participant characteristics, program components) associated with diversion program effectiveness, limiting our understanding of mechanisms accounting for the variable effectiveness of diversion programs. The manuscripts included in this dissertation address these limitations by examining the psychometric properties of risk assessment instruments conducted in this
setting, associations between risk and protective factors vis-à-vis recidivism, and participant- and process-related factors associated with recidivism outcomes.

**Summary of Findings**

**Chapter 2.** The first manuscript investigated the psychometric properties of risk assessments produced by two instruments, the START and LSI-R, in mental health diversion participants. Overall, findings showed good-to-excellent levels of inter-rater reliability for START and LSI-R assessments, adding to the limited research on the inter-rater reliability of assessments completed in correctional settings (Desmarais et al., 2016; Rocque & Plummer-Beale, 2014) and providing some of the first evidence of inter-rater reliability of risk assessments in a mental health diversion setting. Estimates were lower for the piloted START general offending risk estimate, which suggests the need for additional training around this estimate in future pilot studies before incorporation into the START instrument.

Findings showed good support for the predictive validity of LSI-R total scores and START vulnerability scores, both measuring risk factors, vis-à-vis arrests and jail days across most follow-ups. However, START strength scores, measuring protective factors, emerged as the most consistent and robust (i.e., produced the largest effect sizes) predictor across follow-ups. Justice-involved adults with mental illnesses experience a high rate of general risk factors for recidivism (Skeem et al., 2014), and findings from our investigation provide some of the first evidence suggesting protective factors have a role in predicting recidivism in this population. In particular, protective factors may help differentiate among those presenting with similar risk factors and guide treatment considerations (de Ruiter &
Nicholls, 2011). Moreover, although findings showed some evidence for racial bias in the predictive validity of risk assessments across shorter follow-ups, START strength scores showed no evidence of racial bias across any follow-up. Consideration of client strengths or protective factors may contribute to more accurate and unbiased assessments of reoffending risk, and in turn inform fairer risk management strategies (Grove & Meehl, 1996).

Chapter 3. Expanding on findings from the first manuscript, the second manuscript tested five hypothesized models describing associations between risk and protective factors operationalized using START vulnerability and strength total scores, respectively. For both arrests and jail days, risk and protective factors showed direct and independent associations with recidivism, as described in the Compensatory model. These findings are consistent with prior research, which has established direct and independent effects of risk and protective factors with respect to violent reconviction (de Vries Robbé et al., 2013; Ullrich & Coid, 2011) and general offending (Jones et al., 2015; Miller, 2015); however, this is the first study to test these associations in a mental health diversion program and in comparison to other models. There was limited supported for other models, including the Buffer, Challenge, Protective-Protective, and Mediation models. Although interactive effects of risk and protective factors as described in the Buffer model have been tested in previous investigations, our findings are not altogether inconsistent with the existing literature, which has failed to produce meaningful interaction effects on recidivism (de Vries Robbé et al., 2013; Jones et al., 2015; Ullrich & Coid, 2011).
Chapter 4. The third manuscript examined between-group differences in recidivism outcomes as a function of MHC participation and within-group changes in jail days as a function of participant- and process-related factors. Consistent with previous research (Frailing, 2010; Kubiak et al., 2015; Steadman et al., 2011), findings showed MHC participants had fewer jail days relative to offenders experiencing traditional criminal processing. Similar effects were not found for charges or convictions. Few studies have examined charge or conviction outcomes between MHC participants and offenders receiving treatment as usual, and studies to date have produced variable findings (Cosden et al., 2005; McNiel & Binder, 2007; Rossman et al., 2012). In keeping with prior studies (Broner et al., 2009; Herinckx et al., 2005; Steadman et al., 2011), graduation status was associated with greater reductions in offending (i.e., jail days) in MHC participants. However, in contrast to previous investigations (Cosden et al., 2005; Steadman et al., 2011), MHC participants with substance use at intake experienced greater reductions in jail days relative to those with a mental health diagnosis only. Although we investigated within-subject differences in recidivism as a function of several process-related factors, only time in program was associated with reductions in recidivism. Importantly, this effect persisted even after controlling for graduation status, providing evidence against the practice of MHCs to select participants who are likely to graduate (Wolff et al., 2011).

