

Do Risk Assessment Tools Help Manage and Reduce Risk of Violence and Reoffending?
A Systematic Review

Jodi L. Viljoen Dana M. Cochrane Melissa R. Jonnson
Simon Fraser University

Law and Human Behavior

DOI: <http://dx.doi.org/10.1037/lhb0000280>

This article may not exactly replicate the authoritative document published in the APA Journal.
It is not the copy of record.

Author Note:

Jodi L. Viljoen, Dana Cochrane, and Melissa Jonnson, Department of Psychology, Simon Fraser University.

This research was supported by a grant from the Simon Fraser University/Social Sciences and Humanities Research Council of Canada's Institutional Grants Program. The opinions expressed in this article reflect the views of the authors alone. The authors would like to thank Andrew Gray for his work as a research assistant.

Correspondence concerning this article should be address to Jodi Viljoen, Department of Psychology, Simon Fraser University, Burnaby, BC V5A 1S6. Contact: jviljoen@sfu.ca

Abstract

Although it is widely believed that risk assessment tools can help manage risk of violence and offending, it is unclear what evidence exists to support this view. As such, we conducted a systematic review and narrative synthesis. To identify studies, we searched 13 databases, reviewed reference lists, and contacted experts. Through this review, we identified 73 published and unpublished studies ($N = 31,551$ psychiatric patients and offenders, $N = 10,002$ professionals) that examined either professionals' risk management efforts following the use of a tool, or rates of violence or offending following the implementation of a tool. These studies included a variety of populations (e.g., adults, adolescents), tools, and study designs. The primary findings were as follows: (1) despite some promising findings, professionals do not consistently adhere to tools or apply them to guide their risk management efforts; (2) following the use of a tool, match to the risk principle is moderate and match to the needs principle is limited, as many needs remained unaddressed; (3) there is insufficient evidence to conclude that tools directly reduce violence or reoffending, as findings are mixed; and (4) tools appear to have a more beneficial impact on risk management when agencies use careful implementation procedures and provide staff with training and guidelines related to risk management. In sum, although risk assessment tools may be an important starting point, they do not guarantee effective treatment or risk management. However, certain strategies may bolster their utility.

Public Significance Statement

To help guide treatment and management decisions, professionals often assess offenders' and patients' risk of violence or offending using risk assessment tools. However, based on our review, these tools are often not used in an optimal manner. As such, they may not have as large an impact as intended.

Keywords: systematic review, risk assessment, violence, offending, risk management

Do Risk Assessment Tools Help to Manage and Reduce Risk of Violence and Reoffending? A Systematic Review

The field of violence risk assessment has grown rapidly. In the past several decades, researchers have developed over 400 different tools designed to assess risk of violence and offending (Singh, Desmarais, Hurducas, et al., 2014). Professionals, such as psychologists, probation officers, nurses, psychiatrists, and police, have widely adopted these tools in as many as 44 countries (Singh, Desmarais, Hurducas, et al., 2014). In addition, administrators and policymakers have created policies and, in some cases, laws mandating the use of tools (Monahan & Skeem, 2014). For instance, in the United States, at least 20 states use risk assessment tools in sentencing offenders (Starr, 2014), and at least 28 states use tools to determine if an offender should be released on parole (Harcourt, 2007). All 50 states have adopted the use of risk assessment tools in youth probation, with 34 states mandating the use of tools on a state-wide basis (Wachter, 2015). Thus, considering that over 4.5 million American adults and adolescents are placed on probation or parole each year (Kaeble & Bonczar, 2017; Livsey, 2012) and over 1.5 million Americans are incarcerated (Carson & Anderson, 2016), tools are used with many offenders.

Risk assessment tools are also commonly used in forensic psychiatric facilities (Singh, Desmarais, Hurducas, et al., 2014), general psychiatric hospitals (Higgins, Watts, Bindman, Slade, & Thornicroft, 2005), treatment programs (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010), correctional centers (Taxman, Cropsey, Young, & Wexler, 2007), and in a variety of court evaluations, such as those involving the transfer of adolescents to adult court and the civil commitment of individuals who have sexually offended (Viljoen, McLachlan, & Vincent, 2010). Extrapolating from these reports, it could be estimated that, globally, over a million risk assessments are conducted *each year*.

Driving the widespread adoption of risk assessment tools is a belief that such tools can help professionals to better manage and ultimately reduce risk. Broadly speaking, risk management refers to the process of planning and implementing strategies to help prevent violence and other forms of offending. It is carried out by a variety of professionals (e.g., psychologists, probation officers, nurses, police) and encompasses not only treatment (e.g., therapy), but also strategies such as supervision, case management, and placement decisions. For instance, Hart, Douglas, and colleagues (Douglas, Hart, Webster, & Belfrage, 2013; Hart et al., 2003) conceptualize risk management strategies as falling in four domains: (1) treatment such as one-on-one or group counselling, (2) supervision and control efforts such as placements in secure facilities or supervision by probation officers, (3) monitoring of risk factors, and (4) victim-safety planning strategies such as warning potential victims.

There is widespread agreement, among risk assessment researchers, that risk management is a key goal of risk assessment (Andrews & Bonta, 2010; Douglas & Kropp, 2002; Hart & Logan, 2011; Heilbrun, 1997). Some authors have pointed out that the goal of risk management may be more salient in some contexts than others (e.g., modifiable decisions in which courts have continued legal jurisdiction vs. decisions in which there is no continued jurisdiction; Heilbrun, 1997). However, other authors have asserted that risk management is the primary goal or sine qua non of risk assessment regardless of context (Hart & Logan, 2011). For instance,

Hart (1998) states, “To put it simply, the clinical task is violence prevention, not violence prediction” (p. 123). This is because predicting if someone will reoffend, in and of itself, has little value if nothing is done to manage risk. In fact, according to some authors, assessing risk without making efforts to manage it is “harmful and unethical” (Roychowdhury & Adshead, 2014, p. 81).

As a result, the field of risk assessment has evolved to focus increasingly on the management and reduction of risk, rather than solely its prediction (Bonta & Andrews, 2017). Indeed, many risk assessment tools are now “infused with the concept of risk reduction” (Skeem & Lowenkamp, 2016, p. 682). This focus on risk management is evident in tools’ manuals and descriptions. For instance, the authors of the Structured Assessment of Violence Risk in Youth (SAVRY) describe that “the objective of the SAVRY is ultimately to assist in prevention and risk reduction” (Borum, Bartel, & Forth, 2006, p. 5).

Similarly, many tool developers list specific types of risk management tasks for which their tools can be used. These include tasks such as to help judges “decide between alternative sentences” (Kropp & Gibas, 2010, p. 230), determine an “appropriate security classification” (Bonta & Andrews, 2017, p. 202), “inform future risk management and treatment plans” (Douglas, Hart, Webster, & Belfrage, 2013, p. 35), “assist in treatment and daily management” (Webster, Martin, Brink, Nicholls, & Desmarais, 2009, p. 4), and “help parole boards in making release decisions” (Bonta & Andrews, 2017, p. 202). The belief that risk assessment can help manage and reduce reoffending is also evident in some practice guidelines. For instance, the National Working Group of the National Center for State Courts encourages judges to use risk assessment tools in sentencing decisions because tools offer benefits such as “contributing to public safety” (Casey, Warne, & Elek, 2011, p. 7). However, even though many risk assessment tools aim to help manage and reduce risk, it is unclear whether tools do, in fact, achieve this goal. As we describe below, opinions about this differ.

Tools Might Help Reduce Risk by Linking Individuals to Appropriate Interventions

Many tool developers, researchers, and practitioners consider risk assessment to be an important—or even essential—starting point to risk management (e.g., Douglas & Kropp, 2002; Vincent, Paiva-Salisbury, Cook, Guy, & Perrault, 2012). However, in more exact terms, how might the use of risk assessment tools improve risk management practices, or reduce reoffending and violence? Although the mechanism between risk assessment and risk management is not always articulated, the primary hypothesized mechanism appears to be twofold (see Bonta & Andrews, 2017). First, it is thought that tools will increase professionals’ level of adherence to the risk-need-responsivity (RNR) model of offender treatment. Second, it is believed that this will, in turn, decrease offending.

The RNR model (Andrews, Bonta, & Hoge, 1990; Bonta & Andrews, 2017) is the dominant model of offender treatment, and it has considerable research support. A number of meta-analyses have found that programs that adhere to the RNR principles are more effective in reducing reoffending than those that do not adhere to these principles (e.g., Hanson, Bourgon, Helmus, & Hodgson, 2009; Koehler, Lösel, Akoensi, & Humphreys, 2013; Lipsey, 2009; Lowenkamp, Latessa, & Holsinger, 2006; Prendergast, Pearson, Podus, Hamilton, & Greenwell,

2013). As such, increasing adherence to this model is a reasonable goal.

The RNR model has three primary principles, namely the risk, need, and responsivity principles (Andrews & Bonta, 2010; Bonta & Andrews, 2017). According to the risk principle, high risk individuals should receive more intensive interventions than lower risk individuals. According to the need principle, individuals should receive individualized interventions that target modifiable risk factors that contribute to offending (i.e., criminogenic needs), such as substance use treatment for an offender with substance use problems. Finally, according to the responsivity principle, individuals should receive cognitive-behavioral treatment delivered in the context of a strong therapeutic relationship, with treatment being tailored to individual characteristics such strengths and culture.

Risk assessment tools are designed to help collect information that is relevant to the RNR principles. For instance, risk assessment tools, by definition, are designed to help generate estimates of individuals' risk levels (Skeem & Monahan, 2011). In addition, some tools are designed to identify criminogenic needs, such as antisocial attitudes, and responsivity factors, such as strengths (Bonta & Andrews, 2017). By structuring the routine collection of this information, tools might conceivably enable better match to the RNR principles (e.g., Douglas & Kropp, 2002). For instance, by systematically assessing risk level, professionals may be able to make more appropriate decisions about whether individuals can be safely managed in the community rather than in a secure facility. Also, by identifying criminogenic needs, professionals might be more likely to target those needs in treatment.

In sum, tools are viewed as a means by which to deploy interventions that are evidence-based and individually-tailored, thereby avoiding a "one size fits all" approach. In this sense, many modern risk assessment tools arose from a "rehabilitative ideal" that rejected the claim that "nothing works" and instead emphasized the importance of effective, individualized treatments rather than harsh punishments (Maurutto & Hannah-Moffat, 2007, p. 470).

The Impact of Tools May Depend on the Nature of the Tool and Subsequent Interventions

Although many researchers believe that risk assessment tools can help manage and reduce risk, some researchers have noted that the impact of tools may be contingent upon other factors, such as the type of tool (e.g., Douglas & Skeem, 2005), and whether the tool is followed through on to match individuals to appropriate interventions (e.g., Bonta, Rugge, Scott, Bourgon, & Yessine, 2008). Indeed, tools vary in terms of the types of items they include, and how items are combined to form an overall conclusion (Skeem & Monahan, 2011). Whereas some tools focus primarily on historical or static factors, such as history of offending, other tools include dynamic or modifiable factors, such as anger management difficulties (i.e., criminogenic needs; Douglas & Skeem, 2005). In addition, on some tools (i.e., actuarial tools), evaluators add up risk factors to generate a total score (Skeem & Monahan, 2011). On other tools (i.e., structured professional judgement tools), evaluators use their discretion to make a separate summary risk rating after considering the items and additional case-specific considerations. Thus, tools may differ in their utility for various risk management decisions.

