

A meta-analysis of substance misuse intervention programs offered to women offenders

by

Chealsea De Moor

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial
fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

Carleton University
Ottawa, Ontario

©2018
Chealsea De Moor

Abstract

This meta-analytic review examined the effectiveness of substance misuse interventions in reducing recidivism and substance use outcomes among women offenders. A literature search revealed 22 evaluations, reporting 39 effect sizes. The effect estimate for recidivism outcomes revealed 53% to 79% reductions in the odds of recidivism for women participating in interventions. Effect estimates were similar for outcomes relating to substance use, with reductions in the odds of alcohol and drug use ranging from 13% to 82%. Across evaluations reporting recidivism outcomes, gender-neutral and gender-informed interventions were equally effective in reducing recidivism outcomes. Further, sub-group analyses revealed that study quality did not impact reductions in recidivism. Overall, this review lends support to the findings of previous research, suggesting that substance misuse interventions are effective in reducing both recidivism and substance use outcomes among women offenders.

Acknowledgements

This work would not have been possible without the financial support of the Joseph Armand Bombardier Canada Graduate Scholarship awarded by the Social Sciences and Humanities Research Council of Canada.

I would like to express my deepest appreciation to those I have had the pleasure of working with throughout the duration of this project. Each of the members of my thesis Committee has provided me with guidance, and support, and has encouraged me to expand my knowledge and expertise as a forensic psychology researcher. I would especially like to thank my supervisor, Dr. John Weekes. As my mentor, he has provided endless support and encouragement that reaches far beyond the bounds of any Master's thesis. He exceeds any expectations as a mentor and supervisor, and for that I am very grateful.

I would also like to thank Dr. Shelley Brown, my co-supervisor, for providing expert knowledge on issues surrounding women offenders. Your feedback throughout this project has proven to be invaluable. Furthermore, I would like to extend my appreciation to Dr. Leslie Anne Keown, who played a fundamental role in my understanding of statistics, and the complexities of working with applied data.

Lastly, nobody has been more important to me in the pursuit of this project than the members of my family. I would especially like to thank both my parents, for helping me to maintain motivation in times of stress. Their love and endless support remains the driving force to my success, and will continue to motivate me in the next steps of my life.

Table of contents

Title Page	i
Abstract	ii
Acknowledgments	iii
List of Appendices	vi
List of Tables	vii
List of Figures	viii
Introduction.....	9
Substance use among women offenders.....	11
Gender-neutral approach to intervention.....	14
Gender-responsive approach to intervention.....	16
Effectiveness of current intervention programs	21
Cognitive-behavioural therapy.....	22
Therapeutic communities	23
Contingency management.....	23
Drug treatment courts.....	24
Pharmacological interventions	24
Methadone.....	25
Methods	30
Meta-analysis	30
Literature Search	34
Procedure.....	36
Study Descriptors	37
Study Quality	38
Inter-rater reliability	
Aggregation of Data	41
Results.....	44
Description of eligible studies	44
Effect of intervention participation on recidivism.....	47
Subgroup-analyses for recidivism outcomes.....	50
Effect of intervention participation on drug use.....	52
Effect of intervention participation on alcohol use	53
Study Quality	54
Discussion.....	58
Strengths and Limitations.....	66

Conclusion 70

References..... 72

List of Appendices

Appendix A: Data Extraction Form 82

List of Tables

Table 1. Summary of studies and outcomes	45
Table 2. Effect sizes according to outcome	48
Table 3. Odds-ratios and confidence intervals of evaluations presenting recidivism outcomes ...	49
Table 4. Sub-group analyses for recidivism outcomes	51
Table 5 Odds-ratios and confidence intervals of evaluations presenting drug use outcomes	54
Table 6. Odds-ratios and confidence intervals of evaluations presenting alcohol use outcomes .	54
Table 7. Study quality comparison: Recidivism outcomes.....	57

List of Figures

Figure 1. Flow diagram of literature search process..... 36

Introduction

The problematic use of alcohol and other drugs represents a major public health issue with multiple social and economic implications. In 2012, the World Health Organization presented statistics highlighting the prevalence of substance misuse at a global level. Indeed, it is estimated that between 15.6 million and 38.6 million individuals actively engage in problematic substance use worldwide (United Nations Office on Drugs and Crime, 2012). Thus, it is evident that substance misuse is an important issue of concern among the general population – at a global level. What is even more concerning, however, is the high prevalence of serious substance misuse disorders among marginalized populations, such as those involved in the criminal justice system.

A comprehensive report conducted by The National Center on Addiction and Substance Abuse at Columbia University (2010) examining the extent to which alcohol and other drugs are associated with crime and incarceration in the United States, data analysis revealed alarming statistics regarding the prevalence of substance use disorders among offender populations. Rates of substance use among federal, state, and local level systems was 40%, with only 27% of the entire offender population free of substance use disorders (National Center on Addiction Substance Abuse, 2010).

It is the case that male and female offender populations demonstrate similar rates of substance use. Indeed, nearly 80% of Canadian male and female offenders demonstrate diagnosable issues with substance use (Grant & Gileno, 2008; Weekes, Mosher, Ternes, & 2009). Furthermore, these rates are replicated among American men and women offenders, with approximately 60% of women offenders and 70% of men

offenders demonstrating problematic use of alcohol and other drugs (Dorsey & Middleton, 2005; The Sentencing Project, 2007). Despite almost equal rates of substance misuse among male and female offender populations, research indicates that substance misuse may differentially influence future engagement in criminal behaviour based on gender. (Gehring, & Van Voorhis, 2014; Scott, Grella, Denis, & Funk, 2014; Van Voorhis, Wright, Salisbury, & Bauman, 2010).

In addition to presenting a major global health challenge, the misuse of alcohol and other drugs represents an area of ongoing concern among women offender populations. This concern has been reflected within criminal justice systems worldwide, with many correctional institutions offering a variety of substance misuse treatment programs to women offenders. Meta-analyses have been conducted evaluating the efficacy of correctional interventions offered to women involved in the criminal justice system, however, none have focused solely on substance misuse interventions. Given such high rates of substance misuse among this population, it is important that the efficacy of these programs be explored, to ensure that they are in fact, effective in reducing both recidivism and substance misuse outcomes.

This review will begin by illustrating the prevalence of the misuse of alcohol and other drugs among female offender populations. Through the application of theory, the drug-crime nexus will be explored to highlight the importance of recognizing and addressing the problematic use of alcohol and other drugs among incarcerated women. Gender differences in relation to substance misuse and its effect on criminal behaviour will also be explored. Characteristics of effective correctional programming will be discussed and applied to intervention programs currently offered to female offenders to

provide an illustration of these intervention practices. This critical review will provide the foundation for a meta-analytic study of substance use intervention programs offered to women offenders at an international level.

Substance use among women offenders

Women are among one of the fastest growing segments of the criminal justice population. Indeed, research reflects a steady increase in the rates of incarceration for women offenders over the last 20 years, going above and beyond that of their male counterparts (Covington & Bloom, 2007; Harrison & Beck, 2003; Wright, Van Voorhis, Salisbury, & Bauman, 2012). This steady increase in the rates of incarceration for women offenders has been partially attributed to high rates of substance use among women offenders and the implementation of policies that criminalize substance use (Tripodi, Bledsoe, Kim, & Bender, 2011).

Despite the fact that prevalence rates for substance use appear to be similar among men and women offenders, more close examinations of the patterns and severity of substance use highlight the unique differences that exist between men and women offenders who engage in substance use (Sacks, 2004). Certainly, it is the case that incarcerated women often report higher instances of substance use at the time of their offence, as well as higher rates of substance use as defined by structured assessments. (Greenfeld & Snell, 1999).

Peters, Strozier, Murrin, and Kearns (1997) were some of the first to examine gender differences in the use of substances among men and women offenders. Using a sample of 435 female and 1220 male offenders, the authors compared self-reported responses on the Addiction Severity Index, a standardized instrument providing

information regarding the psychosocial functioning of those engaging in substance use (McLellan et al., 1992). The results demonstrated that women were substantially more impaired than men when committing their offence. In addition, women reported more recent, frequent, and chronic substance misuse, and had a tendency to prefer harder drugs, such as cocaine (Peters, Strozier, Murrin, & Kearns, 1997).

Statistics on the prevalence of substance use in Canadian federal correctional institutions highlight similar patterns of use among women, where approximately 77% of women offenders display some degree of problematic substance use. Furthermore, of those women exhibiting problematic use, just over half displayed moderate to severe problematic substance use (Kelly, & Farrell MacDonald, 2015). This is in contrast to men offenders, where approximately 70% would be considered to have problematic substance use, with just over one-third of male offenders displaying moderate to severe problematic use (Kelly & Farrell MacDonald, 2015).

In addition to gender differences in the patterns and severity of substance use among offender populations, women seem to engage in substance use for reasons different than that of their male counterparts. It is frequently the case that women who engage in problematic substance use report using substances to alleviate psychological pain. Survey data obtained from 350 male and female patients registered in chemical dependency centers in the US highlighted the presence of self-medication among dependent users. Responses indicated that just over half of the sample described “escape from negative affective states” as the primary reason for initialing substance use, with a greater number of women endorsing these items compared to men (Corcoran & Corcoran, 2001). This relationship was partially attributed to high rates of co-occurring

mental health and substance use disorders, as indicated by responses to items such as “it helped me feel less depressed,” and “it helped me worry less” (Corcoran & Corcoran, 2001).

Indeed, co-morbidity of mental health disorder and substance use disorder are highly prevalent among women offender populations. A systematic review of the literature examining co-occurring substance use and mental health disorders among women offenders reported that women often come into contact with the criminal justice system through crimes related to their substance use. Furthermore, it is typically upon entry into the criminal justice system that the majority of women receive a formal diagnosis for any co-occurring disorders (Sacks, 2004).

An Australian study illustrated the prevalence of co-occurring substance use and mental health disorders among a sample of 288 arrestees, 222 men and 66 women (Heffernan, Finn, Saunders, & Byrne, 2003). Results indicated that 85% of female arrestees were diagnosed with at least one substance use disorder, while 94% evidenced significant psychological distress. This is in contrast to men offenders, who demonstrated similar rates of substance misuse (86%), however, they did not evidence the same degree of psychological distress (82%). Women arrestees were more likely than men arrestees to reach the minimum cut-off score for psychological impairment on indicators of psychological health (Heffernan et al., 2003).

Rates of co-morbid mental health and substance use disorders among women offender populations lends credibility to the notion that women offenders may engage in substance use as a form of self-medication, with the goal of easing the psychological distress resulting from a co-occurring mental health disorder. Therefore, it is of great

importance to acknowledge the interaction between substance use and mental health disorders when developing and delivering substance misuse interventions for women offenders (Morrissey et al., 2005).

Gender-neutral approach to intervention

Historically, intervention programs offered to women offenders were founded on male models of change and the principles of effective intervention as outlined by the Risk-Need-Responsivity model. The RNR model addresses several key clinical issues including the appropriate target clientele for intervention, the critical targets of intervention that produce the greatest reduction in criminal behaviour, and the style, modes and strategies of treatment delivery best suited for the target clientele (Andrews & Bonta, 2010). Each clinical issue is reflected in one of the three core principles of effective intervention: the risk principle, the criminogenic need principle and the general responsivity principle.

The first core principle of the RNR model, the Risk Principle, addresses two important aspects of effective intervention among offenders. The first is that criminal behaviour can be predicted. The second aspect involves the idea of matching the levels of treatment services to the risk of the offender. For example, offenders who demonstrate higher risk for reoffending require more intensive treatment for significant reductions in recidivism, while those demonstrating low-risk levels for reoffending require minimal or even no intervention (Andrews & Bonta, 2010; Andrews & Bonta, 2017). Despite being a crucial element in the effective intervention of substance use, theory does not always translate into practice. It is the case that some human service workers prefer to work with low risk clients as opposed to high risk clients, as low-risk client are often more

motivated to change and their use is generally easier to manage (Andrews & Bonta, 2010).

The need principle is the second key principle in the RNR model and describes the targets of change required to successfully reduce recidivism. The need principle distinguishes between two types of needs, criminogenic and noncriminogenic (Andrews & Bonta, 1990; Andrews & Bonta 2010). Criminogenic needs are identified as dynamic needs sharing a causal relationship with offending, that when altered through intervention, reduce the probability of recidivism. Noncriminogenic needs are those that do not share a causal relationship with offending (i.e., self-esteem, or a major mental health disorder; Andrews & Bonta, 1990; Andrews & Bonta 2010). Andrews and Bonta (1990; 2010) identified seven criminogenic needs unique to offender populations: procriminal attitudes, antisocial personality, procriminal associates, social achievement, family/marital, substance misuse and leisure/recreation.

The third and final core principle in the RNR model is the Responsivity Principle, which highlights the importance of delivering treatment programming in a style and manner that reflects the abilities and learning styles, orientation and outlook of the target clientele (Andrews & Bonta, 2010). This principle can refer to responsivity in the general sense, which includes the social-learning and cognitive-behavioural approaches to intervention (Andrews & Bonta, 2010). In addition, responsivity can also be specific to each individual and refers to the importance of considering the unique learning style and cognitive abilities of each offender seeking treatment. Accordingly, interventions derived for a group therapy setting may not be appropriate for clients with high levels of social anxiety or challenges with interpersonal sensitivity (Andrews & Bonta, 2010).

This approach is considered to be gender-neutral, meaning that risk factors identified as criminogenic apply to men and women equally, and the strength of their relationship with crime does not differ based on gender (Andrews & Bonta, 1990; Andrews & Bonta 2010).). Research has consistently identified this approach to correctional interventions as the gold standard, resulting in its widespread use among men and women offenders. A point of contention however, is the fact that the majority of research evaluating this approach has focused almost exclusively on male offender samples.

Gender-responsive approach to intervention

The practice of adapting male models of intervention for use among women offenders has received criticism, as the critics (e.g., Bloom, Owen, & Covington, 2003; Hall, Golder, Conley, & Sawning, 2013) argue that this approach does not appropriately acknowledge that women and men differ with respect to their initial risk factors for engaging in criminal behaviour, the ideal content and delivery of the intervention, and the unique barriers women offenders face with regards to community reintegration

Certainly, it is the case that women's pathways into the criminal justice system appear vastly different from that of men. Both qualitative and quantitative approaches have been used to examine women's unique pathways into offending. Daly (1992) was the first to qualitatively describe these pathways. Using presentence investigation reports for 40 women, Daly constructed statistical profiles and biographies by coding information surrounding the offence (i.e., offence type, victim's and defendant's version of the offence), and personal characteristics (i.e., social history, prior record, mental health, substance use) for each of the women in the sample. Five distinct pathways

emerged based on patterns of shared characteristics across women; harmed-and-harming women, battered women, street women, drug-connected women, and other women (Daly, 1992).