Implications for Research

Research on mental health diversion programs has grown substantially in the past decade; however, significant gaps remain. The present studies address key limitations and
have important implications for future research on mental health diversion programs. Primarily, findings suggest that both risk and protective factors have a role in the estimation of recidivism risk in justice-involved adults with mental illnesses. Despite the importance of protective factors as shown in present findings, knowledge of which protective factors are most relevant for this population remains limited. In contrast, prior research has identified specific risk factors for recidivism in this population, including substance use (Baillargeon et al., 2010; Castillo & Alarid, 2011; Wilson et al., 2011), lack of stable housing (Burns et al., 2013; Fischer et al., 2008), and prior criminal history (Hiday et al., 2013; Lovell et al., 2002; Steadman et al., 2011). To date, only treatment-relevant factors (e.g., medication adherence, outpatient treatment use) have been identified as potential contributors to reduced recidivism (Han & Redlich, 2015; Van Dorn et al., 2011, 2013) in this population. In justice-involved adults generally, several protective factors have been shown to reduce reoffending, including social and emotional support, time with family and friends, and closeness to others (Gilgun et al., 2000; Ullrich & Coid, 2011). Despite an emphasis on important relationships as foundational to mental health diversion program practice, such as the participant-caseworker (Canada & Epperson, 2014) and participant-judge (Boldt & Singer, 2006) relationships, there have been few—if any—evaluations of the role of social support or social relationships as a protective factor in diversion participants or justice-involved adults with mental illnesses more broadly. Given high rates of reoffending in this population and the promise of protective factors in estimating reoffending risk, future research is needed to identify protective factors relevant to the risk management of this population.
Further complicating limited research on protective factors in this population, there have been few evaluations in any adult offender population on how specific risk factors relate to protective factors. Our findings overall suggest that risk and protective factors, as uniform constructs, operate independently in the prediction of reoffending. Yet, whether specific protective factors may mitigate the effect of specific risk factors on recidivism remains to be seen. To date, there has been only one investigation in an adult offender population that examined the interactive effects of distinct protective factors (i.e., time with family and friends, social support, emotional support, religion, closeness with others) on risk factors; however, this study employed a measure of cumulative risk and failed to find evidence of interactive contributions (Ullrich & Coid, 2011). In contrast, there have been a handful of investigations in adolescent samples on the independent and interactive effects of specific risk and protective factors on adverse outcomes including physical aggression (Aceves & Cookston, 2006; Farrell et al., 2010; Hollister-Wagner et al., 2001), problem behavior (Richard Jessor et al., 1995), and offending (Salekin et al., 2010). Intuitively, protective factors like social support should offer protection against common criminogenic risk factors, like lack of financial resources or lack of stable housing; however, no study to date has examined these associations in a justice-involved sample, and future research is needed.

Beyond the identification of risk and protective factors in justice-involved adults with mental illnesses, our findings warrant investigation into the relevance of competing models of offender rehabilitation for this population. Specifically the Risk-Need-Responsivity (RNR; Andrews & Bonta, 2010; Andrews, Bonta, et al., 1990) model emphasizes the assessment
and targeting of criminogenic risks and needs as part of offender rehabilitation. In contrast, the Good Lives Model (GLM; Ward, 2002) posits that risk management strategies must additionally develop offender strengths and resources for successful rehabilitation. Given the robust predictive capacity of protective factors in the current studies, future investigations should investigate the utility of the GLM for mental health diversion clients. Although many strengths emphasized in the GLM (e.g., knowledge, excellence in play, excellence in work, excellence in agency, emotion, relationships, social support) are captured in risk assessment instruments measuring protective factors (e.g., STRONG; Barnoski & Drake, 2007; SPROF; de Vogel et al., 2011; SPIIn; Orbis Partners, 2003; START; Webster et al., 2009), the present studies are among the first to explore the predictive utility of protective factors in a diversion sample. Validation of risk assessment instruments, particularly those measuring protective factors, in mental health diversion settings is a critical first step toward informing the relative importance of the RNR model and the GLM in the risk management of this population.

With respect to research on the effectiveness of diversion programs, particularly MHCs, our findings are novel in two regards. First, they suggest that MHCs have the potential to improve outcomes in offenders with co-occurring substance use and mental illness, despite presence of co-occurring substance use being implicated as a risk factor for recidivism in prior research (Dirks-Linhorst & Linhorst, 2012; Steadman et al., 2011). Second, findings suggest that ongoing engagement in the MHC may improve outcomes for participants, even in the absence of successful program completion. Together, both findings
underscore the need for research on the unique risks and needs of diversion program participants, the extent to which those risks and needs are addressed in case management plans, and strategies to increase participant engagement.

Risk management strategies borrowed from the RNR model and the GLM may assist in these goals. Specifically, consistent with the need and responsivity principles of the RNR framework, assessing individual needs and delivering more individually tailored services may improve the effectiveness of diversion programs (Campbell et al., 2015; Hubbard, 2007). Further, consistent with the GLM, consideration of participant strengths may improve the responsiveness of case management strategies (Ward, Melser, & Yates, 2007) and increase engagement in diversion programming by promoting a therapeutic alliance between participants and case managers (Saleebey, 1996). However, only one study to date has examined adherence to these strategies in a mental health diversion program (Campbell et al., 2015), and more work is needed.