For instance, many authors have asserted that tools that include dynamic risk factors will

have greater utility for treatment-planning than tools that include primarily historical factors, as dynamic factors can serve as treatment targets (e.g., Douglas & Kropp, 2002; Douglas & Skeem, 2005; Hanson & Harris, 2000; Hart & Logan, 2011; Roychowdhury & Adhhead, 2014). However, historical factors may be relevant to other aspects of risk management, such as decisions about the intensity of supervision and the need for secure placements (Hart, Webster, & Douglas, 2001; Douglas & Kropp, 2002). Some authors have also argued that structured professional judgment tools may have greater utility for treatment-planning than actuarial tools because existing structured professional judgment tools tend to place a greater focus on dynamic factors than do many actuarial tools (e.g., Hart & Logan, 2011; Roychowdhury & Adhhead, 2014). That said, there nothing inherent in actuarial tools that necessarily precludes attention to dynamic factors (Douglas, Hart, Groscup, & Litwack, 2013; Skeem & Monahan, 2011). Some actuarial tools can and do include dynamic factors.

Besides variation in item content, risk assessment tools also vary in terms of their validity, or their ability to predict reoffending. Some tools have been found to significantly predict reoffending in multiple studies, with effect sizes falling in the moderate range (Hanson & Morton-Bourgon, 2009; Olver, Stockdale, & Wormith, 2009, 2014; Singh, Grann, & Fazel, 2011; Yang, Wong, & Coid, 2010). However, other risk assessment tools have poor predictive validity, or have not yet been tested at all. Presumably, unless a tool has adequate validity in predicting reoffending, it will have limited value for risk management.

Finally, tools vary in the extent to which they include a focus on risk management. Whereas most risk assessment tools do not include explicit instructions or guidance regarding how to manage risk, some tools aim to bridge risk assessment to risk management by providing greater structure and support for risk management, such as by including case management planning forms. Such tools are sometimes referred to as “fourth generation” tools to reflect the notion that that they represent an advancement over earlier tools (Bonta & Andrews, 2017).

Not only does the utility of risk assessment depend on the nature of tool, it could also depend on how professionals *use* and apply tools. In particular, as is true of any type of assessment, the value of risk assessment likely lies primarily in what happens *after* the assessment. Although risk assessment tools may be a *starting point* for treatment and risk management, they are not a treatment in and of themselves. As such, it may be unrealistic to expect that using a tool will help manage risk or reduce reoffending unless: (a) appropriate treatments are, in fact, available, and (b) professionals meaningfully apply tools to match individuals to these treatments.

For this reason, Vincent, Paiva-Salisbury, et al. (2012) note that “simply adopting a risk assessment tool is a necessary but insufficient method” for improving risk management (p. 553). Similarly, as Wong and Olver (2010) describe, violence risk assessment might help reduce the likelihood of violence *if* it occurs “in combination with interventions and treatment” (p. 121). Herein lies a potential barrier in the pathway between risk assessment and risk management. It is unclear to what extent tools are, in fact, optimally used to guide interventions. Some researchers have observed significant gaps between risk assessments and risk management efforts (Bonta et al., 2008; Peterson-Badali, Skilling, & Haqanee, 2015).

Some Have Questioned the Value of Tools for Risk Management

Despite the debate about whether some tools may be better-suited for risk management than others, most risk assessment researchers appear to believe that tools can, in principle, aid in risk management. However, this viewpoint is not embraced by everyone (e.g., Almond, 2012; Hannah-Moffat, Maurutto, & Turnbull, 2009; Harcourt, 2007; Large, Ryan, Callaghan, Paton, & Singh, 2014; Ryan, Nielssen, Paton, & Large, 2010; Rose, 1998; Szmukler & Rose, 2012; Starr, 2014). Some critics have expressed concern that risk assessment tools, in general, are not only ineffective in managing risk, they might even cause harm to patients and offenders (e.g., Large et al., 2014).

In particular, critics have pointed out that tools tend to have high rates of false positives, meaning that many offenders who are rated as high risk do not reoffend (e.g., Ryan et al., 2010). It is thought that this ‘risk inflation’ might lead to stigma, over-incarceration and warehousing of individuals who are *high* risk, and a decreased focus on rehabilitation and therapeutic efforts (Almond, 2012; Fitzgibbon, 2007; Rose, 1998; Ryan et al., 2010). At the same time, there is concern that the use of risk assessment tools might lead to an inappropriate withdrawal of resources from clients who are *low* risk, because these individuals may be falsely viewed as not needing treatment (Large et al., 2014; Rose, 1998; Ryan et al., 2010). Thus, in these critics’ views, risk assessment tools provide a “beguiling, but flawed, rationale for the distribution of resources” (Nielssen, Ryan, & Large, 2011, p. 270). This concern is compounded by the fact that assessments can be costly and could erode resources that might otherwise be available for treatment (Nielssen et al., 2011).

Furthermore, some critics have questioned the motives that underlie the adoption of risk assessment tools. Hannah-Moffat and colleagues (2009) argues that a primary reason that agencies use tools is not to reduce risk or protect others, but rather to protect themselves from lawsuits by creating a “paper trail” (p. 396). Similarly, Undrill (2007) describes risk assessment as a “neurotic organisational attempt to tame anxiety” (p. 294). In other words, there is concern that the use of tools puts the needs of the agency before those of the clients themselves.

Finally, critics have pointed out that despite claims that risk assessment tools help to manage risk, there is little evidence to support such assertions (e.g., Wand, 2012). For instance, in a survey of 1,937 psychiatrists, 87% of respondents endorsed the view that tools provide a false sense of security, as there is little direct evidence that tools help to reduce adverse events (Royal College of Psychiatrists, 2008). Some risk assessment researchers and tool developers have offered similar cautions. For instance, Rice, Harris, and Hilton (2010) noted that, “At this point, no one can promise that the introduction of any formal violence risk assessment (actuarial or nonactuarial) will improve real risk-related decisions. No one really knows what strategies in the application of available empirical evidence will ensure that avoidable violent recidivism and unnecessary restriction of offenders’ freedoms are simultaneously minimized” (p. 113).

The Present Study

In sum, opinions about the value of risk assessment tools for risk management efforts range considerably. Although some researchers view risk assessment tools as invaluable to risk

management, other scholars believe that such tools are ineffective at best and harmful at worst. Still others acknowledge both the promise *and* pitfalls of tools, noting that the value of tools for risk assessment might depend on factors such as the nature of the tool. As such, at the present time, it is unclear what conclusions are justified. Concerningly, many assertions about the utility of tools (or lack of utility) have been offered without reference to research findings.

One possible explanation for the lack of empirical grounding for these assertions is that relevant research simply does not exist, as of yet. However, another possibility is that existing research has not yet been adequately integrated into the literature due to a lack of comprehensive reviews. Indeed, although there have been numerous efforts to synthesize research on the predictive validity of risk assessment tools (e.g., Hanson & Morton-Bourgon, 2009; Olver et al., 2009, 2014; Singh et al., 2011; Yang et al., 2012), systematic reviews on the utility of tools for risk management are lacking.

To our knowledge, only one systematic review has examined the utility of risk assessment tools for risk management, and this review was not designed to focus on risk assessment per se. In particular, Gaynes et al. (2017) conducted a systematic review of strategies to reduce aggression among psychiatric patients in acute settings (e.g., staff training, medication protocols). As part of this review, they identified two randomized control trials (RCTs) which examined whether the implementation of risk assessment tools reduces aggression (i.e., Abderhalden et al., 2008; van de Sande et al., 2011). Gaynes et al. (2017) concluded that although these two studies provide some evidence that risk assessments may reduce aggression in acute psychiatric patients, the strength of this evidence is low due to limitations in research.

The goal of the present systematic review was to expand on Gaynes et al.'s (2017) review. First, rather than focusing exclusively on psychiatric patients in acute care settings, we included a range of populations (e.g., patients, offenders) and settings (e.g., jails, forensic hospitals). Second, rather than restricting our review to RCTS, our review encompassed a variety of designs (e.g., RCTs, pre-post studies). We adopted this broad lens so that we could determine the full scope of extant research. Third, instead of solely examining whether the use of risk assessment tools reduce violence, we also investigated their impact on professional practices. Certainly, violence reduction is one indicator of the success of a tool. For instance, Douglas and Kropp (2002) note that "risk assessment should be considered successful when we can demonstrate reduced rates of violence in connection with risk assessment procedures" (p. 623). However, violence reduction is not the only possible indicator of tools' impact. If risk assessment tools do, indeed, reduce violence and offending, this relationship may be indirect, and contingent upon professionals' risk management practices.

As such, we reviewed research on whether tools facilitate professionals' adherence to the risk and need principles, as these are the hypothesized mechanism by which tools might help manage risk. We also examined whether professionals perceive tools as useful for risk management and whether they use tools to guide their risk management effort; tools are unlikely to be effective if professionals do not apply them or view them as useful. In other words, our goal was to understand, more thoroughly, the pathway between risk assessment and risk management. Furthermore, to develop an agenda for future research, we reviewed studies on strategies to enhance the utility of risk assessment tools for risk management, such as staff

training. Even if risk assessment tools do not have as positive an impact as desired, there may be means to improve their utility. As such, our research questions were as follows:

1. Do professionals perceive risk assessment tools to be useful for risk management?
2. When professionals conduct risk assessments with tools, do these assessments guide risk management efforts?
3. Does the use of risk assessment tools facilitate match to the risk principle?
4. Does the use of risk assessment tools facilitate match to the need principle?
5. Does the implementation of risk assessment tools reduce violence or offending?
6. What strategies help to improve the utility of risk assessment tools for risk management?

Method

Design

To examine our research questions, we chose to conduct a systematic review rather than a traditional literature review because systematic reviews are more transparent, comprehensive, and objective (Green, Johnson, & Adams, 2006). For instance, in a systematic review, authors provide a detailed account of their methodology (e.g., search terms, databases).

To synthesize findings, we used a narrative approach. An empirical synthesis (i.e., meta-analysis) was neither feasible nor appropriate because our review included a wide range of designs (e.g., RCTs, surveys), populations (e.g., offenders, patients), tools (i.e., 34 different risk assessment tools), and outcomes (e.g., supervision, dispositions; see Borenstein, Hedges, Higgins, & Rothstein, 2009; Popay et al., 2006; Ryan & Cochrane Consumers and Communication Review Group, 2013). In particular, to help determine if a meta-analysis was viable, we created a table for each research question that listed the tools in the row headings and the specific outcomes in the column headings. Most cells had only one or two studies listed, indicating that it would be difficult to meaningfully combine studies in an empirical manner. In addition to being more viable, a narrative approach enabled us to provide more detailed examples and descriptions of studies than is typical in an empirical synthesis. This was important because research in this area is at an early stage, and thus relatively little is known about these studies.

To ensure that our systematic review met best practice standards and followed relevant reporting guidelines, we followed the criteria set forth in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Liberati et al., 2009) and the Assessment of Multiple Systematic Reviews (AMSTAR) tool (Shea et al., 2007).

Inclusion and Exclusion Criteria

Studies were required to meet the following inclusion criteria: (1) empirical study that was published or disseminated in English; (2) included a sample of individuals who were assessed with a structured risk assessment tool in real-world practice, and/or a sample of professionals who used risk assessment tools in practice; and (3) included an outcome relevant to at least one of our research questions (e.g., perceived utility, adherence to the risk principle). We defined structured risk assessment tools as tools that included a designated list of risk factors and

an overall rating of risk level for violence or offending (Skeem & Monahan, 2011). Thus, we did not include measures of psychopathy. Also, we did not include qualitative research, as systematic reviews of qualitative studies use different methodologies, nor did we include research on the responsivity principle, as this principle encompasses a wide range of constructs (i.e., culture, trauma, mental health) and is not as well-researched as the risk and need principles. When disseminations were based on the same sample, we selected the study that was the most comprehensive and rigorous (e.g., largest sample).