The first pathway, harmed-and-harming women, experienced high rates of abuse during childhood, and often have histories in juvenile court. Similarly, those following the pathway of battered women also experienced abuse, but in the context of an intimate relationship with a violent partner (Daly, 1992). The street women pathway describes women who left abusive households during childhood, became dependent on substances, and subsequently engaged in criminal behaviours such as theft and prostitution. Drug-connected women are those that misuse, manufacture, and or/distribute drugs via a relationship with a partner or family member. Lastly, the other pathway describes women who engaged in criminal activity out of economic motivation, with very little evidence of a history of early trauma or substance misuse (Daly, 1992).

The typologies created by Daly (1992) revealed some commonalities in terms of risk factors that may influence women's participation in crime. These factors include such things as victimization during childhood or adulthood, substance use, social marginalization, and parental stress. Quantitative approaches to examining female pathways to offending further validated some of the findings in Daly's research. For example, quantitative research consistently identifies high rates of early childhood trauma, housing and financial difficulties, and substance misuse among justice-involved women. (Brennan, Breitenbach, Dieterich, Salisbury, & Van Voorhis; Messina, Grella, Burdon, & Prendergast, 2007; Van Voorhis et al., 2010).

Regardless of the approach, both qualitative and quantitative gender-responsive research underscores the notion that women enter the criminal justice system from circumstances and situations that are different from men. Findings from the pathways literature prompted a movement towards recognizing how risk factors may differentially impact future engagement in crime based on gender, and the importance of incorporating these differences into correctional intervention programs. Evidence suggests that risk factors for offending can be categorized into three main groups; gender-neutral, gender-specific, and gender-salient (Brown & Motiuk, 2005; Brown & Motiuk, 2008).

Gender-neutral risk factors can be defined as risk factors that predict future engagement in crime equally for both men and women. An effect of the same magnitude is present in both genders. Examples of gender-neutral risk factors include being unemployed at the time of the offence, difficulties communicating with others, and negative attitudes towards corrections (Brown & Motiuk, 2005; Brown & Motiuk, 2008).

In contrast, gender-specific risk factors are those that predict criminal behaviour for one gender only. These risk factors must demonstrate a small, moderate, or large effect in predicting criminal behaviour in one gender, but show no effect for the other. Gender-specific risk factors for women include such things as combining the use of alcohol and drugs, being a victim of spousal violence, and demonstrating poor problem solving skills (Brown & Motiuk, 2005; Brown & Motiuk, 2008). Lastly, gender-salient risk factors bridge the gap between both gender-neutral and gender-specific categories, by demonstrating an effect in both genders, but the size of the effect is stronger for one gender. Examples of gender-salient risk factors for women include having an unstable job

history, having many criminal friends, and poor stress management (Brown & Motiuk, 2005; Brown & Motiuk, 2008).

Indeed, research supports the notion that risk factors for criminal involvement may differentially impact men and women offenders. Gehring and Van Voorhis (2014) investigated how pretrial needs, both gender-neutral and gender-specific in nature, may impact pretrial outcomes (i.e., failure to attend court, engaging in criminal activity during the pre-trial period) among men and women offenders. In terms of gender neutral risk factors, bivariate correlations between pre-trial needs and pre-trial outcomes demonstrated that substance misuse and criminal history were significant predictors of pre-trial outcomes for men and women, however, correlations for women were much higher. Despite being categorized as a gender-neutral risk factors under the RNR model, substance misuse and criminal history seem to be stronger predictors of pre-trial outcomes for women than for men (Gehring & Van Voorhis, 2014).

These findings are somewhat in keeping with a 2002 meta-analysis by Dowden and Brown that investigated the that role substance use plays in predicting recidivism for men and women offenders. The authors were interested in examining how different predictor categories of substance use (i.e., alcohol use, drug use, alcohol and drug use) could be used to predict both general and violent recidivism (Dowden & Brown, 2002). In total, 45 studies were found that produced 116 individual effect sizes for general recidivism and 29 effect sizes for violent recidivism. The mean effect size for substance use and general recidivism for men and women combined was 0.10, with alcohol/drug misuse generating the strongest mean effect size ($M=0.22$) followed by drug use ($M=0.19$), parental substance misuse ($M=0.13$) and lastly, alcohol use ($M=0.12$). All

substance use categories were significantly predictive of recidivism among men and women (Dowden & Brown, 2002).

When predicting violent recidivism, substance use was also a significant predictor for both men and women, with an overall mean effect size of 0.06 (Dowden & Brown, 2002). Disaggregating effect sizes by gender, the authors found that substance misuse was equally predictive for men and for women, however, the predictor category of parental substance misuse was uniquely predictive of general recidivism for women offenders. These results, coupled with those presented by Gehring & Van Voorhis (2014) suggests that risk factors such as substance misuse may in fact be best categorized as a gender-salient predictor, as substance misuse can predict recidivism for both men and women offenders, but appears to be more strongly correlated with pretrial outcomes for women than for men (Gehring & Van Voorhis, 2014).

In terms of the gender-specific risk factors and their relationship with outcomes for men, according to Gehring & Van Voorhis (2014), items assessing trauma, mental health, homelessness, and family support were moderate predictors of one of the outcomes of interest, a failure to appear to court, but were not strongly correlated with new arrests or any failures (failure to appear and new arrests combined). With regards to women, these risk factors seemed to portray a somewhat different relationship with pre-trial outcomes. Indeed, the gender-specific scales were consistent predictors of all pretrial outcomes for women. Further, a reported history of abuse emerged as a predictor of all pretrial outcomes for women, where abuse was not related to any pretrial outcomes for men (Gehring & Van Voorhis, 2014).

This research underscores the importance of recognizing that risk factors, whether they be gender-neutral, gender-salient, or gender-specific, can have differential impact on criminal behaviour based on gender. It is therefore important to develop intervention programs specifically for women that both recognize and address these differences. Although some institutions recognize the unique needs of women offenders and are working to meet these needs, it is also true that many institutions do not.

It is the case that the majority of intervention programs were designed and assessed on samples of male offenders, as they represent the majority of the offender population. These results are then generalized and applied to female offenders under the assumption that their needs are the same (Wright et al., 2012). However, according to gender-responsive models, for substance misuse interventions to be effective for women, these interventions must recognize the unique impact that certain risk factors, such as substance misuse, trauma, and victimization, may have on female offending patterns, and how this may differ from male offenders.

Effectiveness of current intervention programs

To date, several substance use intervention programs have been developed for the treatment of substance use disorders among women offenders (Bahr, Masters, & Taylor, 2012). Intervention approaches vary, but generally share common goals with respect to treatment outcomes. By definition, incarceration-based interventions include any interventions that specifically target individuals with substance misuse problems, with the goal of reducing substance misuse and related criminal behaviours (Mitchell, Wilson, & Makenzie, 2007). Thus, in addition to reducing substance use, incarceration-based interventions aim to reduce recidivism and future engagement in criminal activities.

Cognitive-behavioural therapy. One of the most common interventions for the treatment of substance use among offender populations is cognitive-behavioural therapy (CBT). CBT is a structured, goal-oriented intervention that places the focus on immediate problems faced by substance users entering treatment (Bahr et al., 2012). The foundations of CBT draws from two distinct fields of study, cognitive theory, which emphasizes the importance of internal thought processes in substance use, and behavioural theory, which emphasizes external behaviours (Carroll, 2000). The underlying assumption of CBT is that learned processes, such as maladaptive thought patterns and attitudes, contribute to the development, maintenance and reinforcement of substance use disorders. In order to reduce substance use, pre-existing, maladaptive thought patterns and attitudes must be restructured through the use of CBT. By helping the client to recognize situations where they may be at increased risk of engaging in substance use, to avoid high-risk situations and to effectively cope with a variety of problematic behaviours related to substance misuse, substance use disorders can be managed and treated (Carroll, 2000). It is designed to maintain flexibility and allow for individualization of treatment that can be suited to a range of clients and treatment settings (Carroll, 2000).

Offender populations pose a unique challenge for cognitive-behavioural approaches to substance misuse treatment. In addition to addressing pre-existing maladaptive thought patterns and behaviours related to substance use, CBT has the added task of identifying and restructuring criminal attitudes related to substance use (Prendergast, 2009). Since the discovery of the relationship between substance use and criminal behaviour, researchers have emphasized the importance of addressing criminal

attitudes and behaviours as they relate to substance use. Without addressing these needs treatment may be insufficient, as criminal attitudes and behaviours contribute to relapse and recidivism (Prendergast, 2009).

Therapeutic communities. Therapeutic communities (TCs) are highly structured residence programs that promote accountability and a sense of governance among offenders in treatment. TCs take a holistic approach to the understanding of addiction, suggesting that drug use is just a single element of behaviour that contributes to a more complex, behavioural disorder (Bahr et al., 2012). Treatment is delivered in an organized group setting, with peer appointed leaders that recognize peers or community members as the agents of behavioural change (Leukefeld, Gullotta, & Gregrich, 2011). Clients must learn and adopt new pro-social behaviours in order to effectively change existing, maladaptive behaviours. By instilling a sense of responsibility within the offenders for each other's well-being and success, peer pressure builds within the group, fostering compliance for the program rules. TCs can be adapted to include a variety of treatment modalities, such as cognitive interventions, individual counselling and group counselling (Bahr et al., 2012).

Contingency management. As greater emphasis was placed on exploring the neural basis of addiction, researchers began to better understand the role that reward plays in the development and maintenance of a substance use disorder. Indeed, it is inherent reward circuits within the brain that are hypothesized to be altered through the use of psychoactive substances (Chandler, Fletcher & Volkow, 2009). It is these same circuits however, that can also be targeted through intervention to manage substance use. This forms the basic foundation for contingency management (CM), a form of

intervention that administers rewards to reinforce positive behaviours related to the management and treatment of substance use, such as a negative urine sample, attending treatment sessions or completing other obligations related to the treatment process. The reinforcing nature of these rewards work to deter the individual from engaging in subsequent substance use (Bahr et al., 2012).

Drug treatment courts. As greater emphasis was placed on treating and rehabilitating offenders with substance use disorders, diverting offenders from the criminal justice system to a treatment setting became more commonplace. Drug treatment courts use the power of the justice system to encourage individuals to attend treatment and reduce substance use that is directly and indirectly related to criminal offending (Bahr et al., 2012). By diverting offenders to treatment rather than incarceration, offender well-being can be increased, while at the same time reducing the number of individuals incarcerated for substance use (Giacomazzi & Bell, 2007). To participate in drug treatment court, eligible individuals must plead guilty to their charges, with all charges withheld permitting compliance with court requirements (Bahr et al., 2012). Treatment requirements can include judicial monitoring, drug testing and intensive supervision (Giacomazzi & Bell, 2007). Failure to comply with court-appointed treatment requirements may result in incarceration, based on the fact that the individuals must plead guilty to their charges to be eligible for participation in treatment (Bahr et al., 2012).

Pharmacological interventions. In addition to behavioural interventions, pharmacological interventions for substance use are becoming more commonplace in criminal justice settings. Pharmacological approaches to treatment are rooted in harm

reduction theory, aiming to manage substance use by replacing the harmful drug of choice with a less harmful drug substitute (Leukefeld et al., 2011). There are two main ways in which pharmacological interventions can be used in the management and treatment of substance misuse. The first way it can be used is for the management of drug withdrawal and its associated symptoms. In addition, it can also be used as agonist maintenance treatment, the scheduled administration of a similar drug substitute to maintain a stable level of drug effect, without producing a noticeable “high”, or to precipitate withdrawal (World Health Organization, 2009). There are several common pharmacological treatments used to manage substance misuse and dependency among criminal justice populations.

Methadone. Used in the management of opiate dependency, methadone maintenance therapy (MMT) is an agonist drug substitute that can be used in place of opiate-based drugs. The main goal of MMT is to manage opiate dependence by providing a controlled dosage of a less harmful drug substitute, while also improving the quality of life for the substance user (Bahr et al., 2012). Methadone is one of the most frequently used pharmacological treatments for opiate dependence (Lobmann & Verthein, 2009).

With such a wide array of substance use interventions, it is important to try to understand which treatments are the most effective at reducing substance use and recidivism among women offender populations. Since the relationship between substance use and criminal behaviour is well established, it is crucial that substance use interventions are successful at reducing substance use in order to achieve a reduction in recidivism, and vice-versa. Recent emphasis on gender-responsive models and the

unique impact of gender-responsive variables, such as substance use for female offenders, underscores the importance of considering gender-responsive needs in the treatment of substance use disorders. Given that the majority of intervention programs were developed and evaluated on male offender populations, it is crucial that the underlying assumption of gender-neutrality is tested, to verify that current interventions are, in fact, effective for use with female offenders.

To date, three meta-analyses have been conducted examining the effectiveness of correctional interventions for women. Although these studies did not focus solely on substance misuse interventions, all three meta-analyses included studies with substance misuse outcomes.

In 1999, Dowden and Andrews conducted the first meta-analysis evaluating correctional interventions for women offenders. This review included 26 studies, 4 of which were identified as substance misuse interventions. When assessing the impact of intervention programs designed to address substance misuse, the authors found a slight negative correlation ($r = -0.01$) between participating in substance misuse interventions and future recidivism. These results suggest that substance misuse interventions are not associated with reductions in recidivism for women offenders. It is important to note however, that this meta-analysis incorporated a small number of evaluations and focused primarily on recidivism outcomes for juvenile offender populations (Dowden & Andrews, 1999). It is unclear how these results would generalize to other outcome measures, such as the use of alcohol and other drugs, or to samples with a greater number of adult women offenders.

Expanding upon the results presented by Dowden & Andrews (1999), the second meta-analysis conducted by Tripodi, Bledsoe, Kim, and Bender (2011) included 24 studies published from 1988 to 2008, 6 of which were identified as substance misuse interventions. Outcomes of interest were those related to crime (i.e., return to custody, arrests), substance use, mental health, and trauma. In terms of the types of substance misuse interventions, the majority of intervention programs were identified as therapeutic communities, followed by cognitive-behavioural therapy interventions. Six odds-ratio effect sizes were reported for recidivism outcomes (n =1588) ranging from 0.17 to 1.00, suggesting that participation in substance misuse interventions are associated with a reduction in the odds of recidivism for women offenders to varying degrees. In addition, two studies reported odds-ratios for substance misuse outcomes (n =494), with odds-ratios ranging from 0.31 to 1.42. These results suggest that an association exists between participation in correctional interventions programs and reductions in the odds of recidivism (Tripodi et al., 2011).

Lastly, Gobeil, Blanchette, and Stewart (2016) built upon the results of Tripodi et al. (2011), by expanding the time frame for eligible studies and by incorporating meta-regression to better investigate study moderators. In total, 38 studies evaluating correctional interventions for women were included, with 16 studies identified as interventions for substance misuse. In this case, the authors limited their analyses to those studies evaluation recidivism outcomes.