Our findings also provide further evidence on the variable effectiveness of MHCs. Specifically, we found significant effects of MHC participation on only one (i.e., jail days) of three recidivism outcomes. Moreover, only one process-related factor (i.e., time in program) was associated with better recidivism outcomes. Indeed, the substantial variability in the effectiveness of diversion strategies, particularly MHCs, reflects our limited understanding of the potential mechanisms of these programs. Although prior research has established variation in the structure (Redlich, Steadman, Monahan, Robbins, & Petrila, 2006) case processing (Steadman et al., 2005), and selection of participants into MHCs (Wolff et al.,
2011), little is known regarding the case management practices, community-based service referrals, or features of MHC status hearings that constitute MHC participation and may contribute to variable effectiveness. Together with the extant literature, our findings highlight the need for future investigation into sources of variability in the design and practice of MHCs.

**Implications for Practice**

Findings have several implications for criminal justice practice, particularly in the context of mental health diversion programs. Primarily, our results suggest risk assessment instruments have predictive utility and may inform supervision decisions in the context of diversion programming by predicting recidivism risk. However, findings also suggest consideration of risk factors has the potential to be biased against racial minorities, particularly in the short-term. Indeed, critics have raised concerns about the use of actuarial risk assessment instruments to inform sentencing and supervision, suggesting they may contribute to racially biased decisions (Harcourt, 2015; Holder, 2014; Starr, 2014). As such, practitioners employing risk assessments in this setting, particularly in racially diverse communities, should have an awareness of their potential for overclassification of racial minorities.

Beyond risk factors, our findings suggest that consideration of protective factors may result in more accurate assessments of recidivism risk in this population. Further, assessment of protective factors may be less racially biased relative to risk factors and contribute to fairer risk management decisions. There is opportunity for interventions developed in this setting to
assist in the promotion of protective factors such as social skills, prosocial attitudes, or social support, pending further research. Minimally, the current state of the science supports assessing both risk and protective factors and devising individually tailored treatment plans as part of a comprehensive risk management strategy.

Looking beyond a risk-only approach, findings have implications for the selection of participants into mental health diversion programs. MHCs, in particular, have been criticized for selection processes that reduce the pool of eligible defendants to those who are most likely to successfully complete the program (Sarteschi et al., 2011). However, findings suggest a reconsideration of selection processes into MHCs, including whether defendants judged to be at high risk may benefit from MHC participation, even in the absence of a successful outcome. Compared to limiting diversion participation to a select few, strategies to assess high-risk participants, provide tailored services, and increase participant engagement and length of participation may prove more productive in reducing high rates of recidivism in this population.

Implications for Policy

On a policy level, findings may inform both funding prioritization for diversion program practice as well as updated recommendations on best practices in mental health diversion. Primarily, there are opportunities for federal agencies, particularly those that provide technical assistance (e.g., the Substance Abuse and Mental Health Services Administration), to fund strategies that may increase the uniformity of assessment procedures and case management planning in these settings. Although such strategies have been funded
as part of broader grant solicitations to promote collaborations between criminal justice agencies and community-based providers (e.g., Behavioral Health Treatment Court Collaborative), these efforts have been small in scope and not targeted specifically to mental health diversion programs.

Further, although many mental health diversion programs were originally federally funded, programs today are funded primarily through either state or local agencies, as well as operational grants (Case et al., 2009; Goodale, Callahan, & Steadman, 2013). With the expiration of federal funding and current uncertainty of Medicaid expansion and the Affordable Care Act, it is likely that funding sources will continue to diversify, resulting in greater variability in the structure, operation, and key services of mental health diversion programs. As such, there is need for a national consensus on key components of mental health diversion programs or, minimally, key strategies to improve participant outcomes. Current recommendations are based primarily on observations of early mental health diversion schemes (Steadman et al., 1995; Thompson et al., 2008) and do not account for growing diversification of diversion programs or updated research on the multidimensional needs of justice-involved adults with mental illnesses (Wolff et al., 2013). Some states have developed comprehensive state-wide guidelines for the operation of mental health diversion programs (e.g., MHCs in Oregon), which may serve as a starting point for building a national consensus on best practices in this setting.
Limitations and Future Directions

In addition to those mentioned in each chapter, the included studies have several overarching limitations that provide directions for future research on mental health diversion programs. Specifically, although findings herein suggest risk assessments have utility in predicting recidivism in diversion program clients, they do not speak to whether such assessments can be implemented in a diversion setting to improve outcomes. The extent to which assessments can be implemented—and risk and protective factors targeted—in practice is a separate inquiry and an important avenue for future research. Similarly, risk assessments were conducted at a single time point and, as a result, did not account for changes in dynamic risk and protective factors during diversion participation. This may have attenuated the predictive accuracy of assessments given that diversion programs, by their very name, are designed to divert adults to community-based services to address criminogenic risks and needs. Prior investigations have examined changes in assessments over time in other justice-involved samples (e.g., de Vries Robbé, de Vogel, Douglas, & Nijman, 2015; Labrecque et al., 2014; Schlager & Pacheco, 2011), providing evidence for the dynamic nature of certain criminogenic risks and needs. Determining whether risk assessments can capture changes in risk and protective factors and maintain their predictive accuracy in the context of diversion programming is a direction for future research.