Search Methods

Our search procedure is outlined in Figure 1. To select our search terms, we consulted with a research librarian and pilot tested 15 different combinations of terms. Our final set of search terms were as follows: "risk assessment" AND (violen* OR reoffen* OR recidivism OR offen*) AND ("risk management" OR utility OR useful OR "risk-need-responsivity" OR RNR OR "decision-making" OR "case management"). Using these terms, we searched 13 databases, including the following academic databases: PsycINFO, PsycARTICLES, PsycBOOKS, MEDLINE, National Criminal Justice Reference Service, Criminal Justice Abstracts, Sociological Abstracts, Social Sciences Abstracts, Social Sciences Full Text, Social Services Abstracts, and Web of Science.

To reduce the likelihood of publication bias (Ioannidis, 2005; Kicinski, Springate, & Kontopantelis, 2015), we also searched databases that capture more of the gray literature (i.e., ProQuest Dissertations and Theses Global Database, Google Scholar). These searches encompassed January 1990 (i.e., an approximate date of when risk assessment tools were first developed) to the date of the search (September 30, 2016). Given that Google Scholar searches generate an unwieldy number of hits, researchers typically restrict their screening to the first 50 to 100 search records when conducting systematic reviews (Haddaway, Woodcock, Macura, & Collins, 2015). However, we adopted a more comprehensive approach and examined the first 300 Google Scholar search records. Only 15 new studies were screened in through our Google Scholar search and only one of these met inclusion criteria, which suggests that we successfully identified most of the relevant studies through the other databases. Besides searching databases, we reviewed reference lists of relevant studies and emailed 15 authors with multiple studies in this area (i.e., two or more studies included in this review); these searches were completed August 31, 2017. Included studies are listed in Appendix A.

Data Collection

To collect data, we used a six-stage process, and followed the guidelines and examples provided by the Cochrane Collaboration (e.g., Johnson, Sandford, & Tyndall, 2003; Ryan & Cochrane Consumers and Communication Review Group, 2013) and the PRISMA statement (Liberati et al., 2009). In Stage 1, the authors screened the abstracts of the 1,588 articles identified from our searches to determine if they were relevant (see Figure 1). To check the accuracy of our screening, each author screened 75 abstracts for practice; they correctly screened in each of the eligible studies. In Stage 2, the study authors examined full copies of each of the 236 screened-in abstracts to determine if the study met inclusion criteria.

In Stage 3, data from the 73 studies that met inclusion criteria was extracted using a data extraction form (described below). We also appraised each study's limitations. To increase objectivity and replicability, each study was coded by two raters (i.e., study authors [DC, MJ] and one trained research assistant). If the dissemination did not include adequate information (e.g., the type of tool was not adequately specified), we contacted the authors for clarification.

In Stage 4, the raters discussed each study's overall findings and study appraisal to reach consensus; the first author, who reviewed all studies, provided an opinion when consensus could not be reached. In Stage 5, the first author compiled ratings in evidence tables (i.e., tables that summarized the methodology and findings of each study; Green et al., 2006; Popay et al., 2006) and prepared narrative descriptions of the results. These tables and results were reviewed and checked by the other authors, and a second set of consensus meetings were held to finalize ratings. Finally, in Stage 6, each of the authors graded the overall strength of evidence for each research question and arrived at a consensus.

Materials

Data Extraction Form. Our data extraction form included approximately 150 questions, divided into sections that pertained to each of our research questions. Extracted data included: study information (e.g., publication type), sample (e.g., gender), method (e.g., design), risk assessment tool, outcome (e.g., violence), results, and follow-up information (e.g., new studies to screen). We developed this form by reviewing other data extraction forms (e.g., Guy, 2008). We then pilot tested the form with six cases, and adjusted the form accordingly. Cohen's (1960) kappa coefficients for setting (i.e., justice setting, mental health or forensic setting), gender of sample (i.e., male only, female only, mixed), age of sample (i.e., adult, adolescent, mixed), type of professionals (i.e., clinicians, probation officers, police, other), and study design (i.e., RCT, pre-post, post-test only, survey, other) were 1.00, 0.90, 0.90, 0.89, and 0.82, respectively ($n = 73$ cases). Thus, agreement fell in the "almost perfect" range (i.e., $> .80$; Landis & Koch, 1977).

In addition to rating study characteristics, we made summary ratings for each study to help interpret findings. Interrater reliability of summary ratings were assessed using intraclass correlation coefficients ([ICCs], two-way mixed, absolute agreement, average measures; Hallgren, 2012; McGraw & Wong, 1996). For research questions 1 – 4 (i.e., perceived utility, use of tools for risk management, match to the risk and need principles), we made summary ratings of low, mixed/moderate, or high (operational definitions are provided in the Results section and in the tables; ICC = 0.92, $n = 82$). For research question 5 (i.e., violence/reoffending), we made summary ratings of decrease, mixed, or no change (ICC = 1.00, $n = 12$). For research question 6 (i.e., strategies to improve utility), we made summary ratings of improvement, some improvement, or no change (ICC = 0.89, $n = 8$). These ICC values were all in the "excellent" range (i.e., $> .75$; Cicchetti, 1994).

Study Appraisal. We appraised the limitations of RCTs using the Cochrane Risk of Bias tool for RCTs (Higgins et al., 2011). This tool examines various forms of bias (e.g., bias arising from the randomization process, measurement of the outcome). Raters then assessed overall risk of bias as low, medium, or high. ICCs were not calculated because only four studies were RCTs and there was limited variability in ratings. However, to check our ratings, we

independently rated the two RCTs that were included in Gaynes et al. (2017) and then compared our ratings to Gaynes et al.'s ratings to ensure that they were consistent.

We appraised the limitations of observational studies and non-controlled intervention studies (e.g., pre-post studies) using the following items: sample, adequacy of comparison group, adherence to the tool, measurement approach ([a] description, [b] support for measurement approach). These items are based on the National Institute for Health and Care Excellence [NICE] tool (2012), but the wording was adjusted for this context (e.g., replacing the term “intervention” with the term “risk assessment tool”). To help structure ratings, we developed anchors. Each item was rated yes, possibly/partially, no or not reported, or not applicable. Then, an overall judgement was made (i.e., low, medium, or high study limitations). The ICC (mixed, absolute agreement, average measures) for overall judgment was .73 ($n = 68$).

Finally, we appraised survey studies using the following items from the Center for Evidence-Based Medicine's Critical Appraisal of Surveys (N.D.): sample, response rate, and measurement approach ([a] description, [b] support for measurement approach). Once again, each item was rated yes, possibly/partially, no or not reported, or not applicable, and an overall judgement was made (i.e., low, medium, or high study limitations). The ICC for overall judgment was .92 ($n = 20$).

Overall Strength of Evidence for Outcomes. In addition to evaluating *individual* studies, we rated the *overall* strength of evidence for our research questions (e.g., whether tools facilitate match to the risk principle). To do so, we used the system developed by the Evidence-Based Practice Center of the Agency for Healthcare Research and Quality (AHRQ; Berkman et al., 2015). The AHRQ system is one of the most common methods for rating quality of evidence (Kane, Butler, & Ng, 2016), and it draws heavily from the GRADE framework (Grading of Recommendations Assessment, Development, and Evaluating Work Group; Guyatt et al., 2011). Within the AHRQ system, raters evaluate a body of research in five domains, including study limitations (the overall quality of study designs and risk of bias), consistency (whether different studies find the same pattern of results), directness (whether the evidence directly links the intervention to the outcome), precision (the degree of certainty of the estimated effect), and reporting bias (publishing or reporting findings based on their favorability). Then, they grade the overall quality of evidence as high, moderate, low, or insufficient. We did not complete AHRQ ratings for studies on perceived utility or use of tools for risk management, as the AHRQ system was not designed for survey research.

Results

In total, 73 studies met inclusion criteria (see Figure 1 and Appendix A). Sixteen of these studies were unpublished; nine of the unpublished studies were dissertations or theses and the remaining seven were reports by researchers, government, or other organizations. These studies included 31,551 offenders or patients, 10,002 professionals, and 34 risk assessment tools. Most tools included dynamic or modifiable factors (i.e., criminogenic needs; 76.5%, $k = 26$), and were validated (i.e., have been found to significantly predict violence or reoffending; 79.4%, $k = 27$). However, the degree of research support for tools ranged from single studies to large meta-analyses. Appendix B provides a brief description of the tools.

1. Do Professionals Perceive Risk Assessment Tools to be Useful for Risk Management?

Study Characteristics. In 12 studies (six of which were unpublished), researchers surveyed or interviewed professionals about the utility of risk assessment tools for risk management ($N = 6,664$ professionals; see Table 1). These studies examined the extent to which professionals considered tools to be useful for management-related decisions, such as placement and supervision decisions (66.7%, $k = 8$), and/or treatment-related decisions, such as service referrals (33.3%, $k = 4$). Some studies involved probation or parole officers (POs) or other correctional staff (50.0%, $k = 6$), whereas other studies involved mental health clinicians (e.g., psychologists; 41.7%, $k = 5$). Also, a couple of studies involved judges (16.7%, $k = 2$). In most studies, researchers examined tools that have been found to significantly predict offending (83.3%, $k = 10$; see Appendix B). However, in three studies, the authors surveyed professionals about tools in general rather than asking about a specific tool per se. Most of the examined tools included dynamic or modifiable factors (66.7%, $k = 7$). However, in two studies, researchers examined the VRAG, a tool which focuses on historical or static factors.

Study Appraisal. Five studies (41.7%) were rated as having low limitations overall, whereas the remaining studies had medium (41.7%, $k = 5$) or high limitations (16.7%, $k = 2$; see Table 7). Of the specific domains that were appraised (sample representativeness, response rate, description of measure, and support for measure), the most common limitation was sample representativeness. Only four of the 12 studies (33.3%) included clear evidence to support the representativeness of their sample. This means that the results may not generalize to the population of tool users; for instance, survey respondents may have had more positive or negative views about the perceived utility of tools than typical users. In addition, few studies (33.3%, $k = 4$) included clear support for the approach that was used to measure perceived utility, such as evidence of pilot testing or expert feedback.

Results. In most studies (58.3%, $k = 7$), perceived utility was mixed, mixed-low, or mixed-high (see Tables 1 and 7). A rating of mixed was given when professionals rated usefulness as being just above the midpoint, such as a mean of 5 to 7 out of 10 (see Table 1 for rating criteria). For instance, in one survey, corrections staff rated the usefulness of the Y-LSI for identifying treatment needs as 5.5 on a 10-point scale, with 10 being the most positive rating (Flores, Travis, & Latessa, 2004); they rated its usefulness for making placement decisions as 5.2 out of 10. In another study, forensic staff rated the perceived utility of the VRAG for determining transfers to be 7.8 out of 10 (with 10 being very important; McKee, Harris, & Rice, 2007). However, they rated other factors, such as severity of the index offense, to be more important than the VRAG. In a couple of studies (16.7%, $k = 2$), perceived utility was low (i.e., helpfulness was rated below the midpoint). For instance, in a survey of psychiatrists, only 48% of respondents reported that tools lead to better decisions (Royal Colleges of Psychiatrists, 2008). Conversely, in a couple of studies (25.0%, $k = 3$), perceived utility was high (i.e., most participants viewed the tool as helpful). For example, in one study, 78% of clinicians reported that the VRAG was very helpful for making discharge decisions (Muheizen, 2014).