Analyses revealed that, as a collective, substance misuse interventions were associated with a reduction the odds of recidivism for women offenders. Examinations of the specific types of interventions demonstrated that both therapeutic community

interventions and cognitive-behavioural therapy were associated with reduction in the odds of recidivism. When parsing out the results by intervention type (i.e., CBT versus TC), the authors did not provide an overall effect estimate for substance misuse interventions specifically, however, that authors noted that substance misuse interventions contributed the strongest effects to the overall estimates for each intervention type (Gobeil, Blanchette, & Stewart, 2016). Lastly, evaluations of interventions identified as gender-neutral versus those identified as gender-informed revealed that both gender-informed and gender-neutral intervention programs were effective at reducing the odds of recidivism. Again, effect estimates were not reported specifically for substance misuse interventions (Gobeil et al., 2016).

Although a literature base exists providing some indication of which treatments are effective for use among women offender populations as a whole, it remains unclear which approaches to intervention are the most effective in reducing both substance use and recidivism outcomes among women offenders. Previous meta-analyses provide somewhat mixed evidence in terms of the effectiveness of substance misuse interventions in reducing a variety of post-release outcomes. In addition, none of the previous meta-analyses have targeted substance use and recidivism outcomes specifically. By definition, substance misuse interventions are expected to be effective in reducing both substance use and recidivism among women with substance misuse disorders (Mitchell et al., 2007). Therefore, it is of importance to explore the impact of various approaches to substance misuse interventions offered to women offenders, with the goal of determining the overall efficacy of these programs.

Expanding upon the results of previous meta-analytic reviews of correctional interventions for women, this study will evaluate all experimental and quasi-experimental evaluations of incarceration-based substance misuse treatment programs offered to women offenders from 2000 to 2017. The results will help to provide an evaluation of the effectiveness of incarceration-based treatment programs for women and will exemplify advancements made in intervention practices for female offenders who misuse alcohol and other drugs.

Methods

Meta-Analysis

Meta-analyses are a subset of systematic review, which attempt to summarize and integrate empirical evidence via the use of inclusion and exclusion criteria to answer a specific research question (Haidich, 2010). Meta-analyses are conducted to assess the strength of the evidence reported for a disease or intervention. Two aims have been identified with regards to meta-analyses. The first is to determine whether the effect of an intervention is positive or negative, and the second is to obtain a single summary estimate of the overall efficacy of the intervention under review (Haidich, 2010).

Preferred over narrative-style reviews which are largely subjective in nature, this approach applies objective formulas to collate quantitative results from each evaluation into a single effect estimate (Borenstein, Hedges, Higgins, & Rothstein, 2009). Meta-analysis is best suited for use with a group of studies that are sufficiently homogenous with respect to the participants involved, the intervention characteristics examined and the outcome measures of interest (Haidich, 2010). This goal is typically achieved through the application of predetermined, specific inclusion and exclusion criteria used to select studies. Despite efforts to ensure that studies are sufficiently homogenous, typically there is some degree of heterogeneity present across studies. Thus, one of the greatest benefits to the meta-analytic approach is the ability to examine sources of heterogeneity, as the presence of large amount of variability between studies can impact the generalizability of the overall effect estimate (Haidich, 2010).

Meta-analysis is just one of many approaches to conducting a systematic review. Indeed, each approach to systematic review exhibits unique characteristics, as well as

strengths and weaknesses. For example, scoping review, a common pre-cursor to meta-analysis, addresses a broadly defined, complex research question and summarizes data through both objective and subjective methods. This review is recommended for use when integrating a body of literature that is thought to be vast and diverse in terms of methodology, theoretical orientation, and disciplines (Thomas & Harden, 2008). This approach to review is unlikely to include an assessment of study quality (Djikers, 2015). Thematic synthesis is yet another type of systematic review that can be used to summarize a body of literature. This approach is reserved for the synthesis of primary qualitative research, where an insufficient number of randomized-control trials exist to take a quantitative approach to systematic review (Thomas & Harden, 2008).

Given that the research questions of interest in this review is rather specific, and the body of literature evaluating substance misuse intervention programs for women offenders is quite large, meta-analysis would be the best approach to evaluating this body of literature. In addition, the majority of evaluations utilize a treatment-control design, and often report the outcomes of interest as a quantitative measure (i.e., odds-ratio), which is best suited for meta-analysis. In the likely even that heterogeneity is present across evaluations, meta-analysis will allow for exploration of this variation. Lastly, this approach often incorporates an assessment of methodological quality, another point of interest in this review.

Eligibility criteria

1. Eligible studies employed the use of a substance misuse intervention program aimed to reduce substance use, reoffending, or both among adult female offenders or predominantly female adult offender samples with problematic substance use.

2. **Research Respondents.** To be eligible for inclusion, both the treatment and control participants were adult female offenders who are serving a sentence in any formal correctional institution, or serving a sentence within the community (i.e., probation, parole, drug treatment courts). Adult offenders are defined as those 18 years of age or older. Evaluations focusing on female youth populations were excluded. Indeed, research suggests that youth and adult offenders exhibit different needs, and it is often the case that youth offenders require different services while incarcerated (Tripodi et al., 2011). As such, this meta-analysis focused on substance misuse intervention programs for adult women offenders.
3. **Key Variables.** Studies reported at least one outcome measure of substance use, reoffending or both. Since the overall goals of substance misuse intervention programs are to reduce further involvement in criminal activity, as well as cope with high-risk, challenging situations without the use of drugs or alcohol (Doherty, Ternes, & Matheson, 2014), outcome measures of substance use, recidivism, or both were of interest in this case. Substance use could be measured using urinalysis testing (Gonzalez-Menendez, Fernandez, Rodriguez & Villagra, 2014), self-report using an established substance misuse screening tool such as the AUDIT (Newbury-birch et al., 2014), or a combination of measures. Recidivism was defined as either a formal institutional measure, such as re-arrest, reconviction and re-incarceration, or as a self-reported measure of reoffending (Koehler, Humphreys, Akoensi, Sanchez de Ribera, & Losel, 2014). Studies were included if they reported any common statistics or raw data that allowed for the calculation of effect sizes.

4. Research Methods. To be considered for inclusion in the study, the evaluation compared the effect of an intervention, as applied to a particular treatment group, to the level of reoffending, substance use and/or both in an equivalent comparison and/or control group. This corresponds to a level three or more on the Maryland Scientific Methods Scale, which was defined as a comparison between two or more comparable units of analysis, one with and one without the program (Sherman, et al., 1998).

A control group can be defined as a group of participants receiving no treatment, treatment as-usual or placebo. Studies were excluded if the treatment group was compared to any of the following: national statistics of the general offender population, a sample of participants that entered the same program, but through different routes (e.g. comparing treatment effectiveness for participants who entered treatment voluntarily to those who were court-ordered into the same program), and treatment drop-outs (Koehler et al., 2014).

Chemical intervention studies were excluded if the control group comprised subjects who received the same treatment but in small doses. Given that the inclusion criteria require studies to have both a treatment, a control group, which is defined as a group receiving no treatment, treatment as-usual, having a single comparison group simply receiving a smaller dose of medication does not meet the inclusion criteria. If, on the other hand, there are multiple comparison groups, and at least one qualifies as a control group by the stated definition, then the study was included, and comparisons were made between the control group receiving no medication and an average of the two groups receiving medication.

5. Cultural and Linguistic Range. To be eligible for inclusion, studies were reported in English or French.
6. Publication Type. Published and unpublished studies were eligible for inclusion, including dissertations and conference presentations.

Literature Search

Searches of studies from January 2000 to September 2017 were conducted in PsycINFO, Criminal Justice Abstracts, Social Work Abstracts, Social Sciences Full Text and SAGE Journals. This timeframe was chosen based on the notion that women offenders were often neglected in criminal justice research prior to the early 2000s (Gobeil et al., 2016). In addition, it is generally known that any correctional interventions offered to women offenders prior to the turn of the last century were initially developed for use with men offenders, and simply applied to women (Gobeil et al., 2016). Given that it is of interest to explore the differential impact of gender-neutral and gender-informed interventions, evaluations conducted prior to the year 2000 were excluded.

Search terms were derived first by creating a list of common indexed search terms reported in related meta-analyses and other relevant articles. These terms were then searched in the index/thesaurus unique to each database. Search terms included those reflecting gender (women, female), population (perpetrators, prisoners, criminals, incarcerated), target of intervention (drug dependency, drug abuse, drug usage, substance use disorder, alcohol abuse drug addiction, addiction), intervention (behaviour therapy, drug rehabilitation, treatment effectiveness evaluation, intervention, correctional intervention, treatment, treatment outcomes, dialectical behavior therapy, cognitive behavior therapy). No explicit search terms were defined for the In addition, a variety of

government and other websites were searched, including the Australian Institute of Criminology, National Criminal Justice Reference Service, the U.S. National Institute of Corrections, Correctional Service of Canada, Public Safety Canada, the Bureau of Justice Statistics, and the U.K. Ministry of Justice. Studies obtained through these means were assessed for relevance, and for those studies that met the established inclusion criteria, reference lists were examined for possible additional resources.

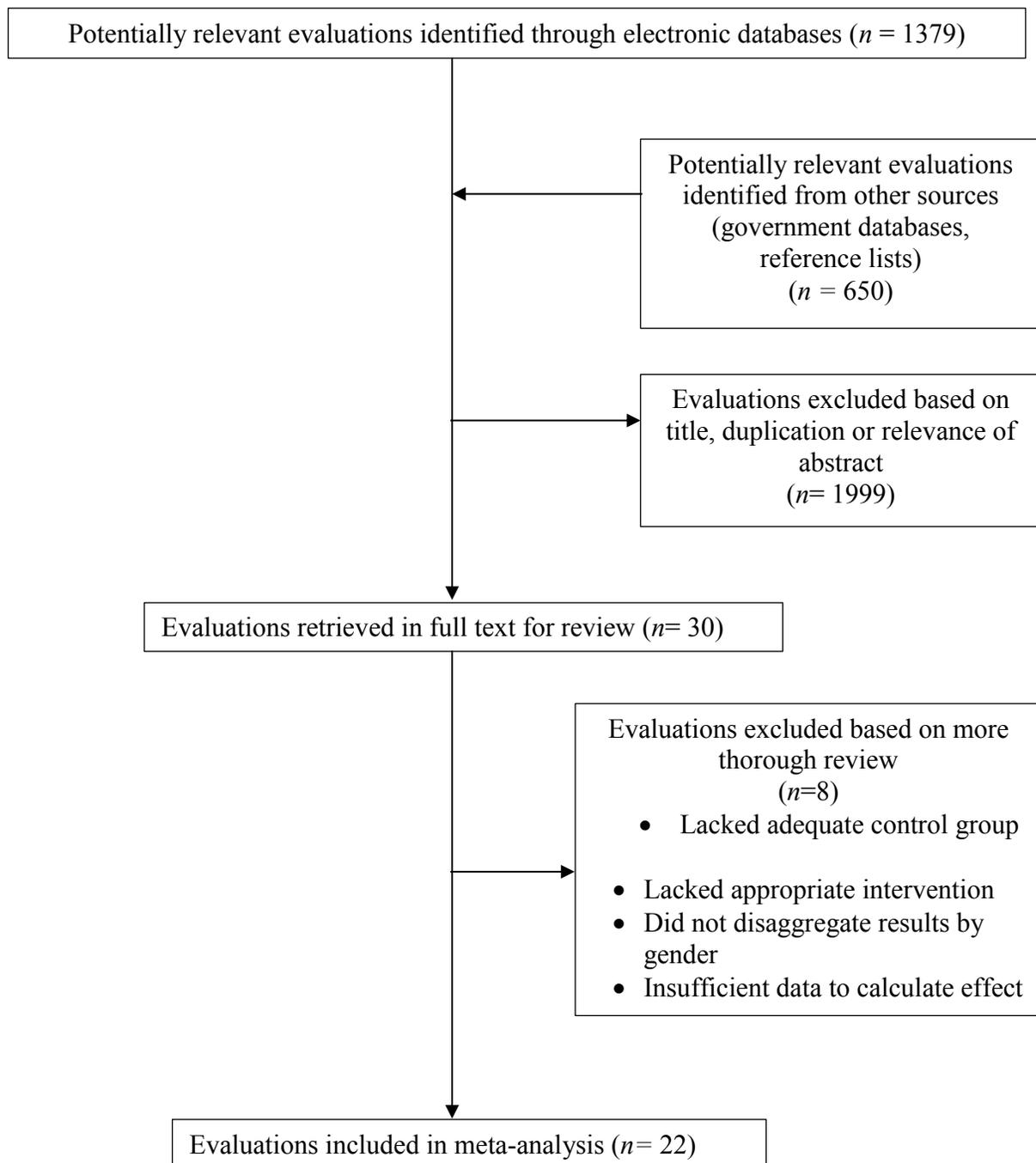


Figure 1. Flow diagram of literature search process

Procedure

Data coding and extraction was completed by the principle investigator. In an attempt to establish a reliable coding manual, first, the principle investigator and a fellow graduate student volunteer reviewed the coding manual and together, coded an evaluation. Then, both independently coded two studies and compared the results. Discrepancies in coded items were discussed and any necessary changes were reflected in the coding manual to enhance overall clarity before formally testing inter-rater reliability on two independent cases. Once suitable levels of inter-rater agreement were established between raters, all studies were then coded by the principle investigator.

Study Descriptors

Data included in the coding manual can be summarized into one of four categories: study design, sample characteristics, intervention characteristics, and study quality. First, information relating to study design was coded including participant assignment to groups (random/quasi-random, non-random but matched, non-random, no matching), recidivism outcomes (new conviction, new charge/arrest, return to custody, any criminal activity), substance use outcomes (relapse, return to use), outcome sources (official records, self-report, urinalysis), sample size, and length of the follow-up period. Information relating to publication status, peer-review status, the year of publication, and country of publication was also noted.

Second, sample characteristics were coded, including participant age, ethnicity, institutional security level (minimum, medium, maximum, multi-level), reason for participation (voluntary, mandated, recommended, other), rate of co-occurring mental-

health disorder(s), most prevalent co-occurring mental disorder(s) and details surrounding the experience of childhood and/or lifetime trauma.

Third, information surrounding intervention characteristics were coded including intervention type (cognitive-behavioural therapy, mindfulness, 12-step, therapeutic community, drug treatment court, boot camp, pharmacological intervention), location of intervention (institution, community, both), implementers (evaluator/researcher, practitioner, volunteer, technology), length of intervention and whether the intervention was gender-informed. Gender-informed interventions were defined as interventions that explicitly consider needs that are particularly salient to women, in addition to traditional elements of evidence-based practices (Gobeil et al., 2016).

Study Quality

Given that evaluations must have a treatment-control design to be eligible for inclusion, studies considered to have low study quality were filtered out of the study selection process. Consequently, all eligible studies included in this review would be rated as a level 3 or greater on the Maryland Scientific Method Scale (SMS; Sherman, Gottfredson, MacKenzie, Eck, Reuter, & Bushway, 1998). A rating of level 3 on the SMS is generally considered to be the minimum level of study quality acceptable for intervention reviews (Madaleno & Waights 2016).