Finally, all three studies were single-site investigations that used data drawn from distinct jurisdictions. Given the wide variability in diversion programs (Redlich et al., 2006; Steadman et al., 1995), findings may have limited generalizability to other sites. The
potential for disparate findings across studies in mental health diversion program research has been confirmed in previous systematic reviews on the effectiveness of mental health diversion programs (Honegger, 2015; Sirotich, 2009). Future research may benefit from consideration of sample characteristics, program components, and study quality that together may contribute to variability in findings. To date, one meta-analytic investigation has been conducted on mental health diversion programs, specifically MHCs (Sarteschi et al., 2011). Although authors reported significant heterogeneity in the effect of MHCs on recidivism, this investigation did not examine factors that may explain variability in findings (e.g., potential study-level moderators).

Research on mental health diversion programs, especially MHCs, has grown considerably in recent years, and there is need for sophisticated research strategies to identify sources of variability in study findings. Minimally, single-site studies should report effect sizes as a function of individual characteristics and program-related factors. Cumulatively, consistent reporting of these characteristics will allow for more sophisticated meta-analytic investigations (e.g., meta-regression) or systematic reviews (e.g., realist synthesis approach; Rycroft-Malone et al., 2012) to identify program components or other mechanisms accounting for the effectiveness of diversion programs. Investigation into these characteristics in future studies may increase our understanding of how diversion programs work, for whom they work, and how they can potentially work better.
Conclusion

Mental health diversion programs were designed to address the overrepresentation of justice-involved adults with mental illnesses in U.S. jails through diversion into community-based treatment services. Although research on such programs has proliferated in recent decades, findings have not provided consistent support for their effectiveness in improving criminal justice outcomes. Rather, effectiveness has been shown to vary considerably across sites and studies. Given the multidimensional risks and needs of justice-involved adults with mental illnesses, risk management strategies may play a role in improving recidivism outcomes in this population. However, few studies to date have examined the use of risk assessments in this setting or investigated the relative importance of risk and protective factors in predicting recidivism. Additionally, there has been limited investigation of potential mechanisms accounting for the effectiveness of diversion programs. To these ends, the manuscripts included in this dissertation examined the psychometric properties of risk assessment instruments conducted in a mental health diversion setting, associations between risk and protective factors vis-à-vis recidivism, and participant- and process-related factors associated with recidivism outcomes. Findings offer support for the predictive utility of risk assessments administered in this setting and suggest that protective factors may be more relevant to the risk management of this population relative to risk factors. Further, findings support the effectiveness of longer-term engagement in mental health diversion programs and suggest that high-risk participants, such as those with co-occurring disorders, may benefit from participation in structured diversion programs, like MHCs. Future research is needed to
inform the development of risk management strategies, including strengths-based approaches, that may increase engagement in diversion services and ultimately improve criminal justice outcomes among mental health diversion participants.
REFERENCES


behavior. *Child Development, 74*(1), 109–126. https://doi.org/10.1111/1467-8624.t01-1-00524


Lowenkamp, C. T., Lovins, B., & Latessa, E. J. (2009). Validating the Level of Service Inventory—Revised and the Level of Service Inventory: Screening Version with a

https://doi.org/10.1177/0032855509334755


1403. https://doi.org/10.1176/appi.ajp.2007.06101664

homelessness, mental disorder, and co-occurring substance abuse. *Psychiatric
https://doi.org/10.1176/appi.ps.56.7.840

sample of pre-release general offenders. *Behavioral Sciences & the Law, 24*(6), 767–
782. https://doi.org/10.1002/bsl.728

Professional manual*. Odessa, FL: Psychological Assessment Resources.

https://doi.org/10.1177/1079063214564389

scores after multiple imputation. *Statistical Methods in Medical Research.*
https://doi.org/10.1177/0962280212445945

https://doi.org/10.1007/s10979-007-9087-7


https://doi.org/10.1080/14999013.2013.791351


https://doi.org/10.1016/j.ijlp.2014.02.017

https://doi.org/10.1037/lhb0000041