Potential Moderators. We did not find evidence of publication bias (i.e., wherein unpublished studies are less positive). However, there was some evidence that the perceived utility of tools varied somewhat by tool. For instance, in a large international survey, Singh,

Desmarais, Hurducas, and colleagues (2014) found that tools with the highest perceived utility for risk management were the FORTES, HKT-30, HCR-20, SAPROF, START, and VRS; these tools all include dynamic risk factors. In contrast, the tool with the lowest perceived utility was the VRAG; this tool focuses on historical factors and/or factors which are assumed to be stable. However, in other studies, tools like the VRAG were considered helpful (McKee et al., 2007; Muheizen, 2014). There were also some differences across professionals. Specifically, judges rated structured decision-making tools as being more useful for post-adjudication placement than did POs (Shook & Saari, 2007). Finally, although this was not directly tested, there was some indirect evidence that perceived utility may be higher when tools are selected or designed with staff input, and fit the specific needs of sites (e.g., are relevant to the population served at the site). For instance, in one study, an agency developed a site-specific risk assessment tool through staff consultation (Richardson, 2009); clinicians viewed the resulting tool as very useful.

Summary. In sum, studies revealed a range in views about the utility of tools for risk management; some professionals view these tools as useful whereas others do not (see Table 7). However, even if professionals view tools as useful, they may not adequately apply tools to their decision-making. Therefore, as a next step, we reviewed studies that examined the *use* of risk assessments for risk management.

2. When Professionals Conduct Risk Assessments with Tools, Do These Assessments Guide Risk Management?

Study Characteristics. We identified 14 studies (including three unpublished studies) in which researchers examined the extent to which findings from risk assessments tools guide or influence risk management efforts (see Table 2; $N = 5,628$ professionals). These studies examined use for management-related decisions (e.g., determining supervision level; 42.9%, $k = 6$), treatment-related decisions (e.g., measuring treatment progress; 35.7%, $k = 5$), and/or general use (e.g., developing “risk management plans,” 28.6%, $k = 4$). Most of the studies were surveys of professionals at agencies in which tools had been implemented (64.3%, $k = 9$; i.e., professionals’ self-reported use). However, in a couple of studies (35.7%, $k = 5$), researchers investigated the use of risk assessment tools through observational methods, such as by examining whether judges referred to tools in their written legal decisions. Many studies focused on POs or other correctional staff (64.3%, $k = 9$), but some studies included clinicians (21.4%, $k = 3$), judges (21.4%, $k = 3$), and/or jurors (14.3%, $k = 2$). In most studies, researchers examined tools which have been found to significantly predict offending and/or violence (64.3%, $k = 9$; see Appendix B), but in three studies, researchers examined tools in general rather than a specific tool, or did not specify the name of the tool, and thus validity could not be determined. Many of the examined tools include a focus on dynamic factors (64.3%, $k = 9$). However, two studies examined tools that focus on historical or static factors (i.e., Static 99, MnSOST-R).

Study Appraisal. Although six studies (42.9%) were rated as having low study limitations, the remaining studies had medium (35.7%, $k = 5$) or high limitations (21.4%, $k = 3$; see Table 7). Of the specific domains that were appraised, the most common limitation was with respect to research support for measurement approaches; only two of the 14 studies (14.3%) included clear evidence that the approach that they used to measure professionals’ use of tools was reliable or valid. As such, these studies may under- or over-estimate the use of tools.

Results. In most studies (57.1%, $k = 8$), the application of risk assessment tools to risk management efforts was mixed, mixed-low, or mixed-high (see Tables 2 and 7). A rating of mixed was given when helpfulness was rated as just above the midpoint (see Table 2 for rating criteria). As an example, Miller and Maloney (2013) classified 40% of tool users as bureaucratic compliers, meaning that although they completed risk/needs assessment tools for assessment purposes, they were less likely to apply these tools to guide their decision-making. Also, in a case law review of adolescent offender cases, judges placed great weight on the risk assessment in 4% of cases, and some weight on it in 58% of cases (Urquhart & Viljoen, 2014). Two studies (14.3%) reported high use of tools for risk management wherein most professionals (e.g., over 70%) used the tool for risk management. For instance, Vincent, Paiva-Salisbury, et al. (2012) found that, following the implementation of the SAVRY and YLS/CMI, 80% of youth probation officers (YPOs) reported that they used the tool in service decisions and 90% used it in supervision decisions.

In contrast, in four studies (28.6%), the use of tools for risk management was low (i.e., less than half of professionals used the tool for risk management). In all but one of these studies, the use of tools was measured via observation or written records, suggesting that observation yields more modest estimates of use than do professionals' self-reports about their use of tools. For instance, in an observational study of 69 PO–probationer interactions, POs only used the tool for case management purposes on one occasion (Viglione, Rudes, & Taxman, 2015). In another study, professionals directly referred to clients' risk assessments in only 8% of plans (Kewley, Beech, Harkins, & Bonsall, 2015). Also, in a study of sexually violent predator trials, jurors' perceptions of risk were not significantly associated with respondents' Static-99 or MnSOST-R risk scores (Boccaccini, Turner, Murrie, Henderson, & Chevalier, 2013).

Potential Moderators. Results of published and unpublished studies were quite similar. However, a couple of studies found that the use of risk assessment tools for risk management may vary by profession (e.g., Haas & DeTardo-Bora, 2009; Shook & Saari, 2007). For instance, in one study, nurses were significantly more likely than doctors to report that risk assessments influenced their judgment or therapeutic actions (Hawley, Gale, Sivakumaran, & Littlechild, 2010). The use of tools for risk management may also vary depending on the particular task. For example, Haas and DeTardo (2009) found that 80% of professionals reported that they used the LSI-R to guide service referrals, but only 42% used it to develop reentry plans. In addition, some tools, such as the HCR-20, appear to be used for risk management more frequently than other tools (Singh, Desmarais, Hurducas, et al., 2014), possibly because professionals view them as more relevant to risk management.

Summary. Overall, this research indicates that although tools guide decisions in some contexts, “slippage” often occurs between assessments and risk management (Peterson-Badali, et al., 2015; see Table 7). Thus, we subsequently investigated, more specifically, how the use of tools impacts match to the risk and need principles, as this is the hypothesized mechanism by which tools are thought to improve risk management.

3. Does the Use of Risk Assessment Tools Facilitate Match to the Risk Principle?

Study Characteristics. According to the risk principle, higher risk individuals should

receive more intensive risk management strategies than lower risk individuals. We identified 36 studies (including nine unpublished studies) in which researchers examined match to the risk principle following the use of risk assessment tools; these studies included a total of 17,597 assessees (e.g., offenders, patients) and 2,507 professionals (see Table 3). Studies examined a variety of risk management decisions, including decisions about placements (e.g., incarceration, 25.0%, $k = 9$), other sentencing outcomes (e.g., dangerous offender designation, 5.6%, $k = 2$), supervision levels (22.2%, $k = 8$), and release or discharge (16.7%, $k = 6$). In addition, several studies examined decisions about the overall *level* or *intensity* of services (e.g., number of referrals, 16.7%, $k = 6$) or the number of strategies included in case management or risk management plans (25.0%, $k = 9$); overall intensity of services is an aspect of match to the risk principle (Bonta & Andrews, 2017).

Many studies were with offenders (80.6%, $k = 29$), and approximately half were with adult samples (47.2%, $k = 17$). In most studies (80.6%, $k = 29$), researchers measured match by examining associations between risk scores and the intensity of risk management strategies (e.g., correlations between risk total scores and decisions to incarcerate). However, in a couple of studies (8.3%, $k = 3$), researchers rated the proportion of individuals who received an acceptable intensity of risk management strategies, as defined by the researchers or policy guidelines.

Many studies examined tools which have demonstrated predictive validity (83.3%, $k = 30$; see Appendix B), but we were unable to identify evidence of empirical support for the predictive validity of a couple of tools (i.e., AJADMS, DSI, L17A, RAI). Most tools included dynamic or modifiable factors (80.6%, $k = 29$). However, six studies examined tools that focus on static or historical factors (i.e., AJADMS, DSI, RAI, Static-99, VRAG).

Study Appraisal. Most studies had medium limitations (66.7%, $k = 24$; see Table 7). Of the specific domains that were appraised (sample representativeness, comparison group, adherence to tool, measurement of need principle), the most common limitation was a lack of comparison groups. Only two of the 36 studies (5.6%) included a comparison group of individuals who were not assessed with a tool. As such, most studies were not designed to directly test whether the use of tools increases match to the risk principle per se. Also, many studies did not include evidence of adherence to the tools (e.g., that professionals completed tools as mandated); only 14 of the studies (38.9%) fully met this criterion.

Results. Based on the two studies that included comparison groups, there is some evidence that match to the risk principle may be better when tools are used than when they are not (see Tables 3 and 7). Bonta and Motiuk (1990) found that low risk inmates at jails that used the LSI were more likely to be transferred to a halfway house than were low risk inmates at a jail that did not use the LSI. In addition, Vincent, Paiva-Salisbury, et al. (2012) found that YPOs were more likely to identify risk level as an important consideration in their supervision decisions following the implementation of the SAVRY and YLS/CMI rather than prior to the implementation of tools. However, their reported use of risk level in service and disposition decisions did not significantly increase following implementation.

Although the remaining 34 studies did not have a comparison group of individuals who were not assessed with a tool, they nevertheless provide descriptive information regarding match

to the risk principle following the use of a tool. In most of these studies (64.7%, $k = 22$), match to the risk principle was moderate following the use of risk assessment tools (i.e., significant associations between risk total and intervention intensity which fell in the small or moderate range; see Table 4 for rating criteria). Furthermore, in five studies (14.7%), match to the risk principle following the use of a tool was high. For instance, in one study, offenders with a life sentence were significantly less likely to be granted parole if they were rated as high risk than if they were rated as low risk with the HCR-20 and LS/CMI (Guy, Kusaj, Packer, & Douglas, 2015). Not only were high risk individuals given more restrictive penalties, on average, they were also referred to more services, and professionals recommended that they receive more risk management strategies than low risk individuals. For example, in one study, police officers recommended a median of four risk management strategies for intimate partner violence offenders who were rated as high risk on the SARA, compared to a median of two strategies for offenders rated as low risk (Belfrage et al., 2012). Whereas match to the risk principle was generally moderate or high following the use of tools, in seven studies (20.6%), it was low. For instance, in a couple of studies, VRAG scores failed to predict tribunal decisions (Hilton & Simmons, 2001; McKee et al., 2007).

Potential Moderators. Overall, the findings in the published and unpublished studies were similar. There was some evidence that risk assessments might have a greater influence on the behavior of individuals who conducted the risk assessments themselves (e.g., clinicians) than on the behavior of external consumers of risk assessments (e.g., review boards who were provided with clinicians' risk assessments). For instance, in McKee et al. (2007), VRAG scores predicted psychiatrist recommendations and team recommendations, but not tribunal decisions. In other words, the influence of risk assessments may be weaker when assessors have an intermediary role rather than a direct influence on risk management.