In addition to incorporating study methodology into the process of study selection, it was of interest to explore further the impact of study quality on eligible evaluations. To do this, the Collaborative Outcome Data Committee's Guidelines for the Evaluation of Sexual Offender Treatment Outcome Research (Collaborative Outcome Data Committee, 2007) were applied to each evaluation. The CODC guidelines consist of

21 individual items, grouped into 13 over-arching categories. Scores on each item range from 0 to 2, with 0 indicative of little or no adherence to the quality item in question, 1 indicative of some/partial adherence to the quality item in question and 2 indicative of complete adherence to the quality item in question (CODC, 2007). Using a structured-judgment approach, a global quality rating is then created based on a combination of the item scores.

Global ratings are expressed as one of four categories of study quality: strong, good, weak, and rejected. A global quality rating of strong describes an evaluation that has minimal bias in estimating the effectiveness of the intervention. The evaluation is well-designed, well-executed and has convincing results. Second, a good quality evaluations is an intervention that demonstrates no more than minimal bias. Reasonable efforts have been made to address threats to validity, but much remains unknown (CODC, 2007). Characteristics of stronger evaluations would include high levels of confidence in the results of the evaluation. Evaluations include such characteristics as clear definitions and descriptions of the treatment and control groups, sample sizes of 50 or more in both treatment and control groups, low rates of attrition, and an average follow-up period of 3 years or more (CODC, 2007).

Further, a quality rating of weak describes an evaluation that has no more than a small amount of bias, but demonstrates significant flaws. These studies may still be of relevance to the question of treatment effectiveness, however the evidence is weak at best. Lastly, a study with a global quality rating of rejected is one that demonstrates a considerable amount of bias. The study has multiple significant flaws, but may still be of

relevance to the question of treatment effectiveness. Again, the evidence is weak at best. (CODC, 2007).

Characteristics common to studies with weak or rejected quality include uncontrolled miscellaneous factors that may bias the results, sample sizes less than 50 in both the treatment and intervention groups, large participant attrition during intervention and during the follow-up period, and an average follow-up period of less than three years. Confidence in the results of these evaluations is generally low (CODC, 2007).

In this case, the global quality ratings produced for each evaluation were used to investigate heterogeneity across studies, rather than to filter out evaluations of low or unacceptable study quality. Given that the eligibility criteria inherently exclude studies of unacceptable quality, the global quality ratings produced by the CODC reflect evaluations already considered to meet the standards for inclusion in intervention reviews.

Inter-rater reliability

In order to examine inter-rater reliability, 2 studies were independently coded by the principle investigator and a fellow graduate student volunteer. Agreement for categorical variables was assessed using Cohen's Kappa, while continuous variables were assessed using intra-class correlation coefficients (ICC). Kappa statistics were calculated for 9 of the 23 categorical variables included in the coding manual. Statistics could not be calculated for the remaining 14 categorical variables, as there was no variation in responses between raters, and thus, a measure of association could not be computed. Eight of the nine variables had a Kappa value of 1.00, indicative of perfect agreement between raters. The remaining variable received a Kappa value of 0.33, indicative of fair

agreement between raters (McHugh, 2012). In this case however, the Kappa value was based on a single comparison between raters, as responses from both raters were the same for the second observation. Again, this lack of variation between responses did not allow for a measure of association to be computed for the second observation. Intra-class correlations could be calculated for 13 of the 14 continuous variables. ICC values were 1.0 for all variables eligible for assessment. The remaining variable had zero variation between raters, and thus, an association between responses could not be measured.

Aggregation of Data

At least one effect size was coded from each evaluation. In the event that evaluations reported multiple outcomes for recidivism and substance use, a single effect size was coded for each of the three outcome categories; recidivism, alcohol use, and drug use. Thus, each study could contribute up to three independent effect sizes. Effect sizes were coded in such a way that positive effects indicated that the treatment group had a more favourable outcome than the comparison group (i.e., less re-offending or substance use during the follow-up period).

For dichotomous outcomes, odds-ratios were used. Outcomes based on non-dichotomous scales were coded using the standardized mean difference effect size (SMD; Lipsey & Wilson, 2001). In order to combine these two effect sizes, SMDs were converted to logged-odds ratios (Chinn, 2000). Then for the purpose of data aggregation, all odds-ratios were transformed to logged odds-ratios. Finally, the logged odds-ratios were combined using the inverse-variance method. This method of analysis weights studies by the inverse of their variance, ultimately giving greater weight to larger studies with small standard errors, and less weight to smaller studies with larger standard errors

(The Cochrane Collaboration, 2011). The final average effect size was then transformed back to an odds-ratio, as this effect estimate best reflects the outcomes of interest.

In the event that studies included multiple comparisons within a single study, data from both treatment groups were collapsed, and this new combined data was used to calculate an effect size and variance. This effect size was then compared to that reported for the control group (Borenstein et al., 2009). This allowed for the study to contribute a single combined effect size reflecting data from both treatment groups.

Both fixed-effect and random-effects models were computed (Lipsey & Wilson, 2000). The fixed-effect model assumes that there is one true effect size underlying all of the studies in the analysis, and that any differences in observed effect are due to sampling error. This model assumes that intervention effects are not influenced by between-study characteristics such as study population, differences in subject selection criteria, and treatment implementation (Borenstein et al., 2009; Haidich, 2010). This limits the application of these results, as the fixed-effect model reflects only those studies included in the analysis. The results of this model can only be used to make conditional inferences, those where that are only applicable to the set of studies under review. The weights used to combine effect sizes under the fixed-effect model is determined solely by within-study variance. Given that this model assumes one true effect size underlying all studies, between-study variation would not exist under this model (Borenstein et al., 2009;).

By contrast, the random-effects model assumes that the true effect size can vary from study to study, and thus, results can be used to estimate effects within the full population of possible studies (Borenstein et al., 2009). This model recognizes the potential impact of between-study variability, and thus provides a measure of this

variability and attempts to account for this variability when producing the overall effect estimate. Under this model, the weights used to combine studies are determined by both within-study and between-study variation. By incorporating both types of variation in the model, this allows for broader generalizations to be made to studies outside of the scope of the review by allowing for and anticipating greater differences across evaluations (Borenstein et al., 2009).

In addition to computing an overall summary effect, it was also of interest to explore patterns of variation across studies. In order to assess heterogeneity between studies, the Q statistic was calculated. The Q statistic is computed by summing the squared deviations of each study's effect estimate from the overall effect estimate, and then weighting the contribution of each study by the inverse of its variance. A significant Q-statistic would suggest that the studies included in the meta-analysis have greater variation than would be expected by chance, and that this variation should be explored further (Heudo-Medina, Sanches-Meca, Marin-Martinez, & Botella. 2006).

Given that there are three outcome variables of interest; recidivism, alcohol use, and drug use, three separate meta-analyses were performed. This allowed for each study to contribute up to three unique effect sizes, without counting participant data twice in a single analysis. Double-counting participants creates error within the analysis, as the correlation between estimated intervention effects cannot be accounted for (The Cochrane Collaboration, 2011).

Results

Description of eligible studies

In total, 22 studies reporting 39 unique effect sizes (20 recidivism, 8 alcohol, 11 drug use) were included in the analyses (see Table 1). Of these, 17 were conducted in the United States, 3 in Canada and 2 in Spain. The majority of the evaluations were coded from studies published in peer-reviewed journals (n=17), followed by government reports (n=5) and finally, one dissertation. In terms of participant assignment to treatment groups, 12 studies were identified as random-assignment, while 11 were identified as quasi-random. 95% of the evaluations assessed were behaviour-based interventions, with the majority coded as a TC or CBT program. Only one study evaluated the effects of a pharmacological intervention on substance use.

Table 1. Summary of studies and outcomes

Study ID	Study	Sample Size		Outcome	Intervention	Quality
		Treatment	Control			
1	Dowden & Blanchette (2002)	58	40	Recidivism	Unclear	Weak
2	Hall, Prendergast, Wellisch, Patten, & Cao (2004)	101	79	Recidivism, drug use	CBT	Rejected
3	Harrel, Roman, & Sack (2001)	110	26	Recidivism, drug use, alcohol use	DTC	Weak
4	Johnson, Friedmann, Green, Harrington, & Taxman (2011)	39	38	Recidivism, drug use, alcohol use	CBT	Good
5	Villagr� Lanza & Gonz�lez-Men�ndez (2013)	18	13	Drug use, alcohol use	ACT	Weak
6	Villagr� Lanza, Fern�ndez Garc�a, Rodr�guez Lamelas, & Gonz�lez-Men�ndez (2014)	37	13	Drug use, alcohol use	ACT	Weak
7	MacSwain, Farrell-MacDonald, & Cheverie (2014)	92	45	Recidivism	MMT	Good
8	Matheson, Doherty, & Grant (2008)	318	242	Recidivism	CBT	Weak
9	Messina, Burdon, & Prendergast (2006)	107	96	Recidivism	TC	Weak
10	Messina, Grella, Cartier, & Torres (2010)	60	55	Recidivism, drug use, alcohol use	TC	Weak
11	Miller, Adams, Trunk, & Drogosz (2010)	122	121	Recidivism	TC	Weak
12	Mosher & Phillips (2006)	322	322	Recidivism	TC	Rejected
13	Pelissier et al. (2003)	245	228	Recidivism, drug use	CBT	Weak
14	Robbins, Martin, & Surratt (2009)	160	116	Recidivism, drug use	TC	Weak
15	Sacks et al. (2008)	163	151	Recidivism, drug use, alcohol use	TC	Good
16	Sacks, McKendrick, & Hamilton (2012)	207	163	Recidivism, drug use	TC	Good
17	Scott & Dennis (2012)	224	238	Recidivism, alcohol use	RMC	Good
18	Scott, Dennis, & Lurigio (2017)	224	238	Recidivism	RMC	Good
19	Stein, Caviness, Anderson, Hebert, & Clarke (2009)	125	120	Recidivism	MI	Weak
20	Waid (2010)	2753	4396	Recidivism	TC	Good

21	Watson, Adkins, Cook, & Stageberg (2010)	164	173	Recidivism	TC	Weak
22	Zlotnick, Johnson, & Najavits (2009)	23	21	Recidivism, drug use, alcohol use	CBT	Weak

Note. ACT= Acceptance Commitment Therapy, CBT= Cognitive Behavioural Therapy, DTC= Drug Treatment Court, MMT= Methadone Maintenance Therapy, RMC= Recovery Management Checkups, MI= Motivational Interviewing. Quality ratings were determined by applying the CODC's guidelines for evaluation of sexual offender treatment outcome research. Global quality ratings serve as an overall summary of study quality and are categorized as follows: rejected, weak, good and strong. (Collaborative Outcome Data Committee, 2007).

Effect of intervention participation on recidivism

Table 2 presents the results of both the fixed-effect and random-effects models for all three outcomes of interest. With regards to the effect of intervention on recidivism, 20 studies reported outcome data relating to recidivism outcomes. Odds-ratios for each study reporting a recidivism outcome can be found in table 3. In total, 80% of the 20 odds ratios for re-offending indicated that the treatment group recidivated at a lower rate than the comparison group. The fixed-effects mean OR was 1.53 (95% CI, 1.36 to 1.73). The random-effects model resulted in a slightly larger effect, with an OR of 1.79 (95% CI, 1.34 to 2.40). These findings indicate that, in general, participation in substance misuse intervention programs, either during incarceration or while on probation/parole, was associated with a reduction in the odds of reoffending during the follow-up period. The variability among studies was statistically significant ($Q=97.51$, $p<0.001$), and the value obtained for I^2 indicated that 81% of the observed variability between studies was more than would be expected by chance. This suggests that there may be important differences in research methodology, sample, and/or interventions that may account for this variability in effect size.

Typically, to explore these variations, intervention-level variables would be examined to determine their association with recidivism outcomes. Despite trying to capture important study characteristics by coding information from each evaluation, efforts to extract relevant study characteristics were hindered by the quality of study descriptions provided by the authors. This in turn, limited the number of subgroup-analyses that could be performed to explore potential moderating variables.

Table 2. Effect sizes according to outcome

Outcome	Fixed-Effect		Random-Effect		Q	I ²	K	n	Studies
	OR	95% CI	OR	95%CI					
Recidivism	1.53	[1.36, 1.73]	1.79	[1.34, 2.40]	97.51**	81	20	12,525	1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
Drug use	1.48	[1.21, 1.81]	1.53	[1.13, 2.08]	17.81*	44	11	2066	2, 3, 4, 5, 6, 10, 13, 14, 15, 16, 22
Alcohol use	1.13	[0.88, 1.46]	1.82	[1.00, 3.31]	24.87**	72	8	1229	3, 4, 5, 6, 10, 15, 17, 22

Note. OR= Odds-Ratio, CI= Confidence Interval. * $p < 0.05$, ** $p < 0.001$

Table 3. Odds-ratios and confidence intervals of evaluations presenting recidivism outcomes

Study ID	Study	Odds-ratio	Lower Limit	Upper Limit
1	Dowden & Blanchette, 2002	4.64	1.65	13.07
2	Hall et al., 2004	1.67	0.92	3.04
3	Harrell et al., 2001	5.04	1.98	12.85
4	Johnson et al., 2011	7.54	1.54	36.83
7	MacSwain et al., 2010	9.30	2.81	30.38
8	Matheson et al., 2008	1.28	0.91	1.80
9	Messina et al., 2006*	0.65	0.36	1.18
10	Messina et al., 2010	1.80	0.84	3.85
11	Miller et al., 2010	1.74	0.90	3.35
12	Mosher & Phillips, 2009	2.96	1.98	4.42
13	Pelissier et al., 2003	1.28	0.85	1.92
14	Robbins et al., 2009*	0.61	0.39	0.95
15	Sacks et al., 2008	1.53	0.98	2.39
16	Sacks et al., 2012	3.96	2.30	6.81
17	Scott & Dennis, 2012*	0.69	0.44	1.08
18	Scott et al., 2017	2.58	1.66	4.02
19	Stein et al., 2010	1.63	0.72	3.70
20	Waid, 2010*	0.86	0.55	1.35
21	Watson et al., 2010	1.95	1.19	3.20
22	Zlotnick et al., 2009	2.92	0.85	10.05

Note. * refers to studies that present treatment outcomes in favour of the control condition

Subgroup-analyses for recidivism outcomes

Two subgroup analyses were performed to explore potential sources of variability among studies (See Table 4). The first sub-group analysis compared gender-informed interventions (those coded as either having clear evidence or partial evidence of gender-informed characteristics) with gender-neutral interventions (those that had no evidence of gender-informed characteristics). Results indicated that interventions with gender-informed characteristics were equally as effective in reducing reoffending compared to interventions with no gender-informed characteristics ($X^2= 1.43, p= 0.23$). It is important to keep in mind that these analyses were performed on a small subset of studies (11 for gender-informed, 8 for gender-neutral). In addition, with such a large amount of variability (I^2 values of 77% for gender-informed, and 81% for gender-neutral), it is important to interpret and apply these results with caution.