Match to the risk principle may also vary depending on the characteristics of the individuals being assessed, such as gender. However, results are unclear. Storey and Strand (2013) found lower match to the risk principle for female than male offenders, but Singh, Desmarais, Sellers, et al. (2014) found higher match for females, and Holloway (2015) found no significant gender differences. Finally, match may vary based on the nature of the decision. For instance, whereas studies generally found a high match in *planned* risk management strategies, match was lower for studies that examined *actual* number of supervision contacts (Auditor General of Ontario, 2012) or services *received* (McCormick, Peterson-Badali, & Skilling, 2017). This slippage could be due to various factors; for instance, professionals may not follow through on their recommendations and/or high-risk individuals might have lower engagement and attendance.

Summary and Overall Strength of Evidence. Overall, the research in this area suggests moderate match to the risk principle following the use of risk assessment tools. However, based on the AHRQ system (Berkman et al., 2015), there was insufficient evidence to determine if tools *improve* match, as only two studies directly tested this (see Table 7). Next, we examined match to the need principle.

4. Does the Use of Risk Assessment Tools Facilitate Match to the Need Principle?

Study Characteristics. According to the need principle, risk management strategies should target criminogenic needs that contribute to each individual's offending (i.e., antisocial personality pattern, procriminal attitudes, procriminal associates, substance abuse, family/marital, school/work, leisure/recreation; Bonta & Andrews, 2017). We identified 17 non-overlapping studies (including five unpublished studies) in which researchers examined match to the need principle following the use of risk assessment tools; these studies included a total of 8,202 assessees and 1,278 professionals (see Table 4). In most studies (58.8%, $k = 10$), match was rated at the stage of the case management or intervention plan (e.g., services planned), such as whether professionals mentioned in their plans that they intended to address specific needs that were identified in the risk assessment (e.g., stating that they would refer a person with substance use difficulties to substance use treatment). However, in some studies, researchers examined whether assessees actually received these services (23.5%, $k = 4$), such as whether people who were identified as having substance use difficulties obtained substance use treatment. Finally, in two studies, researchers examined audiotaped PO-offender probation sessions to determine whether POs discussed and addressed needs that were identified in the risk assessment in their probation sessions with offenders (11.8%, $k = 2$). Most identified studies were with offender populations ($n = 13$); only one study was with psychiatric patients. Eight studies (47.1%) focused on adolescents (i.e., 18 or younger).

All of the studies examined tools which include dynamic or modifiable factors (100%, $k = 17$; see Appendix B). For instance, eight studies included versions of the LSI (e.g., LSI-R, YLS/CMI). In addition, most of the included tools have been found to significantly predict offending and/or violence (82.4%, $k = 14$). In two studies, the name of the tool was not reported or the researchers examined tools in general rather than a specific tool.

Study Appraisal. Most studies had medium (64.7%, $k = 11$) or high (29.4%, $k = 5$) limitations (see Table 7). The most common limitation was a lack of comparison groups. None of the studies had a comparison group of individuals who were not assessed with a risk assessment tool. Also, given the lack of standardized approaches to measure match to the need principle, researchers often developed their own measurement approaches. Many of these measurement approaches had limited or no research support (64.7%, $k = 11$), and descriptions of measures were often vague or missing (64.7%, $k = 11$). For instance, in many studies, researchers rated match as adequate or inadequate using criteria that they developed (70.6%, $k = 12$); inferential statistics were rarely used (11.8%, $k = 2$). In addition, there often was little information about adherence to tools, such as whether professionals received training on the risk assessment tool (64.7%, $k = 11$).

Results. In nearly half of the studies (41.2%, $k = 7$), overall match to the need principle was low (e.g., less than half of needs were addressed; see Tables 4 and 7). In one study, for instance, only 1 to 5% of youths with high scores in Y-LSI family, education, drug, or peer domains received interventions in these domains (Flores et al., 2004). In another study, an average of 35% of youths' needs were addressed in probation services (Vieira, Skilling, & Peterson-Badali, 2009). In almost all remaining studies (52.9%, $k = 9$), match to the need principle was "moderate" following the use of risk assessment tools. However, we defined moderate fairly leniently to mean slightly more than half of needs [50-69%] were addressed (see Table 4 for rating criteria). For instance, in one study that was rated as showing moderate match,

intervention plans addressed 52% of youths' critical needs on the START:AV (Singh, Desmarais, Sellers, et al., 2014). There was only one study (5.9%) in which match to the need principle study was rated as high, and that study relied on POs' self-report. In that study, POs self-reported that they frequently targeted needs in their case management efforts (Miller & Maloney, 2013).

Potential Moderators. The results of published and unpublished studies were comparable. However, there was some evidence that rates of match to the need principle may be higher in initial decisions, such as intervention plans, than in later decisions, such as the actual receipt of services. For instance, Peterson-Badali et al. (2015) found that, when youth had needs in areas such as family or education/employment, clinicians commonly mentioned these needs in the assessment reports that they wrote for POs. However, after receiving these reports, POs addressed only 31% of youths' needs through the provision of services.

Match to the need principle may also vary depending on how the need principle is defined. For instance, some researchers interpreted over-servicing as a poor match to the need principle (i.e., providing low-risk individuals more services than needed; Luong & Wormith, 2011), whereas other researchers focused solely on underservicing (i.e., providing high-risk individuals fewer services than needed; Peterson-Badali et al., 2015). Similarly, some researchers rated any intervention that targeted a need as an adequate match (e.g., court order to refrain from substance use; Flores et al., 2004), whereas other researchers rated a match as adequate only if the intervention was adequate in *quality* and *intensity* (i.e., regular attendance at an evidence-based substance use program; Peterson-Badali et al., 2015). Finally, some needs may be addressed more commonly than others. However, patterns differed considerably across studies. For example, in some studies, rates of match to family needs was over 70% (Holloway, 2015; Luong & Wormith, 2011). In other research, it was 10% or less (Flores et al., 2004).

Summary and Overall Strength of Evidence. Overall, research suggests limited match to the need principle following the use of risk assessment tools. However, it is unclear if this is better, worse, or the same as what it would have been had a tool not been used, as none of the studies included a comparison group. As such, based on the AHRQ system (Berkman et al., 2015), the strength of evidence to determine if tools *improve* match was deemed insufficient (see Table 7). Whereas our findings, thus far, have focused on the impact of tools on *professionals'* behaviors, our next question focused on the impact of tools on *assesseees'* behaviors, namely whether tools reduce violence and offending among people in mental health and justice settings.

5. Does the Implementation of Risk Assessment Tools Reduce Violence or Offending?

Study Characteristics. After removing overlapping studies, 12 studies examined the association between risk assessment tools and either violence (91.7%, $k = 11$) and/or general offending (33.3%, $k = 4$). Most studies were published (83.3%, $k = 12$). Studies included a total of 7,350 patients or offenders (see Table 5). Many of the studies were with general psychiatric or forensic psychiatric populations (66.7%, $k = 8$), primarily adults (66.7%, $k = 8$). Some studies had rigorous designs. Specifically, three studies (25.0%) used cluster RCTs, wherein wards were randomly assigned to an intervention group (i.e., a risk assessment tool was used), or a comparison group (i.e., a tool was not used). The remaining studies used a non-randomized

design, such as a pre-post design (66.7%, $k = 8$), or a non-randomized comparison group design (8.3%, $k = 1$). However, two of these non-randomized studies used propensity score matching to help ensure that the intervention and comparison groups were similar in characteristics such as race, gender, and offense history (i.e., Guy, Vincent, Grisso, & Perrault, 2015; Vincent, Guy, Perrault, & Gershenson, 2016). Nearly all studies examined tools which include dynamic or modifiable factors (91.7%, $k = 11$; see Appendix B), but one study examined the PSA, a tool which has only historical factors. In addition, most of the included tools have been found to significantly predict offending and/or violence (83.3%, $k = 10$). However, we were unable to find support for the predictive validity of two tools (i.e., CANVAS, PSA). The findings from the cluster RCTs and non-randomized designs are presented separately, consistent with guidelines (Berkman et al., 2015).

A. Cluster RCTs.

Study Appraisal. All of the cluster RCT studies were rated as having medium risk of bias (see Table 7). A common limitation was that, in each of the studies, professionals who conducted the risk assessment also collected data on violence or reoffending, thus creating the possibility of bias. Adherence to the tool was often unclear or inadequate. For instance, in Troquete et al. (2013) only 65% of patients who were supposed to be assessed with the START were assessed. In addition, despite randomization, the groups were not fully comparable in two studies. For example, in Abderhalden et al. (2008), patients in the wards that implemented the BVC-CH had significantly higher baseline levels of aggression than the comparison wards.

Findings. Each of the cluster RCTs involved tools that have support for their predictive validity (see Appendix B). In two of the three cluster RCTs (66.7%), the use of the risk assessment tool was associated with declines in aggressive incidents (see Tables 5 and 7). Both of these studies used the BVC, either alone (Abderhalden et al., 2008) or as part of a larger battery of tools (van de Sande et al., 2011). However, in one RCT (33.3%), findings were non-significant (Troquete et al., 2013). Given that this study found low adherence to the tool (i.e., the START), the authors reanalyzed results with only those who received the START. However, this did not alter the findings and, as such, the results do not appear fully attributable to poor adherence.

Potential Moderators. In two of the cluster RCTs, the tool was implemented alongside other risk management strategies (e.g., guidelines, policies). Although it was not directly tested, these strategies may have influenced results. Specifically, in Abderhalden et al. (2008), nurses were provided with a list of preventive measures for patients who scored above 7 on the BVC (e.g., relaxation exercise, one-to-one observation) and encouraged to consult with a multidisciplinary team for patients who scored above 10. Although this study reported decreases in aggressive incidents, it is unclear if this positive result would have been achieved had these guidelines not been in place. In Troquete et al. (2013), a process of joint care planning was introduced at the same time the START was implemented, wherein patients and case managers jointly developed a plan to address agreed upon treatment targets. The authors noted that the joint care plan may have led to a greater focus on treatment needs that may not be directly relevant to the reduction of violence, which could, in turn, have contributed to their null results.

Summary and Overall Strength of Evidence. Based on the AHRQ system (Berkman et al., 2015), there is insufficient evidence to draw a conclusion regarding whether tools reduce violence or offending because existing findings are mixed (see Table 7). Although RCTs should be given more weight than non-randomized designs (Oxford Centre for Evidence-based Medicine [OCEBM] Levels of Evidence Working Group, 2011), we also reviewed findings from non-randomized studies to determine if these studies yield similar results ($k = 9$).

B. Non-Randomized Studies.

Study Appraisal. Most of the non-randomized studies had moderate (55.6%, $k = 5$) limitations (see Table 7). Of the specific domains assessed, the most common limitation was with respect to comparison groups. Although all studies had comparison groups, these comparison groups were less than ideal (i.e., seven of the nine studies [77.8%] did not fully meet this criteria). For instance, studies generally did not test if the baseline risk levels of the pre- and post-samples differed, thus creating a possible confound; if individuals in the post-implementation sample were lower risk than those in the pre-implementation sample, it may appear as though the tool led to decreased offending when, in fact, this finding was due to differences in initial risk levels.

Findings. In two studies (22.2%), rates of violence or reoffending decreased following the implementation of a risk assessment tool (see Tables 5 and 7). For instance, in a study by the Laura and John Arnold Foundation (2014), rates of pretrial crime decreased by 15% following the implementation of a brief pretrial risk assessment tool, the PSA. No inferential tests were provided, and we were unable to find evidence for the predictive validity of this tool. Also, this study did not control for confounds. Thus, it is possible that the decrease in pretrial crime may be due to other factors, such as general decreases in crime rates, rather than the tool per se. Indeed, it may not be realistic to expect that a non-validated nine-item tool, focused on offense history, could directly reduce crime in and of itself.