The second subgroup analysis compared Cognitive Behavioural Therapy (CBT) interventions with Therapeutic Community (TC). Results suggested that CBT interventions were equally as effective in reducing the odds of reoffending during post-release when compared to TC interventions ($X^2= 0.10, p= 0.76$). Again, it is important to keep in mind that these sub-group analyses were performed on a small subset of evaluations (5 CBT studies, and 9 TC studies).

Table 4. Sub-group analyses for recidivism outcomes

Focus	Fixed-Effect		Random-Effect		Q	I ²	K	n	Studies
	OR	95% CI	OR	95% CI					
Overall (Recidivism)	1.53	[1.36, 1.73]	1.79	[1.34, 2.40]	97.51**	81	20	12,525	1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
Gender Informed									
Partial/Clear Evidence	1.54	[1.34, 1.79]	1.60	[1.16, 2.20]	44.13**	77	11	10,462	2, 8, 10, 12, 13, 14, 15, 16, 20, 21, 22
No evidence	1.28	[0.98, 1.67]	2.09	[1.06, 4.12]	37.70**	81	8	1,762	3, 4, 7, 9, 11, 17, 18, 19
CBT versus TC									
CBT	1.42	[1.13, 1.80]	1.54	[1.10, 2.17]	6.46	38	5	1,334	2, 4, 8, 13, 22
TC	1.49	[1.27, 1.75]	1.45	[0.96, 2.17]	49.69**	84	9	9,651	9, 10, 11, 12, 14, 15, 16, 20, 21

Note. OR= Odds-Ratio, CI= Confidence Interval, CBT= Cognitive Behavioural Therapy, TC= Therapeutic Community, * $p < 0.05$, ** $p < 0.001$.

Effect of intervention participation on drug use

The results of the fixed and random effects model for drug use outcomes can be found in table 2. In order to examine the effect of intervention on drug use, 11 unique effect sizes were analysed. Table 5 presents the odds-ratios and confidence intervals of each of the 11 studies that reported drug use outcomes. Overall, 91% of the odd ratios for drug use suggested that substance misuse interventions reduced the likelihood of drug use after release from custody. Indeed, the results confirmed lower rates of drug use among women who participated in substance misuse interventions either during incarceration, or while on probation/parole, with a fixed-effect mean OR of 1.48 (95% CI, 1.21 to 1.81) and a random-effect OR of 1.53 (95% CI, 1.13 to 2.08). The variability among studies however was statistically significant ($Q=17.81, p=0.06$), and the I^2 value indicated that 44% of the observed variability between studies was more than would be expected by chance.

Given such a small number of studies, and lack of sufficient data, it was not possible to conduct moderator analyses to explore this variation further. The Cochrane Handbook for Systematic reviews of Interventions (2011) recommends having at least 10 studies in each group when performing subgroup analyses to investigate heterogeneity. In addition, there is a positive relationship between the number of characteristics examined using subgroup analyses and the likelihood of a false positive result. Therefore it is recommended to limit subgroup analyses to those with adequate data and be conservative in the number of characteristics explored through subgroup-analyses (The Cochrane, 2011).

Effect of intervention participation on alcohol use

Lastly, the overall effect estimates produced by the fixed and random-effects models can be found in Table 2. To examine the impact of intervention on alcohol use, the effects of 8 studies were combined. Table 6 presents the odds-ratios and confidence intervals of each of the 8 studies that reported alcohol use outcomes. In total, three quarters of the studies indicated that substance misuse treatment reduced the likelihood of alcohol use after release from a correctional institution. Meta-analytic results indicated less alcohol use post-release among women offenders who participated in an intervention program, with a fixed-effects mean OR of 1.13 (95% CI, 0.88 to 1.46). Similar to the previous results, the random-effects model produced a slightly larger overall effect, with an OR of 1.82 (95% CI, 1.00 to 3.31). Again, the variability between studies was statistically significant ($Q=24.87, p<0.001$), and just over three-quarters of the variability was more than would be expected by chance ($I^2=72\%$). Given the small number of studies, it was not possible to conduct moderator analyses to explore this variation further.

Table 5 Odds-ratios and confidence intervals of evaluations presenting drug use outcomes

Study ID	Study	Odds-ratio	Lower Limit	Upper Limit
2	Hall et al., 2004	3.22	1.34	7.76
3	Harrel et al., 2001	1.69	0.71	4.02
4	Johnson et al., 2011	1.64	0.42	6.34
5	Lanza & González-Menéndez, 2013	1.16	0.06	22.55
6	Lanza et al., 2014	3.02	0.87	10.55
10	Messina et al, 2010*	0.59	0.30	1.16
13	Pelissier et al., 2003	1.39	0.96	2.01
14	Robbins et al., 2009	2.83	1.57	5.10
15	Sacks et al., 2008	1.06	0.62	1.80
16	Sacks et al., 2012	1.51	0.96	2.38
22	Zlotnick et al., 2009	1.34	0.14	12.61

Note. * refers to studies that present treatment outcomes in favour of the control condition

Table 6. Odds-ratios and confidence intervals of evaluations presenting alcohol use outcomes

Study ID	Study	Odds-ratio	Lower Limit	Upper Limit
3	Harrel et al., 2001	5.20	1.58	17.12
4	Johnson et al., 2011	7.54	1.61	35.35
5	Lanza & González-Menéndez, 2013	1.20	0.05	31.64
6	Lanza et al., 2014	3.19	0.99	10.32
10	Messina et al., 2010	1.89	0.96	3.71
15	Sacks et al., 2008*	0.71	0.41	1.21
17	Scott et al., 2012	0.79	0.54	1.15
22	Zlotnick et al., 2009	2.43	0.61	9.66

Note. * refers to studies that present treatment outcomes in favour of the control condition

Study Quality

To evaluate study quality, the Collaborative Data Committees guidelines for the evaluation of sexual offender treatment outcome research were applied to each study to create a global rating of study quality (CODC, 2007). Global quality ratings are categorized as follows: rejected (low confidence in results, and study has considerable bias), weak (some confidence that the study has a small amount of bias, but the design is significantly flawed), good (high confidence that the study has a small amount of bias, and efforts were made to address validity concerns) and strong (high confidence that the study has minimal bias, well-designed and well executed). Just over half of the studies received a global quality rating of weak (57%), followed by a global quality rating of good (33%), and finally, a global rating of rejected (10%). None of the evaluations received a global rating of strong.

In an attempt to further investigate sources of variability between studies, a subgroup-analysis was conducted to compare evaluations rated as good quality, with those rated as weak or rejected. Due to the small number of evaluations included in the drug and alcohol meta-analyses, comparisons across study quality could only be performed on evaluations reporting a recidivism outcome. The results of this subgroup analysis are presented in Table 4, and show that evaluations receiving a global quality rating of good produced similar reductions in recidivism (random-effect OR of 2.14, 95% CI, 1.16 to 3.95) compared to those evaluations receiving a quality rating of weak or rejected (random-effect OR of 1.67, 95% CI, 1.21 to 2.29). The test for subgroup differences was not significant ($\text{Chi}^2 = 0.49, p = 0.48$), meaning that higher quality studies did not significantly reduce the odds of reoffending above that of low quality studies.

Given such a small number of studies in the subgroup analysis, (7 good quality studies and 13 low/rejected quality studies), this limits the overall level of confidence one can have in the results. In addition, the variability among studies was statistically significant, suggesting more variation between studies than would be expected by chance. Again, given a lack of comprehensive data surrounding study level moderators, it was not possible to explore this variation further.

Table 7. Study quality comparison: Recidivism outcomes

Global Quality Rating	Fixed-Effect		Random-Effect		Q	I ²	K	n	Studies
	OR	95% CI	OR	95%CI					
Good	1.59	[1.30, 1.94]	2.14	[1.16, 3.95]	48.05**	88	7	1,881	4, 6, 14, 15, 16, 17, 19
Weak/Rejected	1.55	[1.33, 1.81]	1.67	[1.21, 2.29]	44.63**	73%	13	3,309	1, 2, 3, 8, 9, 10, 11, 12, 13, 14, 21, 22

Note. OR= Odds-Ratio, CI= Confidence Interval. * $p < 0.05$, ** $p < 0.001$. Quality ratings were determined by applying the CODC's guidelines for evaluation of sexual offender treatment outcome research. (Collaborative Outcome Data Committee, 2000)

Discussion

These results demonstrated that women who participated in substance misuse intervention programs, either during incarceration or while on community supervision (probation, parole, etc.) were more successful on every outcome measure of interest when compared to non-participants, or participants in alternative treatment. In terms of recidivism, meta-analytic results demonstrated that women who participated in these interventions had 56% to 79% greater odds of remaining offence-free during the follow-up period. In addition, the impact of intervention on substance use outcomes was also promising. Women offenders who participated in substance misuse intervention had 47% to 51% greater odds of remaining drug-free, and 69% to 114% greater odds of remaining alcohol-free during the follow-up period.

Comparisons between interventions identified as gender-informed (either partially or completely) and those identified as gender-neutral suggest that participation in any intervention, regardless of gender focus, reduced the odds of reoffending during the follow-up period. These results are partially in keeping with a 2016 meta-analysis evaluating correctional interventions for women offenders, where the authors noted that participation in either gender-neutral or gender-informed interventions were associated with significant recidivism outcomes (Gobeil et al., 2016). Although this meta-analysis did not have a specific focus on substance misuse interventions, substance misuse interventions contributed the strongest effects to the overall effect size estimate.

It is generally known that in order to effectively reduce recidivism, correctional interventions must target needs that are directly related to offending (i.e., criminogenic needs; Andrews et al., 1990). Indeed, interventions are of little value if participants are

not receiving the appropriate services suited to their needs (Wright, Van Voorhis, Salisbury, & Bauman, 2012). This is thought to be even more crucial when rehabilitating women offenders, as research suggests that women may become involved in crime for difference reasons, and often present with needs different from those of men. In order to achieve this, it is crucial that institutions triage women offenders using established risk assessment instruments to assess the needs unique to each woman, and identify the appropriate intervention to best suit their needs (Wright et al., 2012).

A recent study evaluating gender-responsive interventions for juvenile female offenders highlighted the importance of assessing risk prior to assigning participants to intervention. Anderson et al. (2016) compared recidivism outcomes for 172 juvenile female offenders residing in a group-home based, gender-responsive intervention, versus 814 juvenile female offenders who received standard probation services. Participants were mandated to treatment via the court system, with moderate and high-risk participants assigned to the gender-responsive group homes, while low-risk participants received standard probation services (Anderson et al., 2016). The results demonstrated that those juvenile female offenders receiving treatment in gender-responsive group homes had significantly lower 2-year recidivism rates compared to those receiving standard probation services (28.4% vs 42.0%). The authors attributed the success of the intervention to its focus on criminogenic needs (e.g., family and peer relationships, education), along with gender-responsive components (e.g., history of trauma and mental health needs), that were relevant to participant sample. Indeed, those receiving group-home treatment demonstrated significantly higher criminogenic risk scores upon entering the criminal justice system. It was expected that those participating in the gender-

responsive group homes would present with greater needs, given that participants were mandated to intervention based on risk level (Anderson et al., 2016).

Saxena, Messina, and Grella (2014) also underscored the importance of assessing risk and treatment needs prior to intervention, but in this case, among a sample of adult women offenders. Women were randomized into one of two treatment conditions: gender responsive intervention or therapeutic community intervention. Treatment outcomes of interest included experiencing depression and the use of substances during the follow-up period (Saxena, Messina, & Grella, 2014). Analyses showed that participants assigned to gender-responsive intervention who also reported experiencing prior physical/sexual abuse, demonstrated reduced odds of depression and rates of substances used. To contrast, those not reporting prior abuse but assigned to gender-responsive treatment demonstrated greater odds of depression and substances used (Saxena et al., 2014). Gender-responsive intervention seemed to negatively moderate the relationship between prior experiences of abuse and the outcomes of interest. These results suggest that targeting the specific needs of the participants is crucial for effective intervention, and that focusing on needs that are not problematic may not be beneficial for treatment (Saxena, Messina, & Grella et al., 2014).

Indeed, it was the case that the majority of evaluations included in the meta-analysis did not assess participants using validated risk assessment instruments. In fact, there was very little reference made at all to criminogenic risk and evaluating needs prior to intervention. This suggests that interventions may not have been tailored to suit the needs of the participants, potentially reducing the impact of gender-informed interventions over that of gender-neutral interventions. Although gender-responsive

interventions appeared to be no more effective than gender-neutral interventions, effectiveness may have been hindered by a lack of proper assessment and thus, inappropriate assignment to intervention.

In reviewing the literature surrounding the efficacy of substance misuse interventions offered to women offenders, it is evident that very few studies have compared cognitive behavioural interventions with therapeutic community interventions. Despite representing some of the largest subgroups of interventions offered to offender populations (Bahr et al., 2012; Gobeil et al., 2016; Mitchell, Wilson, & MacKenzie, 2007), investigation into differences surrounding the overall effectiveness of these two types of intervention is generally lacking. Therefore, it was of value to conduct subgroup comparisons between CBT and TC interventions to investigate any potential differences in effectiveness. Due to a small number of studies included in the overall meta-analysis, comparisons between CBT and TC interventions were limited to those studies that reported recidivism outcomes. Analyses demonstrated that both types of interventions reduced the odds of recidivism during the follow-up period, however, differences between these two types of intervention were not significant. This suggests that neither type of intervention was superior to the other in its ability to reduce the odds of recidivism among women offenders.

Subgroup analyses for therapeutic community interventions revealed that women participating in TC interventions had a 45% to 49% reduction in the odds of recidivism. This is in keeping with Gobeil et al.'s 2016 meta-analysis of correctional interventions for women offenders, which demonstrated that TC interventions successfully reduced the odds of reoffending. Although this meta-analysis did not focus on substance misuse

interventions specifically, the majority of the TC interventions focused on substance use as a treatment target (Gobeil et al., 2016).

It is suggested that therapeutic community interventions are effective for women in part because they foster an environment conducive to emotional expression and support from others (Wright et al., 2012). This is especially critical in the early stages of intervention, as receiving empathy and validation when sharing experiences can help to empower those participating in therapeutic communities, and create a stronger sense of self in connection to recovery (Covington & Bloom, 2006). Indeed, women offenders participating in substance misuse intervention state that they value opportunities to safely share experiences in an all-women environment. This is especially crucial when interpersonal abuse and victimization are discussed, two factors salient to women offenders (Twaddle, Setpaul, Leon Guerrer, Manibusan, & Riddle, 2008; Van Voorhis, Wright, Salisbury, & Bauman, 2010).

Given a lack of sufficient studies, it was not possible to explore the impact of TC interventions on substance use outcomes. Despite the results suggesting that TC interventions have a positive effect on recidivism outcomes, one cannot assume that TC interventions will have the same impact on substance use outcomes (Mitchell, Wilson, & MaKenzie). Previous research does provide some evidence that TC interventions reduce the use of substances in those diagnosed with a substance misuse disorder. A meta-analysis evaluating the impact of TC interventions on adults diagnosed with co-occurring disorders demonstrated a moderate, consistent reduction in substance use across four studies. Although this meta-analysis did not focus on offender populations specifically, one of the four studies was conducted on a male offender sample (Sacks, McKendrick,

Sacks, & Cleland, 2010). Clearly, there is a need for more research to determine the efficacy of TC interventions in reducing substance use among women offenders.

Assessment of the impact of CBT interventions on recidivism outcomes for women offenders demonstrated that CBT interventions successfully reduced the odds of reoffending. Women participating in CBT interventions for substance misuse had 42% to 54% greater odds of not engaging in criminal activities during the follow-up period. Gobeil et al. (2016) reported similar findings in their meta-analysis of correctional interventions for women, where cognitive behavioural interventions successfully reduced that odds of recidivism by approximately 32%. As mentioned, this meta-analysis did not focus on substance misuse interventions specifically, however half of the CBT-based interventions focused on substance misuse as a treatment target (Gobeil et al., 2016).

Although this study could not evaluate the impact of CBT programs on substance use outcomes, previous research has identified CBT as an effective intervention for reducing substance misuse among women. In a meta-analysis evaluating the impact of CBT on adults diagnosed with a substance use disorder, results indicated that CBT had a small, but significant treatment effect on substance use outcomes. Of more importance was the fact that CBT seemed to provide greater treatment benefits for women diagnosed with substance use disorders than for men (Magill & Roy, 2008).

A noted strength of CBT is its ability to easily incorporate individual treatment needs into the overall CBT framework. Comprehensive assessment, both prior to implementation, as well as throughout the therapeutic process is a key component of CBT, and helps to ensure that all individual treatment needs can be addressed. As previously mentioned, this is an important factor to consider when working with women

offenders, as they often present with unique treatment needs not typically addressed in gender-neutral interventions. This is also in keeping with the principles of effective intervention, which state that interventions should be tailored to the needs of the individual offender by way of structured assessment (Andrews & Bonta, 2010).

Indeed, upon reviewing the treatment curricula of the five CBT interventions included in the subgroup analysis, it became evident that in addition to providing the basic foundations of CBT, each CBT program incorporated gender-responsive features to some degree. Given that comprehensive assessment is noted as a core feature of CBT interventions, one could assume that assessments were conducted prior to assignment to treatment. It is unclear however to what degree participants were assessed prior to treatment implementation, as very few authors made mention of risk or need assessments in the program descriptions, or when discussing assignment to treatment.

A key feature of meta-analysis is that results are the product of the management of the meta-analytic process (i.e., aggregation of effect-sizes), in combination with specific features of the studies included (i.e., study-level characteristics). It is often the case that the scientific rigour of primary studies can vary considerably within a meta-analysis. A common objection of meta-analysis is that they combine results from studies of differing levels of study quality. Therefore, it has become common practice when conducting a meta-analysis to assess study quality, and attempt to control for any potential variations in study quality that may impact the overall effect estimate (La Torre, Chiaradia, Gianfagna, De Laurentis, Boccia, & Ricciardi, 2006).

When it comes to methodology, generally speaking, studies evaluating substance misuse interventions for women offenders tend to be methodology weak. Indeed, when

applying the CODC's guidelines for the evaluation of sexual offender treatment outcome research (CODC, 2007), nearly two-thirds of evaluations were categorized as having a global quality rating of rejected, or low. The remaining evaluations were categorized as having a global quality rating of good. No evaluations met the criteria for a global quality rating of strong. This is in keeping with ratings of study quality among criminal justice research as a whole, where authors concluded that 73% of criminal justice evaluations were rated as having low, or standard levels of methodological rigour, leaving only a small percentage of evaluations assessed as having strong methodological rigour (Weisburd, Lum, & Petrosino, 2001). This speaks to the overall lack of methodological quality implemented within criminal justice research, including but not limited to studies evaluating substance misuse interventions.

A major concern in working with a body of literature that is generally defined as methodologically weak is the potential for overestimation of treatment effects. Research has consistently demonstrated that methodological strength can be correlated with the ability to accurately detect treatment effects (Boruch, Snyder, & DeMoya, 2000; Campbell & Boruch, 1975). Indeed, Weisburg, Lum, & Petrosino (2011) highlighted the inverse relationship that exists between study quality and treatment outcomes among criminal justice research. Using the Maryland Scientific Method Scale to assess the methodological rigour of 308 studies, the authors found that weaker study designs were more likely to report an outcome in favour of treatment, and less likely to report a harmful effect of treatment (Weisburg, Lum, & Petrosino, 2011).

The last subgroup analysis focused on methodological strength and the impact of study quality on overall treatment success. Given a lack of sufficient data, comparisons

were limited to those studies that reported recidivism outcomes. Comparisons between interventions identified as having weak methodology (overall quality ratings of rejected or low) versus those identified as having stronger methodology (overall quality ratings of good) revealed no significant differences in reduction of recidivism odds. Studies identified as having stronger methodological quality were not significantly better at reducing the odds of recidivism for women offenders, when compared to those with weaker methodology. This suggests that lower rates of reoffending observed among women offenders participating in any kind of intervention was largely robust to methodological variations. Substance misuse intervention seems to be effective, regardless of methodology.

Strengths and Limitations

The results of this meta-analysis contribute important information to the existing body of literature by being the first to use meta-analytic techniques to evaluate the impact of substance misuse interventions on both recidivism and substance misuse outcomes for women offenders. In addition, by running both fixed-effect and random-effects models, measures of between-study variation could be obtained and the resulting variation explored via subgroup analyses. More specifically, this study was also the first to compare CBT and TC interventions for women offenders with substance misuse disorders through subgroup analyses. Given that the majority of correctional interventions for women offenders with substance misuse disorders are either TC or CBT based interventions, it was of importance to evaluate the effectiveness of these interventions, both separately and in contrast with each other. Despite finding no significant difference between these two types of interventions, both TC and CBT

interventions successfully reduced the odds of recidivism for women offenders participating in these programs. This provides support for the efficacy of both CBT and TC interventions in reducing recidivism outcomes among women offenders with substance misuse disorders.

In addition to noting the strengths of this study, it is also of important to keep in the mind the limitations of meta-analytic techniques. First, it is important to discuss the issue of publication bias. A critical concern when summarizing a body of literature into an overall effect is the prospect of missing evaluations. It is widely known among the research community that studies with significant results are more likely to be published than those with negative or null results. Therefore, it is possible that evaluations may be missed by conventional search methods, as the majority of evaluations are drawn from peer-reviewed journals found in academic databases.

Given that the large majority of the evaluations included in this review were peer-reviewed publications, it is possible that the overall effect estimates may not accurately reflect all evaluations conducted in the field. With such a large number of peer-reviewed publications contributing to the overall effect estimates, it is possible that this bias may skew the overall effect estimates in favour of the treatment group. These estimates may not reflect what has been found in unpublished evaluations that may have been unaccounted for in the review. Thus, it is important to keep in mind how publication status may influence the overall effect estimate, and whether this estimate is in fact, an accurate representation of the body of literature under review.

Second, this approach may lead to an aggregation of results from interventions that are in fact, not similar to each other. This is commonly referred to as the apples and

oranges problem (Lipsey & Wilson, 2000). Indeed, analyses demonstrated large amounts of unexplained variability across studies.

Ideally, this variability would be explored using study level moderators in an attempt to identify potential sources of variability. However, information reported by authors with regard to study characteristics, participant characteristics, and intervention details varied considerably across evaluations, making it nearly impossible to investigate potential moderating variables.

In addition, this study was also limited by a lack of sufficient evaluations to conduct subgroup analysis with regards to substance misuse outcomes. Although previous research indicates that substance misuse intervention programs can be successful in reducing substance use outcomes in samples other than women offenders (Magill & Roy, 2008; Sacks, McKendrick, Sacks, & Cleland, 2010), it was not feasible to investigate the impact of specific types of interventions on substance misuse outcomes.

When discussing significant amounts of heterogeneity across studies, it is important to note how this extreme variation may impact the generalizability of the results. Given that the underlying principle of the fixed-effects model is that studies are homogenous with regards to intervention and participant characteristics, the fixed-effects model may not be appropriate to describe the overall effects in this case. Given that there were significant amounts of variability for each of the overall effect estimates, the random-effects model would better fit the characteristics of the sample of evaluations, as this model assumes that studies will vary with regards to between-study characteristics.

Further, one cannot discuss generalizability without addressing the implications of combining the results of evaluations from different cultural groups. Given that this meta-

analysis included studies from both North American (Canada and the United States of American) and Europe, it is important to discuss how combining the outcomes of these evaluations may limit the applicability of the results. Although the majority of the evaluations were conducted in North America, two of the evaluations were conducted in Spain. This becomes especially problematic with regards to the operationalization of outcome measures, such as recidivism.

Indeed, it is often the case that countries have different ways of defining recidivism (Fazel & Wolf, 2015). With different laws in place across countries that dictate what is considered to be criminal behaviour, it is possible that the outcome measure of recidivism may be operationalized in different ways across countries and cultures. This may result in different rates of recidivism across the countries represented within this meta-analysis (USA, Canada and Spain), rendering the overall effect estimate for recidivism an under-estimate or over-estimate.

The majority of the evaluations included in this review were conducted in the USA ($k=17$) or Canada ($k=3$). Research suggests that recidivism rates are similar across Canada and the USA when recidivism was defined as both a reconviction and a reimprisonment (Fazel & Wolf, 2015). Despite differences in the operationalization of recidivism, rates seem to be stable across North America. It is unclear however how these rates map onto those established in Spain and other countries outside of North America. Therefore, it is important to take caution when applying these results to countries outside of North America and cultures that differ from those represented across Canada and the USA.

Assessments of the methodological quality of the studies included in this meta-analysis underscores the need for more high-quality evaluations to be conducted. Future

meta-analyses will benefit from continued work in the field to produce higher-quality evaluations that consistently report information on moderating variables, allowing for more meaningful exploration of heterogeneity across studies. In doing so, key factors that may influence the efficacy of interventions can be identified and addressed, resulting in greater treatment efficacy, and ideally, greater reductions in recidivism and substance misuse outcomes.

Specific ways in which overall study quality can be improved include defining treatment and comparison groups more clearly and consistently, and ensuring that any pre-existing differences among participants, such as risk for recidivism or other factors potentially related to recidivism, are controlled for. In addition, evaluations can be improved by ensuring adequate sample sizes for both treatment and comparison groups (50 participants for each group), and follow-up periods of at least three years. Lastly, greater attention must be paid to miscellaneous incidental factors that may impact the efficacy of an intervention, such as differences in the conditions and delivery of intervention, or the provision of services. By incorporating these aspects into the study design, researchers can produce higher-quality evaluations with fewer sources of variability between evaluations.

Conclusion

Results of this meta-analysis demonstrate that correctional interventions for substance misuse are effective in reducing both recidivism and substance use among women offenders with substance misuse disorders. Preliminary evidence suggests that intervention programs that allow for the integration of treatment needs, whether they be needs specific to women offenders, or to the broader offender population, are important

to the successful rehabilitation of women offenders. Therefore, it is in the best interest of policy makers to continue to support and implement correctional interventions that are comprehensive, and address the wide array of treatment needs often reported by women offenders. As more high-quality research is produced, prospective meta-analyses should aim to build upon these results, by incorporating more detailed information regarding potential moderating variables, such as participant and intervention characteristics. In doing so, the variability that exists between evaluations can be explored further, and factors that may influence the efficacy of interventions may be addressed, to increase the overall positive impact of substance misuse interventions on recidivism and substance use outcomes for women offenders.

References

- Anderson, V.R., Walerych, B.M., Campbell, N.A., Barnes, A.R., Davidson, W.S., Campbell, C.A., Onifade, E., Pertersen, J.L. (2016). Gender-responsive intervention for female juvenile offenders: A quasi-experimental outcome evaluation. *Feminist Criminology*, 1-21.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- Andrews, D., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy and Law*, 16, 39-55. doi: 10.1037/a0018362
- Andrews, D., & Bonta, J. (2017). *The psychology of criminal conduct*. New York, NY:Routledge.
- Bahr, S., Masters, A., & Taylor, B. (2012). What works in substance abuse treatment programs for offenders? *The Prison Journal*, 92, 155-174. doi: 10.1177/0032885512438836
- Bloom, B. O., Owen, B., & Covington, S. (2003). *Gender-responsive strategies: Research, practice, and guiding principles for women offenders*. Washington, DC: U.S. Department of Justice.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta analysis (computer program)*. Englewood, NJ: Biostat.
- Boruch, R., Snyder, B., & DeMoya, D. (2000). The importance of randomized field trials. *Crime & Delinquency*, 46, 156-180. doi: 10.1177/0011128700046002002
- Brennan, T., Dieterich, M., Salisbury, E.J., & Van Voorhis, P. (2012). Women's pathways to serious and habitual crime: A person-centered analysis incorporating gender responsive factors. *Criminal Justice and Behavior*, 39, 1481-1508. doi: 10.1177/0093854812456777
- Brown, S., & Motiuk, L. (2008). *Using dynamic risk factors to predict recidivism in a sample of male and female offenders* [PDF document]. Retrieved from https://carleton.ca/gendercrimelab/wp-content/uploads/CPA_2008-DFIA.pdf

- Brown, S.L. & Motiuk, L.L. (2005). The Dynamic Factor Component of the Offender Intake and Assessment Process: A psychometric, meta-analytic and field review (R-164). Research Branch, Correctional Service of Canada, Ottawa: Ontario
- Campbell, D.P., & Boruch, R.F. (1975). Making the case for randomized assignment to treatments by considering the alternatives: Six ways in which quasi-experimental evaluations in compensatory education tend to underestimate effects. *Evaluation and Experiment: Some Critical Issues in Assessing Social Programs*, ed. New York: Academic Press.
- Canada., & Beech, A. (2007). The Collaborative Outcome Data Committee's guidelines for the evaluation of sexual offender treatment outcome research: Part 2 : CODC guidelines. Ottawa: Public Safety Canada.
carleton.ca/viewdoc.html?id=/ebooks/ebooks2/springer/2011-11-21/1/9781441994707
- Carrol, K.M. (1998). *Therapy manual for drug addiction series. A cognitive behavioral approach, treating cocaine addiction*. National Institute on Drug Abuse, National Institute of Health Publication No. 98-4308, USA.
- Carroll, K. M. (2000). *A cognitive-behavioral approach: Treating cocaine addiction* (NIH Publication No. 00-4308). Retrieved from <https://archives.drugabuse.gov/sites/default/files/cbt.pdf>
- Chandler, R., Fletcher, B., & Volkow, N. (2009). Treating drug abuse and addiction in the criminal justice system: Improving public health and safety. *The Journal of the American Medical Association*, 301, 183-190. doi: 10.1001/jama.2008.976
- Chinn, S. (2000). A simple method for converting an odds ratio to an effect size for use in meta-analysis. *Medicine*, 19, 3127-3131. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/11113947>
- Corcoran, M., & Corcoran, J. (2001). Retrospective reasons for the initiation of substance use. *Journal of Ethnic and Cultural Diversity in Social Work*, 10, 69-83. doi: 10.1300/J051v10n01_04
- Covington, S.S., & Bloom, B.E. (2008). Gender responsive treatment and services in correctional settings. *Women & Therapy*, 29, 9-33. doi: 10.1300/J015v29n03_02

- Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *Review of Law and Women's Studies*, 2, 11-52.
- Djikers, M. (2015). What is a scoping review? *KT Update*, 4, 1-5. Retrieved from <http://ktdrr.org/products/update/v4n1/>
- Dorsey, T.L., & Middleton, P. (2005). *Drugs and crime facts*. Retrieved from <http://www.bjs.gov/content/pub/pdf/DCF.pdf>
- Dowden, C., & Andrews, D. (1999). What works with female offenders: A meta-analytic review. *Crime & Delinquency*, 45, 438-452. doi:<https://doi.org/10.1177/0011128799045004002>
- Dowden, C., & Blanchette, K. (2002). An evaluation of the effectiveness of substance abuse programming for female offenders. *International Journal of Offender Therapy and Comparative Criminology*, 46, 220-230. doi: 10.1177/0306624X02462007Evaluation_Report.pdf
- Dowden, C., & Brown, S. L. (2002). The role of substance abuse factors in predicting recidivism: A meta-analysis. *Psychology, Crime, and Law*, 8, 243-264. doi: 10.1080/10683160208401818
- Gehring, K.S., & Van Voorhis, P. (2014). Needs and pretrial failure: Additional risk factors for female and male pretrial defendants. *Criminal Justice and Behavior*, 41, 943-970. doi: 10.1177/0093854814538022
- Giacomazzi, A., & Bell, V. (2007). Drug court program monitoring. *Criminal Justice Policy Review*, 18, 294-312. doi: 10.1177/0887403407301494
- Gobeil, R., Blanchette, K., & Stewart, L. (2016). A meta-analytic review of correctional interventions for women offenders. Gender-neutral versus gender-informed approaches. *Criminal Justice and Behavior*, 43, 301-322. doi: 10.1177/0093854815621100
- González-Menéndez, A., Fernández, P., Rodríguez, F., & Villagrà, P. (2014). Long-term outcomes of Acceptance Commitment Therapy in drug-dependent female inmates: A randomized control trial. [gov/TXManuals/CBT/CBT1.html](http://www.txmanuals.com/CBT/CBT1.html)
- Grant, B. & Gileno, J. (2008). *The Changing Federal Offender Population* Ottawa: Correctional Service Canada.

- Greenfeld L.A., Snell T.L. Women offenders. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, 1999. Retrieved from <https://www.bjs.gov/content/pub/pdf/wo.pdf>
- Haidich, A.E. (2010). Meta-analysis in medical research. *Hippokratia*, 14 29-37. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049418/>
- Hall, E. A., Prendergast, M., Wellisch, J., Patten, M., & Cao, Y. (2004). Treating drug-abusing women prisoners: An out-comes evaluation of the Forever Free program. *The Prison Journal*, 84, 81-105. doi: 10.1177/0032885503262456
- Hall, M.T., Golder, S., Conley, C.L., & Sawning, S. (2013). Designing programming and interventions for women in the criminal justice system. *American Journal of Criminal Justice*, 38, 27-50. doi: 10.1007/s12103-012-9158-2
- Harrell, A., Roman, J., & Sack, E. (2001). *Drug court services for female offenders, 1996-1999: Evaluation of the Brooklyn Treatment Court*. Washington, DC: The Urban Institute. Retrieved from http://www.urban.org/uploadedPDF/410356_Brooklyn-Treatment-Court.pdf
- Harrison, P. M., & Beck, A. J. (2002). Bureau of Justice Statistics bulletin. Prisoners in 2001 (NCJ 195189). Retrieved from <http://www.ojp.usdoj.gov/bjs/>
- Hegdes, L.V., & Vevea, J.L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486-504. doi: 10B2-989X/98/J3.00
- Heffernan, E.B., Finn, J., Saunders, J.B., & Byrne, G. (2003). Substance-use disorders and psychological distress among police arrestees. *The Medical Journal of Australia*, 179, 408-411. Retrieved from <https://www.mja.com.au/journal/2003/179/8/substance-use-disorders-and-psychological-distress-among-police-arrestees>.
- Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Huedo-Medina, T., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *CHIP Documents*, 19. Retrieved from http://digitalcommons.uconn.edu/chip_docs/

- Johnson, J. E., Friedmann, P. D., Green, T. C., Harrington, M., & Taxman, F. S. (2011). Gender and treatment response in substance use treatment-mandated parolees. *Journal of Substance Abuse Treatment, 40*, 313-321. doi: 10.1016/j.jsat.2010.11.013
- Kelly, L. & Farrell MacDonald, S. *Comparing lifetime substance use patterns of men and women offenders*. (RIB-14-44). Ottawa, ON: CSC.
- Koehler, J., Humphreys, D., Akoensi, T., Sánchez de Ribera, & Losel, F. (2014). A systematic review and meta-analysis on the effects of European drug treatment programmes on reoffending. *Psychology, Crime & Law, 20*, 584-602. doi: 10.1080/1068316X.2013.804921
- La Torre, G., Chiaradia, G., Gianfagna, F., De Laurentis, A., Boccia, S., & Ricciardi, W. (2006). Quality assessment in meta-analysis. *Italian Journal of Public Health, 3*, 44-50. doi: 10.2427/5937
- Leukefeld, C.G., Gullotta, T.P., & Gregrich, J. (2011). Evidence based substance abuse treatment in criminal justice settings. Available from <http://books1.scholarsportal.info.proxylibrary>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis: Applied social research method series*. (Vol. 49). Thousand Oaks, CA: SAGE.
- Lobmann, R., & Verthein, U. (2009). Explaining the effectiveness of heroin-assisted treatment on crime reductions. *Law and Human Behavior, 33*, 83-95. doi: 10.1007/s10979-008-9138-8
- MacSwain, M.-A., Farrell MacDonald, S., & Cheverie, M. (2015). Post-release outcomes of Methadone Maintenance Treatment Program (MMTP) participants: A comparative study (R-322). Ottawa, Ontario: Correctional Service of Canada. Retrieved from <http://www.csc-scc.gc.ca/research/005008-0322-eng.shtml>
- Magill, M., & Ray, L. A. (2009). Cognitive-behavioral treatment with adult alcohol and illicit drug users: A meta-analysis of randomized controlled trials. *Journal of Studies in Alcohol and Drugs, 70*, 516-527. doi:
- Matheson, F. I., Doherty, S., & Grant, B. A. (2009). Women offender substance abuse programming & community reintegration (R-202). Ottawa, Ontario: Correctional

- Service of Canada. Retrieved from <http://www.csc-scc.gc.ca/research/092/r202-eng.pdf>
- McHugh, M.L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22, 276-82. doi: 10.11613/BM.2012.031
- McLellan, A.T., Kushner, H., & Metzger, D. (1992). The fifth edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment*, 9, 199-213. doi: 10.1016/0740-5472(92)90062-S
- Messina, N., Burdon, W., & Prendergast, M. (2006). Prison-based treatment for drug-dependent women offenders: Treatment versus no treatment. *Journal of Psychoactive Drugs*, 3, 333-343. doi: 10.1080/02791072.2006.10400597
- Messina, N., Grella, C. E., Cartier, J., & Torres, S. (2010). A randomized experimental study of gender-responsive substance abuse treatment for women in prison. *Journal of Substance Abuse Treatment*, 38, 97-107. doi: 10.1016/j.jsat.2009.09.004
- Messina, N., Grella, C., & Prendergast, M. (2007). Childhood adverse events and current traumatic distress: A comparison of men and women drug-dependent prisoners. *Criminal Justice and Behavior*, 34, 1385-1401. doi: 10.1177/0093854807305150
- Miller, P. (2010). The impact of prison-based substance abuse treatment on rates of recidivism among female offenders. (Unpublished doctoral dissertation). Capella University, Minneapolis, MN. Available from <https://search-proquest-com.proxy.library.carleton.ca/docview/518695601?pq-origsite=summon>
- Mitchell, O., Wilson, D., & MacKenzie, D. (2007). Does incarceration-based drug treatment reduce recidivism? A meta-analytic synthesis of the research. *Journal of Experimental Criminology*, 3, 353-375. doi: 10.1007/s11292-007-9040-2
- Moher, D, Cook, D, Jadad, A, et al. Assessing the quality reports of randomised trials: Implications for the conduct meta-analyses. *Health Technology Assessment* 1999; 3(12): 1-98.
- Morrissey, J.P., Ellis, A.R., Gatz, M., Amaro, H., Glover Reed, B., Savage, A., Finkelstein, N., Mazelis, R., Brown, V., Jackson, E. W., & Banks, S. (2005). Outcomes for women with co-occurring disorders and trauma: Programs and

- person-level effects. *Journal of Substance Abuse Treatment*, 28, 121-133. doi: 10.1016/j.jsat.2004.08.012
- Mosher, C., & Phillips, D. (2006). The dynamics of a prison-based therapeutic community for women offenders: Retention, completion, and outcomes. *The Prison Journal*, 86, 6-31. doi: 10.1177/0032885505283871
- Newbury-Birch, D., Coulton, S., Bland, M., Cassidy, P., Dale, V., Deluca, P., Gilvarry, E., Godfrey, C., Heather, N., McGovern, E.K.R., Myles, J., Oyefeso, A., Parrott, S., Patton, R., Perryman, K., Phillips, T., Shepherd, J., & Drummond, C. (2014). Alcohol Screening and Brief Interventions for Offenders in the Probation Setting (SIPS Trial): a Pragmatic Multicentre Cluster Randomized Controlled Trial. *Alcohol and Alcoholism*, 49, 540-548. doi: <https://doi.org/10.1093/alcalc/agu046>.
- Pelissier, B. M. M., Camp, S. D., Gaes, G. G., Saylor, W. G., & Rhodes, W. (2003). Gender differences in outcomes from prison-based residential treatment. *Journal of Substance Abuse Treatment*, 24, 149-160. doi: 10.1016/S0740-5472(02)00353-7
- Peters, R.H., Strozier, A.L., Murrin, M.R., & Kearns, W.D. (1997). Treatment of substance-abusing jail inmates: Examination of gender differences. *Journal of Substance Abuse Treatment*, 14, 339-349. doi: [https://doi.org/10.1016/S0740-5472\(97\)00003-2](https://doi.org/10.1016/S0740-5472(97)00003-2)
- Prendergast, M. (2009). Interventions to promote successful re-entry among drug-abusing parolees. *Addiction Science & Clinical Practice*, 5, 4-13. doi: 10.1151/ascp09514
Research%20Design%20Affect%20Study%20Outcomes.pdf
- Robbins, C. A., Martin, S. S., & Surratt, H. L. (2009). Substance abuse treatment, anticipated maternal roles, and re-entry success of drug-involved women prisoners. *Crime & Delinquency*, 55, 388-411. doi: 10.1177/0011128707306688
- Sacks, J.Y. (2004). Women with co-occurring substance use and mental disorders (COD) in the criminal justice system: A research review.
- Sacks, J. Y., McKendrick, K., & Hamilton, Z. (2012). A randomized clinical trial of a therapeutic community treatment for female inmates: Outcomes at 6 and 12 months after prison release. *Journal of Addictive Diseases*, 31, 258-269. doi: 10.1080/10550887.2012.694601

- Sacks, S., McKendrick, K., Sacks, J.Y., & Cleland, C.M. (2010). Modified therapeutic community for co-occurring disorders: Single investigator meta-analysis, *Substance Abuse*, 31, 146-161, DOI: 10.1080/08897077.2010.495662
- Sacks, J.Y., Sacks, S., McKendrick, K., Banks, S., Schoenberger, M., Hamilton, Z., Stommel, J., & Shoemaker, J. (2008). Prison therapeutic community treatment for female offenders: profiles and preliminary findings for mental health and other variables (crime, substance use and HIV risk). *Probation and Parole: Current Issues*, 233-251. doi: doi:10.1080/10509670802143680
- Saxena, P., Messina, N. J., & Grella, C. E. (2014). Who benefits from gender-responsive treatment? Accounting for abuse history on longitudinal outcomes for women in prison. *Criminal Justice and Behavior*, 41, 417-432.
- Scott, C. K., & Dennis, M. L. (2012). The first 90 days following release from jail: Findings from the Recovery Management Checkups for Women Offenders (RMCWO) experiment. *Drug and Alcohol Dependence*, 125, 110-118. doi: <http://dx.doi.org/10.1016/j.drugalcdep.2012.03.025>
- Scott, C.K., Dennis, M.L., & Lurigio, A.J. (2017). The effects of specialized probation management checkups (RMCs) on treatment participation, substance use, HIV risk behaviors, and recidivism among female offenders: Main findings of a 3-year experiment using subject by intervention analysis. *Journal of Experimental Criminology*, 13, 53-77. doi: 10.1007/s11292-016-9281-z
- Scott, C.K., Grella, C.E., Dennis, M.L., & Funk, R.R. (2014). Predictors of recidivism over 3 years among substance-using women released from jail. *Criminal Justice and Behavior*, 41, 1257-1289
- Sherman, L.W., Gottfredson, D.C., MacKenzie, D.L., Eck, J., Rueter, P., & Bushway, S.D. (1998). *Preventing crime: What works and what doesn't what's promising*. Washington, D.C: National Institute of Justice, U.S., Department of Justice. Retrieved from <https://www.ncjrs.gov/pdffiles/171676.pdf>
- Stein, M.D., Caviness, C.M., Anderson, B.J., Hebert, M., & Clarke, J.G. (2010). A brief alcohol intervention for hazardously drinking incarcerated women. *Addiction*, 105, 466-475. doi: doi:10.1111/j.1360-0443.2009.02813.x