In six studies (66.7%), there was no overall change in violence or reoffending when a risk assessment tool was implemented (i.e., violence or offending did not significantly change in all or almost all analyses). For instance, Vincent, Guy, Perrault and Gershenson (2016) conducted a pre-post study (using propensity score matching) at six adolescent probation offices. At five of the six sites, new petitions (i.e., charges) did not significantly decrease following the implementation of the SAVRY or YLS/CMI, even though these tools have been well-validated and include an emphasis on dynamic factors (see Appendix B).

In one study (11.1%), the findings were mixed. In particular, Guy, Vincent, Grisso, and Perrault (2015) compared sites that used the SAVRY (i.e., experimental condition) to sites that did not use any risk assessment tool (i.e., comparison condition). They found that adolescents in the experimental condition were less likely than those in the comparison condition to receive a new petition for a violent offense. However, there were no significant differences in new petitions for any offense, non-violent offenses, status offenses, or probation violations.

Potential Moderators. There was some evidence that risk assessment tools may have a larger impact in settings in which pre-existing base rates of violence are high. Specifically, in

one study, the implementation of a risk assessment protocol (i.e., the Alert System) was associated with larger initial reductions in violence in higher risk departments (e.g., psychiatry, emergency) than in lower risk departments (Kling, Yassi, Smailes, Lovato, & Koehoorn, 2011). Although not tested directly, the impact of risk assessment tools may also depend on the availability of appropriate treatment. As some authors noted, even if an individual was assessed to be high risk, it did not guarantee that they received the necessary treatment (Kling et al., 2011; Troquete et al., 2013). Finally, the impact of tools may be affected by jurisdictional policies and how offending is measured. For instance, Vincent et al. (2016) found that, at one site, new adjudications (i.e., convictions) significantly increased following the implementation of a tool, although new petitions (i.e., charges) did not. The authors explained that this might be due to a new district attorney, whose platform was to toughen sanctions for crime (Vincent, Guy, et al., 2012). In other words, the increase in new adjudications could reflect increased pressure to adjudicate adolescents rather than changes in actual rates of reoffending.

Summary and Overall Strength of Evidence. Similar to findings from the RCT studies, there was insufficient evidence from non-randomized studies to conclude that risk assessment tools reduce violence and offending (see Table 7). In sum, our results thus far indicate that use of risk assessment tools does not consistently improve assessees' outcomes, nor does it consistently improve professionals' risk management practices. Thus, as a final research question, we examined strategies that may help to rectify gaps between risk assessment and risk management.

6. What Strategies Help to Improve the Utility of Risk Assessment Tools for Risk Management?

Study Characteristics. We identified eight published studies ($N = 2,617$ offenders or patients, and 102 professionals; see Table 6), in which researchers examined strategies to improve the ability of risk assessment tools to either (a) facilitate adherence to the risk and/or need principles (75.0%, $k = 6$), and/or (b) enhance their effectiveness in reducing violence or offending (25.0%, $k = 2$). Three studies examined training (i.e., tools alone vs. tools plus training), three studies examined risk management guidelines (i.e., tools alone vs. tools plus guidelines), and two studies examined the impact of implementation strategies (e.g., policies). Five studies focused on POs or YPOs, whereas the other studies involved police and nurses. Most studies used a pre-post design (62.5%, $k = 6$). However, one study (12.5%) used an RCT design and two studies (25.0%) used a vignette design wherein professionals developed case plans based on vignettes. All tools included in these studies had support for their predictive validity, and included dynamic or modifiable factors (see Appendix B).

Study Appraisal. Most identified studies were rated as having medium (62.5%, $k = 5$) limitations. The most common limitation was with respect to the quality of comparison groups (only three studies fully met this criteria), as it was often unclear whether the comparison groups were sufficiently similar in variables such as risk level.

Findings. Of the three studies that examined the impact of training, training was associated with clear or partial improvements in risk management in two studies (66.7%, see Tables 6 and 7). For instance, Bonta et al. (2011) found that POs who completed an intensive

training program (i.e., Strategic Training in Community Supervision; STICS) spent more time discussing criminogenic needs, in general, and attitudes, in particular, with probationers. Although they were less likely than POs in the control group to discuss education/employment, within the STICS program, attitudes are prioritized as being more important to address than other needs. Furthermore, POs in the STICS program showed better intervention skills and their clients were less likely to reoffend. Storey, Gibas, Reeves and Hart (2011) found improved match to the need principle, but not the risk principle, following an eight-day course on risk management. In contrast to these generally positive findings, Needham et al. (2004) found that a five-day training course did not significantly enhance the ability of the BVC to reduce rates of aggressive incidents.

Of the three studies that examined the impact of risk management guidelines, guidelines were associated with some improvements in risk management in two studies (66.7%). Specifically, Bosker and colleagues compared results before and after guidelines were added to a risk assessment tool called the RISC (Bosker, Witteman, Hermanns, & Heij, 2015; Bosker & Witteman, 2016). They found that guidelines improved consistency in decision-making about risk and needs, and led to more appropriate levels of planned supervision. In contrast, Daffern et al. (2009) found that guidelines did not enhance the ability of the DASA to reduce aggression.

Finally, two studies from a large, multisite study found that higher quality implementation of tools was associated with improved match to the risk principle. For instance, Vincent, Guy, et al. (2012) found that, after professionals completed standard SAVRY training, there were no significant associations between risk level and decisions; it was only after policies and protocols were in place that risk level significantly predicted POs' management decisions.

Potential Moderators. Although research in this area is preliminary, one factor that may moderate the effectiveness of strategies, such as training, is staff engagement (Daffern et al., 2009). For instance, Bonta et al. (2008) found that POs who participated in ongoing support activities were more likely to benefit from the STICS program. Specifically, they were less likely than those with poor participation to focus on offenders' *non*criminogenic needs.

Summary. We did not complete an overall AHRQ strength of evidence rating for this question because this question encompassed a range of strategies (e.g., training, guidelines). However, these preliminary findings suggest that training, guidelines, and implementation efforts might improve match to the risk and need principles (see Table 7). That said, these approaches do not necessarily enhance the ability of tools to reduce violence and offending.

Discussion

Although risk assessment tools are thought to be an important starting point to risk management, it is unclear whether research evidence exists to support this view. As such, we conducted a systematic review of the research. Although we identified more relevant studies than anticipated ($k = 73$), research on the utility of tools for risk management is nevertheless scarce in comparison to research on the predictive validity of tools. Furthermore, our ability to draw conclusions from this research is limited for several reasons. First, much of the existing evidence on the impact of tools on professional practices is indirect. For instance, the vast majority of studies on adherence to the risk and need principles did not directly compare

professionals' risk management practices with and without a tool.

Second, many studies had medium or high limitations or possible confounds, as this type of research is challenging to conduct. For instance, adherence to risk assessment tools was often poor or unclear, which created a significant confound; it is difficult to meaningfully judge if tools can reduce violence if tools are not being used or are not used *properly*.

Finally, findings for some outcomes (e.g., violence and offending) were inconsistent. This suggests that important moderators could influence the effectiveness of risk assessment tools. However, at this point, it is unclear which variables may act as moderators, as few studies directly tested potential moderators. As an example, although many researchers have hypothesized that some tools may be better suited for risk management than others, for the most part, this has not yet been directly tested. As such, it is not possible to draw conclusions. With these caveats in mind, our key findings are discussed below.

Key Findings

Professionals' Reported Varying Opinions About the Utility of Tools. If professionals do not “buy-in” to tools or perceive them as useful, they may not adequately utilize them. Thus, as an initial step, we examined professionals' attitudes towards risk assessment tools. We found that, although some professionals held positive views about tools, in many studies, professionals had mixed views about the utility of tools for risk management (e.g., treatment-planning, placement decisions). This is not particularly surprising; professionals often feel reluctant to adopt new assessment and intervention approaches even when these approaches have strong research support (Jensen-Doss & Hawley, 2010; Lilienfeld, Ritschel, Lynn, Cautin, & Latzman, 2013). Furthermore, manuals and training for risk assessment tools often focus on how to complete item ratings rather than how to apply the tool to risk management efforts. As such, professionals' questions about the utility of tools may be understandable.

Use of Tools for Risk Management Was Mixed. Not only did professionals have mixed views about the utility of tools for risk management, in most of the identified studies, the use of risk assessment tools for risk management was mixed. Specifically, although some professionals reported that they relied on tools to guide their risk management decisions (e.g., decisions about services or placements), others reported that they did not use tools, even when employers mandated their use. As such, these findings illustrate that risk assessments do not necessarily flow through to risk management efforts. Slippage might be more likely to occur when risk assessors do not have direct control over risk management decisions, but instead act as intermediaries to decision-makers (e.g., judges). In such cases, the application of tools to risk management may depend not only on evaluators' use of tools, but also on whether subsequent decision-makers perceive tools as useful and relevant. In addition, future research should test whether the application of tools to risk management might differ by tool.

Adherence to the Risk Principle Was Moderate Following the Use of Tools. Despite the mixed application of tools to risk management efforts overall, match to the risk principle was moderate following the use of risk assessments tools. In a number of studies, high-risk individuals were referred to more services than low-risk individuals. They were also more likely

to receive secure placements. However, most studies did not have a comparison group of individuals who did not receive risk assessments. As such, it is difficult to determine if such findings are attributable to the use of the tool; some research suggests that high-risk individuals receive more intensive risk management strategies than low-risk individuals even when a risk assessment tool is not used (e.g., Campbell et al., 2015; Crocker, Braithwaite, Côté, Nicholls, & Seto, 2011). To address this possible confound, two studies directly compared match to the risk principle when a tool is and is not used (i.e., Bonta & Motiuk, 1990; Vincent, Paiva-Salisbury et al., 2012). These studies provide some initial evidence that tools may improve match. For instance, in one study, jails that used the LSI transferred a greater proportion of low-risk inmates to halfway houses compared to a jail that did not use the LSI (Bonta & Motiuk, 1990). This research is important as it suggests that, contrary to critics' concerns, tools might sometimes help reduce restrictive sanctions for individuals who are unlikely to reoffend. However, further research is needed, particularly studies with comparison groups in which risk assessment tools were and were not used.

Adherence to the Need Principle Was Limited Following the Use of Tools. Contrary to the positive findings relating to the risk principle, match to the need principle appeared limited following the use of risk assessment tools (match was rated as mixed or low in all but one study). This means that many of offenders' and patients' needs remained unaddressed even when risk assessment tools were used. This could indicate that professionals are not paying adequate attention to risk assessments when they are making decisions about services. Alternatively, these low rates of overall match could occur because professionals opt to focus on only a couple "high impact" needs at a given time, as it may not be feasible to simultaneously target all needs (Haqanee et al., 2015). Another possibility is that low rates of match occur because services to address needs are simply not available (Haqanee, Peterson-Badali, & Skilling, 2015). Clearly, identifying needs has limited value if there are no viable means by which to address these needs. Finally, low match to the needs principle may, in part, arise from limited compliance; offenders and patients may not necessarily attend or engage in the services to which they are referred.