- The National Center on Addiction and Substance Abuse at Columbia University (NCASA). (2010). *Behind bars II: Substance abuse and America's prison population*. Retrieved from <http://www.centeronaddiction.org/addiction-research/reports/substance-abuse-prison-system-2010>
- The Sentencing Project. (2007). *Women in the criminal justice system: Briefing sheets*. Retrieved from <http://www.sentencingproject.org/publications/women-in-the-criminal-justice-system-briefing-sheets/>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 1-10. doi: 10.1186/1471-2288-8-45
- Tripodi, S.J., Ledsoe, S.E., Kim, J.S., & Bender, K. (2011). *Research on Social Work Practice*, 21, 15-31. doi: 10.1177/1049731509352337
- Twaddle, I. K. B., Setpaul, R., Guerrero, V. E. L., Manibusan, A. I., & Riddle, J. A. (2006). Countering correctional discourse: Development of a feminist support group for women prisoners in Guam. *Women & Therapy*, 29, 215-237.
- United Nations Office on Drugs and Crime. (2012). *World Drug Report*. Retrieved from https://www.unodc.org/documents/data-and-analysis/WDR2012/WDR_2012_web_small.pdf
- Van Voorhis, Wright, Salisbury, & Bauman. (2010). Women's risk factors and their contributions to existing risk/needs assessment: The current state of a gender-responsive supplement. *Criminal Justice and Behavior*, 37, 261-288.
- Villagr  Lanza, P., Fern ndez Garc a, P., Rodr guez Lamelas, F., & Gonz lez-Men ndez, A. (2014). Acceptance and commitment therapy versus cognitive behavioral therapy in the treatment of substance use disorder with incarcerated women. *Journal of Clinical Psychology*, 00, 1-14. doi: 10.1002/jclp.22060
- Villagr  Lanza, P., & Gonz lez-Men ndez, A. (2013). Acceptance and Commitment Therapy for drug use in incarcerated women. *Psicothema*, 25, 307-312. doi: 10.7334/psicothema2012.292
- Waid, C. A. (2012). An assessment of substance abuse treatment programs in Florida's correctional institutions for women (Unpublished doctoral dissertation). The

- Florida State University, Tallahassee. Available at <https://search-proquest-com.proxy.library.carleton.ca/docview/871903118?pq-origsite=summon>
- Watson, L., Adkins, G., Cook, M. D., & Stageberg, P. (2010). *Process and outcome evaluation of the STAR (Sisters Together Achieving Recovery) program*. Des Moines: Iowa Department of Human Rights Division of Criminal and Juvenile Justice Planning. Retrieved from [http://www.humanrights.iowa.gov/cjjp/Images/pdf/STAR_Weekes, J., Moser, A., Ternes, M., & Kunic, D. \(2009\). *Substance abuse among male offenders* \(RS 09-02\). Ottawa, ON: Correctional Service Canada.](http://www.humanrights.iowa.gov/cjjp/Images/pdf/STAR_Weekes, J., Moser, A., Ternes, M., & Kunic, D. (2009). Substance abuse among male offenders (RS 09-02). Ottawa, ON: Correctional Service Canada.)
- Weisburd, D., Lum, C. M., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice. *Annals of the American Academy of Political and Social Sciences*, 578, 50–70. Retrieved from [http://cebcp.org/wp-content/publications/Does%20World Health Organization. \(2009\). *World Health Statistics 2009*. Retrieved from \[http://www.who.int/gho/publications/world_health_statistics/EN_WHS09_Full.pdf\]\(http://www.who.int/gho/publications/world_health_statistics/EN_WHS09_Full.pdf\)](http://cebcp.org/wp-content/publications/Does%20World Health Organization. (2009). World Health Statistics 2009. Retrieved from http://www.who.int/gho/publications/world_health_statistics/EN_WHS09_Full.pdf)
- Wright, E.M., Van Voorhis, P., Salisbury, E.J., & Bauman, A. (2012). Gender-responsive lessons learned and policy implications for women in prison. *Criminal Justice & Behavior*, 39, 1612-1632.
- Zlotnick, C., Johnson, J., & Njavits, L. M. (2009). Randomized controlled pilot study of cognitive-behavioral therapy in a sample of incarcerated women with substance use disorder and PTSD. *Behavior Therapy*, 40, 325-336. doi: 10.1016/j.beth.2008.0

Appendix A: Data Extraction Form

Data Extraction Form

Study Identification Number:

Study Details

Methodological features

Study Title:

Author(s):

Year of publication:

Publication status: (1) Published
(0) Unpublished

Type of publication: (1) Peer reviewed article
(2) Non-peer reviewed article
(3) Dissertation
(4) Conference presentation
(5) Other _____

Study methodology:

Identify the method used to assign participants to groups. Methods are described as follows:

Random or quasi-random

- The entire sample is matched or blocked first, then assigned to treatment and comparison groups within pairs or blocks.
- Participants are randomly assigned to treatment and comparison groups without matching.
- Quasi-random procedure presumed to produce comparable groups, such as individuals randomly assigned by some naturally occurring process (Polanin & Tanner-Smith, 2014).

Non-random but matched

- Matched or statistically controlled only on pre-test measures of some or all variables used as outcomes measures.
- Matched or statistically controlled on pretest measures AND other personal characteristics, such as demographics.
- Matched or statistically controlled only on demographics (Polanin & Tanner-Smith, 2014).

Non-random, no matching

- Non-random, not matched, but pre-treatment equivalence information is available (Polanin & Tanner-Smith, 2014)

- (1) Random/ quasi-random
- (2) Non-random but matched assignment
- (3) Non-random, no matching
- (4) Other: _____

Outcome(s):

Indicate the primary outcome measure(s) reported in the study.

Relapse is defined as a return to substance use after a period of abstinence and can be measured using urinalysis testing (Gonzalez-Menendez, Fernandez, Rodriguez & Villagra, 2014), self-report using an established substance abuse screening tool such as the AUDIT (Newbury-birch et al., 2014), or a combination of measures. Recidivism is defined as a return to criminal behaviour after an offence-free period. Recidivism is understood to be either a formal institutional measure, such as rearrest, reconviction and reincarceration, or as a self-reported measure of reoffending (Koehler, Humphreys, Akoensi, Sanchez de Ribera, & Losel, 2014).

- (1) Relapse
- (2) Recidivism
- (3) Both

Sample size:

Indicate the total sample size for the study (including both intervention and control groups). If the sample includes both men and women, please indicate the size of the sample for women-only.

N= _____

Participant Characteristics

Age (average in years): _____

Ethnicity:

Indicate the ethnic background that represents the largest majority of the sample.

- (1) Caucasian
- (2) Black
- (3) Indigenous
- (4) Hispanic
- (5) Asian

- (6) Other _____
- (7) Not reported

Security level

Indicate the security level of the institution that houses the sample. If the sample if was recruited from the community (i.e., residential treatment) choose “not applicable.”

- (1) Minimum
- (2) Medium
- (3) Maximum
- (4) Multi-level
- (5) Other _____
- (6) Not applicable (i.e. community sample)
- (7) Not reported

Reason for participation:

Indicate the reason for participation in treatment.

- (1) Voluntary
- (2) Mandated (i.e., drug treatment court)
- (3) Recommended (i.e., by a physician)
- (4) Mixed
- (5) Other: _____
- (6) Not reported

Concurrent mental health diagnosis:

Indicate the proportion of the sample that were diagnosed as having a concurrent mental health diagnosis at the time of the study (current) as recognized by the Diagnostic and Statistical Manual Mental Disorders Fifth Edition (American Psychological Association, 2000) or any other widely recognized diagnostic tool (i.e., Mini International Psychiatric Interview). If the author(s) did not report on this information, choose “not reported.”

Percentage: _____

Not reported

Most prevalent diagnosis:

Indicate the most prevalent concurrent mental health diagnosis (current) for the sample as recognized by the Diagnostic and Statistical Manual Mental Disorders Fifth Edition (American Psychological Association, 2000). If the author(s) did not report on this information, choose “not reported.”

- (1) Mood disorder
- (2) Personality disorder
- (3) Schizophrenia
- (4) OCD
- (5) PTSD
- (6) Not reported
- (7) Other: _____

Trauma (Lifetime):

Indicate the percentage of the sample that reported experiencing trauma in their lifetime.

Trauma

is defined as the experience of potentially traumatic events including crime, general disaster, and various

forms of abuse (Trauma History Questionnaire, 1996; NorVold Abuse Questionnaire, 2003). If the

author(s) did not report on this information, choose “not reported.”

Percentage: _____

Not reported

Most common abuse reported:

Indicate the type(s) of trauma reported by the sample. If the author(s) did not report on this information, choose “not reported.” If more than one type of trauma was reported, check all that apply.

- Crime related events (i.e. robbery, mugging, break and enter)
- General disaster and trauma (i.e. serious injury, natural disaster, unexpected death)
- Physical abuse
- Sexual abuse
- Emotional abuse
- Neglect
- Not reported
- Other: _____

Trauma (Childhood):

Indicate the percentage of the sample that reported experiencing trauma in their childhood (prior to the age of 18). Trauma is defined as the experience of potentially traumatic events including crime, general disaster, and various forms of abuse (Trauma

History Questionnaire, 1996; NorVold Abuse Questionnaire, 2003). If the author(s) did not report on this information, choose "not reported."

Percentage: _____

Not reported

Most common abuse reported:

Indicate the type(s) of trauma reported by the sample. If the author(s) did not report on this information, choose "not reported." If more than one type of trauma was reported, check all that apply.

Crime related events (i.e. robbery, mugging, break and enter)

General disaster and trauma (i.e. serious injury, natural disaster, unexpected death)

Physical abuse

Sexual abuse

Emotional abuse

Neglect

Not reported

Other: _____

Intervention details

Type of intervention:

Indicate the type of intervention that was administered to the treatment group. If components of multiple interventions were combined to create an intervention program, indicate all types of interventions included in the program.

Cognitive-behavioural therapy

Mindfulness-based intervention

Pharmacological intervention (i.e., methadone maintenance therapy, buprenorphine, Antabuse)

12-step program (i.e., alcoholic anonymous, narcotics anonymous)

Therapeutic community

Contingency management

Drug-treatment court

Boot camp

Other: _____

Pharmacological Intervention:

Indicate the type of pharmacological intervention administered to the intervention group.

If the intervention was not pharmacological in nature, choose "not applicable."

- (1) Methadone maintenance therapy
- (2) Buprenorphine
- (3) Naltrexone
- (4) Antabuse
- (5) Other: _____
- (6) Not applicable

Country:

Indicate the country where the intervention was administered: United States

Location of intervention:

Indicate the location where the intervention was administered to the intervention group.

- (1) Institution
- (2) Community
- (3) Combined
- 4) Not reported
- (5) Other: _____

Level of intervention:

Indicate the level of treatment delivery to the intervention group.

- (1) Individual level treatment
- (2) Group level treatment
- (3) Mixed

Implementers:

Indicate who implemented the intervention program to the intervention group.

- Evaluator/researcher
- Other researchers
- Practitioners
- Volunteers
- Technology

- Other: Department of corrections; Addictions Research and Treatment Services
- Not reported

Primary drug of abuse

Indicate the substance most commonly cited as the primary drug of abuse.

- (1) Alcohol
- (2) Cannabis
- (3) Cocaine
- (4) Amphetamines
- (5) Opioids
- (6) Other: _____

Length of intervention

Indicate the length of the intervention program in months: _____

Intervention sessions:

Indicate the total number of intervention sessions included in the intervention program.

Number of sessions: _____

Frequency of intervention sessions:

Indicate how frequently the intervention is administered to the intervention group in a given week.

Frequency of sessions: _____

Duration of intervention sessions:

Indicate the duration of each intervention session in minutes.

Duration: _____

Dosage (Pharmacological interventions)

Indicate the treatment dosage if the intervention was previously identified as a pharmacological intervention. If the intervention was behaviour-based, indicate "not applicable"

Dosage: _____

Not applicable

Gender informed:

Indicate if the intervention was gender-informed. Gender informed interventions can be defined as interventions that explicitly consider needs that are particularly salient to

women, in addition to traditional elements of evidence-based practices (Gobeil, Blachette, & Stewart, 2016). Interventions should be coded as having clear evidence if the authors specifically stated that the program was gender-informed, or if the intervention included a number of gender-informed concepts, including gendered/unique pathways to offending, relational-cultural theory, feminist theories, trauma and parenting. Interventions were coded as having some/partial evidence if the authors made mention of a single one of these concepts, and/or suggested the intervention was suited to fit women's needs (Gobeil, et al., 2016).

- (1) No evidence of gender-responsive intervention
- (2) Some/partial evidence
- (3) Clear evidence of gender-responsive intervention

Control Condition Details

Control condition

Indicate what type of treatment, if any, the control group received.

- No alternative treatment
- "Treatment as usual"
- Waitlist control
- Other: CBT

Location of control condition:

Indicate the location where the control group received the intervention.

- (1) Institution
- (2) Community
- (3) Combined
- (4) Unknown
- (5) Other: _____

Control Condition #2 Details

Control condition

Indicate what type of treatment, if any, the control group received.

- "Treatment as usual"
- Alternative treatment
- Waitlist control
- No alternative treatment

Other: _____

Location of control condition:

Indicate the location where the control group received the intervention.

- (1) Institution
- (2) Community
- (3) Combined
- (4) Unknown
- (5) Other: _____

Study Findings

If multiple outcomes were reported (i.e., recidivism and relapse outcomes were reported in the same study), code information for both outcomes.

Recidivism outcomes

Dependent measure

Indicate how recidivism was operationalized in the present study.

- New conviction
- New charge/arrest
- Any return to custody
- Criminal activity (not specified)

Data Source:

Indicate the source of the recidivism data. If the author(s) did not report this information, indicate "not reported"

- Official records
- Self-report
- Not reported
- Other: _____

Follow-up period:

Indicate the length of the follow-up period post-intervention.

- 1 month
- 3 months
- 6 months

Control group #2: Sample size: _____ Mean: _____ SD:

Correlation:

N (tx): _____ N(cont): _____ r: _____

T-test/ F-test:

N (tx): _____ N(cont): _____ t: _____ F: _____

ROC areas:

AUC: _____

Odds ratios:

Odds ratio: _____ Confidence interval: _____ p: _____

Odds ratios:

Odds ratio: _____ Confidence interval: _____ p: _____

Odds ratios:

Odds ratio: _____ Confidence interval: _____ p: _____

Relapse outcomes

Data source:

Indicate the source of the relapse data. If the author(s) did not report this information, indicate "not reported"

- (1) Self-report
- (2) Urinalysis
- (3) Structured clinical interview
- (4) Other: _____

Follow-up period

Indicate the length of the follow-up period post-intervention.

- 1 month
- 3 months
- 6 months
- 12 months
- 18 months

- 24 months
- Other: _____

Sample sizes, means, and standard deviations:

Outcome 1:

Treatment group: Sample size: _____ Mean: _____ SD: _____
 _____ Control group #1: Sample size: _____ Mean: _____
 SD: _____
 Control group #2: Sample size: _____ Mean: _____ SD: _____

Outcome 2:

Treatment group: Sample size: _____ Mean: _____ SD: _____

 Control group #1: Sample size: _____ Mean: _____ SD: _____

 Control group #2: Sample size: _____ Mean: _____ SD: _____

Outcome 3

Treatment group: Sample size: _____ Mean: _____ SD: _____

 Control group #1: Sample size: _____ Mean: _____ SD: _____

 Control group #2: Sample size: _____ Mean: _____ SD: _____

Correlation:

N (tx): _____ N(cont): _____ r: _____

T-test/ F-test:

N (tx): _____ N(cont): _____ t: _____ F: _____

ROC areas:

AUC: _____

Odds ratios:

Odds ratio: _____ Confidence interval: _____ p: _____

Odds ratios:

Odds ratio: _____ Confidence interval: _____ p: _____