Evidence on Whether Tools Reduce Violence and Offending Was Inconsistent. In light of the above findings, it is perhaps not surprising that evidence on whether tools reduce violence and offending was inconsistent. Although two RCTs found that the use of the BVC resulted in decreases in violence, another RCT did not find significant changes in violence or other criminal incidents when another tool, the START, was implemented. In addition, although two pre-post studies found that the implementation of risk assessment tools was associated with decreases in violence or offending, the bulk of pre-post studies ($k = 7$) did not. Thus, at present, there is insufficient evidence to conclude that tools reduce violence or offending.

One possible explanation for these findings is that it may be unrealistic to expect risk assessment tools to directly reduce violence or offending. As an analogy, would assessing whether someone is depressed directly reduce depression? Probably not. Instead, it may depend largely on what happens after the assessment. The same may be true for risk assessment tools; tools might be effective only if they enhance the likelihood that individuals receive appropriate, empirically-supported interventions.

Another possible explanation for the modest findings is that the effectiveness of tools

might vary by factors such as the setting, population, or tool. For instance, the two cluster RCT studies that found positive results (Abderhalden et al., 2008; van de Sande et al., 2011) were both conducted in acute psychiatric wards. This could suggest that tools may have a more positive impact in acute psychiatric settings than in other settings, perhaps because it is a more controlled environment. However, this may not entirely explain results because some studies in controlled psychiatric settings found null results (e.g., Gunenc, O'Shea, & Dickens, 2017).

It is also possible that some tools may be more effective than other tools. For instance, the two RCTs with positive results included the BVC, a brief tool that includes many dynamic or changeable risk factors. Again, though, this may not fully explain results because other studies in which other validated dynamic risk assessment tools were implemented (e.g., BVC-R, DASA, SAVRY, START, YLS/CMI) failed to find significant decreases in violence or offending.

Training, Guidelines, and Implementation Quality May Enhance the Utility of Tools. Overall, our results suggest that there is a disconnect between the theory of risk assessment and what actually happens in real-world practice (Peterson-Badali et al., 2015). However, although research is limited, preliminary evidence suggests that it may be possible to enhance professionals' risk management practices by combining the use of risk assessment tools with approaches such as risk management training (e.g., Bonta et al., 2011) and structured risk management guidelines (e.g., Bosker & Witteman, 2016). Though such approaches may not directly reduce violence and offending, they have been found to improve match to the risk and need principles, thus providing potential avenues by which to enhance the utility of tools. Sound implementation practices, such as policies and protocols to guide the use of tools, are also critical (e.g., Vincent et al., 2016). For instance, to assist agencies, Vincent, Guy, and Grisso (2012) have developed a guidebook for implementing risk assessment tools.

Limitations of Systematic Review

In interpreting our findings, several limitations of this review are important to note. Although we systematically searched 13 databases, reviewed reference lists, and contacted experts, our review likely missed some studies, such as studies written in languages other than English. In addition, although we attempted to summarize our findings with terms such as low, mixed, or high, definitions of such terms are somewhat subjective by nature. As such, to increase transparency and objectivity, we provided operational definitions of our summary terms, and summarized study findings in more detail using evidence tables (see Tables 1 – 6). In addition, two independent raters coded each study and we conducted consensus ratings. Another limitation of this systematic review is that there is a lack of appropriate tools for appraising risk of bias in risk assessment studies. As such, we drew items from other tools, and adapted the wording for this context. However, the approach that we used to appraise observational and survey studies was brief and, as such, our review likely failed to capture some relevant study limitations. Finally, although we examined differences in general patterns of results across published and unpublished studies, it is difficult to evaluate publication bias in narrative reviews.

Implications for Practice and Research

Our findings suggest that risk assessment tools are not sufficient to guarantee sound risk

management practices or reductions in violence (e.g., Peterson-Badali et al., 2015). Thus, researchers and tool developers should be careful to not overstate the potential value of risk assessment for risk management. Likewise, professionals, agencies, and policy-makers should not rely on risk assessment tools as their sole or primary risk management strategy. Instead, agencies who use risk assessment tools should work to build staff buy-in, regularly monitor adherence, and ensure that they are providing effective treatment, rather than funneling all their resources into assessment.

That said, even though risk assessment tools have limitations, they remain a best available practice. Although tools may not reduce violence or offending in and of themselves, there is no evidence that alternative approaches, such as assessing risk via unstructured clinical judgement or not assessing risk at all, do so either. For instance, one study found that when tools are not used, judges relied on unstructured risk assessments in making sentencing decisions, even though these assessments did not accurately predict violence (Lodewijks, Doreleijers, & de Ruiter, 2008). Indeed, numerous studies indicate that tools can help to predict violence and reoffending (e.g., Olver et al., 2014; Singh et al., 2011), and outperform unstructured approaches (Ægisdóttir et al., 2006; Hanson & Morton-Bourgon, 2009).

Based on this review, there are number of important areas for future research. First, many studies have lacked appropriate comparison groups, making it impossible to determine if tools improve practices per se. As such, there is a strong need for further research with comparison groups of individuals who were not assessed with a risk assessment tool, including studies with both mental health and justice populations. Second, to determine if certain tools may have a more beneficial impact on risk management than other tools, head-to-head comparisons of tools are needed. Third, adherence to risk assessment tools appears to be poor in some cases, making it difficult to evaluate the impact of tools. As such, research should routinely measure and report adherence. Fourth, given that the utility of tools for risk management likely depends heavily on what happens *after* the risk assessment, researchers examine the pathway between risk assessment and risk management, such as by developing and testing conceptual models. Finally, researchers should create and evaluate approaches to improve the utility of tools for risk management, such as training initiatives, structured risk management guidelines, and quality improvement or audit systems.

Conclusion

Overall, this review suggests that even though risk assessment tools may be a starting point for risk management, they are not sufficient in and of themselves (Vincent, Paiva-Salisbury, et al., 2012). Although some studies found positive results, indicating that tools might help achieve better match to the risk principle or even reductions in violence in some circumstances, the findings also revealed that “there is no guarantee that the results of these protocols flow through to front line service provision” (Haqanee et al., 2015, p. 39). For risk assessment researchers, the results of this review may be somewhat disappointing. Even though agencies pay “lip service” to using risk assessment to guide their risk management efforts (Skeem, 2013, p. 302), there is often “a lack of follow through between the assessment and case management” (Bonta et al., 2008, p. 266).

To ensure that risk assessment instruments are optimally used and do not degenerate into merely a bureaucratic exercise (Webster, 2011), further efforts are needed. In particular, rather

than focusing exclusively on predictive validity studies and the development of new tools, researchers need to pay greater attention to how tools are applied to guide real-world decisions, such as by testing the pathways between risk assessment and risk management, identifying areas of slippage, and developing strategies to facilitate the ability of risk assessments to translate into better risk management efforts. Such initiatives are essential to ensuring that the potential value of risk assessment is more fully realized.

References

- Abderhalden, C., Needham, I., Dassen, T., Halfens, R., Haug, H., & Fischer, J. E. (2008). Structured risk assessment and violence in acute psychiatric wards: Randomised controlled trial. *British Journal of Psychiatry*, *193*(1), 44-50. doi:10.1192/bjp.bp.107.045534
- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., & ... Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *Counseling Psychologist*, *34*(3), 341-382. doi:10.1177/0011000005285875
- Almond, T. (2012). Asset: An assessment tool that safeguards or stigmatizes young offenders? *Probation Journal*, *59*(2), 138-150. doi:10.1177/0264550512438255
- Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, *16*(1), 39-55. doi:10.1037/a0018362
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*(1), 19-52. doi:10.1177/0093854890017001004
- Auditor General of Ontario. (2012). *Annual report: Section 3.13. Youth Justice Services Program*. Retrieved from http://www.auditor.on.ca/en/reports_2012_en.htm
- Belfrage, H., Strand, S., Storey, J. E., Gibas, A. L., Kropp, P. R., & Hart, S. D. (2012). Assessment and management of risk for intimate partner violence by police officers using the Spousal Assault Risk Assessment Guide. *Law and Human Behavior*, *36*(1), 60-67. doi:10.1037/h0093948
- Berkman, N. D., Lohr, K. N., Ansari, M. T., Balk, E. M., Kane, R., McDonagh, M., & ... Chang, S. (2015). Grading the strength of a body of evidence when assessing health care interventions: An EPC update. *Journal of Clinical Epidemiology*, *68*(11), 1312-1324. doi:10.1016/j.jclinepi.2014.11.023
- Boccaccini, M. T., Turner, D. B., Murrie, D. C., Henderson, C. E., & Chevalier, C. (2013). Do scores from risk measures matter to jurors? *Psychology, Public Policy, and Law*, *19*(2), 259-269. doi:10.1037/a0031354
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). New York, NY: Routledge.
- Bonta, J., Bourgon, G., Rugge, T., Scott, T.-L., Yessine, A. K., Gutierrez, L., & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior*, *38*, 1127-1148. doi:10.1177/0093854811420678
- Bonta, J., & Motiuk, L. L. (1990). Classification to half-way houses: Quasi-experimental evaluation. *Criminology*, *28*(3), 497-506. doi:10.1111/j.1745-9125.1990.tb01336.x
- Bonta, J., Rugge, T., Scott, T., Bourgon, G., & Yessine, A. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation*, *47*, 248-270. doi:10.1080/10509670802134085
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley. doi:10.1002/9780470743386
- Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY)*. Odessa, FL: Psychological Assessment Resources.
- Bosker, J., & Witteman, C. (2016). Finding the right focus: Improving the link between

- risk/needs assessment and case management in probation. *Psychology, Public Policy, And Law*, 22(2), 221-233. doi:10.1037/law0000075
- Bosker, J., Witteman, C., Hermanns, J., & Heij, D. (2015). Improving agreement about intervention plans in probation by decision support. *International Journal of Offender Therapy & Comparative Criminology*, 59(13), 1459-1473. doi:10.1177/0306624X14539127
- Campbell, M. A., Canales, D. D., Wei, R., Totten, A. E., Macaulay, W. A., & Wershler, J. L. (2015). Multidimensional evaluation of a mental health court: Adherence to the risk-need-responsivity model. *Law and Human Behavior*, 39(5), 489-502. doi:10.1037/lhb0000135
- Carson, E. A., & Anderson, E. (2016). *Prisoners in 2015*. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Casey, P. M., Warren, R. K., & Elek, J. K. (2011). *Using offender risk and needs assessment information at sentencing: Guidance for courts from a National Working Group*. Williamsburg, VA: National Center for State Courts.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. doi:10.1037/1040-3590.6.4.284
- Cohen, J. (1960). Coefficient of agreement for nominal scales. *Educational & Psychological Measurement*, 20(1), 37-46. doi:10.1177/001316446002000104
- Crocker, A. G., Braithwaite, E., Côté, G., Nicholls, T. L., & Seto, M. C. (2011). To detain or to release? Correlates of dispositions for individuals declared not criminally responsible on account of mental disorder. *Canadian Journal of Psychiatry*, 56(5), 293-302. doi:10.1177/070674371105600508
- Daffern, M., Howells, K., Hamilton, L., Mannion, A., Howard, R., & Lilly, M. (2009). The impact of structured risk assessments followed by management recommendations on aggression in patients with personality disorder. *Journal of Forensic Psychiatry and Psychology*, 20(5), 661-679. doi:10.1080/14789940903173990
- Douglas, K. S., Hart, S. D., Groscup, J. L., & Litwack, T. R. (2013). Assessing violence risk. In I. B. Weiner, R. K. Otto (Eds.), *The handbook of forensic psychology* (4th ed.) (pp. 385-441). Hoboken, NJ: Wiley.
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 (Version 3): Assessing risk for violence*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29(5), 617-658. doi:10.1177/009385402236735
- Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, 11(3), 347-383. doi:10.1037/1076-8971.11.3.347
- Fitzgibbon, D. W. M. (2007). Risk analysis and the new practitioner: Myth or reality? *Punishment & Society*, 9, 87-97. doi:10.1177/1462474507070554
- Flores, A. W., Travis, L. F., & Latessa, E. J. (2004). *Case classification for juvenile corrections: An assessment of the youth level of service/case management inventory (YLS/CMI), final report* (Document No. 204005). Cincinnati, OH: University of Cincinnati.
- Gaynes, B. N., Brown, C. L., Lux, L. J., Brownley, K. A., Van Dorn, R. A., Edlund, M. J., & ...

- Lohr, K. N. (2017). Preventing and de-escalating aggressive behavior among adult psychiatric patients: A systematic review of the evidence. *Psychiatric Services*, *68*(8), 819-831. doi:10.1176/appi.ps.201600314
- Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *Journal of Chiropractic Medicine*, *5*(3), 101-117. doi:10.1016/S0899-3467(07)60142-6
- Gunenc, C., O'Shea, L. E., & Dickens, G. L. (2017). Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: Use of the START. *Criminal Behaviour and Mental Health*, doi:10.1002/cbm.2036
- Guy, L. S. (2008). Performance indicators of the structured professional judgment approach for assessing risk for violence to others: A meta-analytic survey (Order No. NR58733). Available from ProQuest Dissertations & Theses Global. (89265011). Retrieved from <http://ezproxy.library.ubc.ca/login?url=http://search.proquest.com/docview/89265011?accountid=14656>
- Guy, L. S., Kusaj, C., Packer, I. K., & Douglas, K. S. (2015). Influence of the HCR-20, LS/CMI, and PCL-R on decisions about parole suitability among lifers. *Law and Human Behavior*, *39*(3), 232-243. doi:10.1037/lhb0000111
- Guy, L. S., Nelson, R. J., Fusco-Morin, S. L., & Vincent, G. M. (2014). What do juvenile probation officers think of using the SAVRY and YLS/CMI for case management, and do they use the instruments properly? *International Journal of Forensic Mental Health*, *13*(3), 227-241. doi:10.1080/14999013.2014.939789
- Guy, L. S., Vincent, G. M., Grisso, T., & Perrault, R. (2015). *Advancing use of risk assessment in juvenile probation*. Technical Report submitted to the Office of Juvenile Justice and Delinquency Prevention. Worcester, MA: University of Massachusetts.
- Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., & ... Schünemann, H. J. (2011). GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology*, *64*(4), 383-394. doi:10.1016/j.jclinepi.2010.04.026
- Haas, S. M., & DeTardo-Bora, K. A. (2009, Spring). Inmate reentry and the utility of the LSI-R in case planning. *Corrections Compendium*, *11-16*, 49-52.
- Haddaway, N. R., Woodcock, P., Macura, B., & Collins, A. (2015). Making literature reviews more reliable through application of lessons from systematic reviews. *Conservation Biology*, *29*(6), 1596-1605. doi:10.1111/cobi.12541
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23-34. doi:10.20982/tqmp.08.1.p023
- Hannah-Moffat, K., Maurutto, P., & Turnbull, S. (2009). Negotiated risk: Actuarial illusions and discretion in probation. *Canadian Journal of Law & Society/Revue Canadienne Droit Et Societe (University of Toronto Press)*, *24*(3), 391-409.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, *36*(9), 865-891. doi:10.1177/0093854809338545
- Hanson, R. K., & Harris, A. R. (2000). Where should we intervene? Dynamic predictors of sexual assault recidivism. *Criminal Justice and Behavior*, *27*(1), 6-35. doi:10.1177/0093854800027001002
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments

- for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, 21(1), 1-21. doi:10.1037/a0014421
- Haqanee, Z., Peterson-Badali, M., & Skilling, T. (2015). Making 'what works' work: Examining probation officers' experiences addressing the criminogenic needs of juvenile offenders. *Journal of Offender Rehabilitation*, 54(1), 37-59. doi:10.1080/10509674.2014.980485
- Harcourt, B. E. (2007). *Against prediction: Profiling, policing, and punishing in an actuarial age*. Chicago, IL: University of Chicago Press.
- Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal & Criminological Psychology*, 3(1), 121-137. doi:10.1111/j.2044-8333.1998.tb00354.x
- Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University; Pacific Psychological Assessment Corporation; and the British Columbia Institute Against Family Violence.
- Hart, S. D., & Logan, C. (2011). Formulation of violence risk using evidence-based assessments: The structured professional judgment approach. In P. Sturmey & M. McMurrin (Eds.), *Forensic case formulation* (pp. 83–106). Chichester, England: Wiley.
- Hart, S. D., Webster, C. D., & Douglas, K. S. (2001). Risk management using the HCR-20: A general overview focusing on historical factors. In K. S. Douglas, C. D. Webster, S. D. Hart, D. Eaves, & J. R. P. Ogloff (Eds.), *HCR-20 violence risk management companion guide* (pp. 27-40). Burnaby, Canada/Tampa, FL: Simon Fraser University, Mental Health Law and Policy Institute/University of South Florida, Department of Mental Health Law & Policy.
- Hawley, C. J., Gale, T. M., Sivakumaran, T., & Littlechild, B. (2010). Risk assessment in mental health: Staff attitudes and an estimate of time cost. *Journal of Mental Health*, 19(1), 88–98. doi:10.3109/09638230802523005
- Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior*, 21(4), 347-359. doi:10.1023/A:1024851017947
- Higgins, J. T., Altman, D. G., Gøtzsche, P. C., Juni, P., Moher, D., Oxman, A. D., & ... Sterne, J. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Clinical Research Ed.)*, 343d5928. doi:10.1136/bmj.d5928
- Higgins, N., Watts, D., Bindman, J., Slade, M., & Thornicroft, G. (2005). Assessing violence risk in general adult psychiatry. *Psychiatric Bulletin*, 29(4), 131-133. doi:10.1192/pb.29.4.131
- Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior*, 25(4), 393-408. doi:10.1023/A:1010607719239
- Holloway, E. D. (2015). *Risk assessment and case planning by juvenile probation officers: The role of trauma*. (M.A., Fordham University). ProQuest Dissertations and Theses. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/1733682593?accountid=13800>
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *Plos Medicine*, 2(8), 696-701. doi:10.1371/journal.pmed.0020124

- Jensen-Doss, A., & Hawley, K. M. (2010). Understanding barriers to evidence-based assessment: Clinician attitudes toward standardized assessment tools. *Journal of Clinical Child and Adolescent Psychology, 39*(6), 885-896. doi:10.1080/15374416.2010.517169
- Johnson A., Sandford J., & Tyndall J (2003). Written and verbal information versus verbal information only for patients being discharged from acute hospital settings to home. *Cochrane Database of Systematic Reviews* (4). doi:10.1002/14651858.CD003716
- Kaebel, D., & Bonczar, T. P. (2017). *Probation and parole in the United States, 2015*. Washington, DC: Bureau of Justice Statistics.
- Kane, R. L., Butler, M., & Ng, W. (2016). Examining the quality of evidence to support the effectiveness of interventions: An analysis of systematic reviews. *BMJ Open, 6*(5), 1-5. doi:10.1136/bmjopen-2016-011051
- Kewley, S., Beech, A., Harkins, L., & Bonsall, H. (2015). Effective risk management planning for those convicted of sexual offending. *Journal of Aggression, Conflict and Peace Research, 7*(4), 237-257. doi:10.1108/JACPR-05-2015-0171
- Kicinski, M., Springate, D. A., & Kontopantelis, E. (2015). Publication bias in meta-analyses from the Cochrane Database of Systematic Reviews. *Statistics in Medicine, 34*(20), 2781-2793. doi:10.1002/sim.6525
- Kling, R. N., Yassi, A., Smailes, E., Lovato, C. Y., & Koehoorn, M. (2011). Evaluation of a violence risk assessment system (the alert system) for reducing violence in an acute hospital: A before and after study. *International Journal of Nursing Studies, 48*(5), 534-539. doi:10.1016/j.ijnurstu.2010.10.006
- Koehler, J. A., Lösel, F., Akoensi, T. D., & Humphreys, D. K. (2013). A systematic review and meta-analysis on the effects of young offender treatment programs in Europe. *Journal of Experimental Criminology, 9*(1), 19-43. doi:10.1007/s11292-012-9159-7
- Kropp, P. R., & Gibas, A. (2010). The Spousal Assault Risk Assessment Guide (SARA). In R. K. Otto, K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 227-250). New York, NY: Routledge.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. doi:10.2307/2529310
- Large, M. M., Ryan, C. J., Callaghan, S., Paton, M. B., & Singh, S. P. (2014). Can violence risk assessment really assist in clinical decision-making? *Australian and New Zealand Journal of Psychiatry, 48*(3), 286-288. doi:10.1177/0004867413498275
- Laura and John Arnold Foundation (2014). *Results from the First Six Months of the Public Safety Assessment – Court™ in Kentucky*. Laura and John Arnold Foundation.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. A., & ... Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Plos Medicine, 6*(7), 1-29. doi:10.1371/journal.pmed.1000100
- Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzman, R. D. (2013). Why many clinical psychologists are resistant to evidence-based practice: Root causes and constructive remedies. *Clinical Psychology Review, 33*(7), 883-900. doi:10.1016/j.cpr.2012.09.008
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims & Offenders, 4*(2), 124-147. doi:10.1080/15564880802612573

- Livsey, S. (2012). *Juvenile delinquency probation caseload, 2009*. U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- Lodewijks, H. P. B., Doreleijers, T. A. H., & de Ruiter, C. (2008). SAVRY risk assessment in violent Dutch adolescents. *Criminal Justice & Behavior, 35*(6), 696-709. doi:10.1177/0093854808316146
- Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M. (2006). The risk principle in action: What have we learned from 13,676 offenders and 97 correctional programs? *Crime & Delinquency, 52*(1), 77-93. doi:10.1177/0011128705281747
- Luong, D., & Wormith, J. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior, 38*(12), 1177-1199. doi:10.1177/0093854811421596
- Maurutto, P., & Hannah-Moffat, K. (2007). Understanding risk in the context of the Youth Criminal Justice Act. *Canadian Journal of Criminology & Criminal Justice, 49*(4), 465-491. doi:10.3138/cjccj.49.4.465
- McCormick, S., Peterson-Badali, M., & Skilling, T. (2017). Mental health and specific responsibility: Its role in juvenile justice rehabilitation. *Law & Human Behavior, 41*(1), 55-67. <http://dx.doi.org/10.1037/lhb0000228>
- McGrath, R., J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American Survey*. Brandon, VT: Safer Society Press.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46. doi:10.1037/1082-989X.1.1.30
- McKee, S. A., Harris, G. T., & Rice, M. E. (2007). Improving forensic tribunal decisions: The role of the clinician. *Behavioral Sciences & the Law, 25*(4), 485-506. doi:10.1002/bsl.768
- Miller, J., & Maloney, C. (2013). Practitioner compliance with risk/needs assessment tools: A theoretical and empirical assessment. *Criminal Justice and Behavior, 40*(7), 716-736. doi:10.1177/0093854812468883
- Monahan, J., & Skeem, J. L. (2014). Risk redux: The resurgence of risk assessment in criminal sentencing. *Federal Sentencing Reporter, 26*(3), 158-166. doi:10.1525/fsr.2014.26.3.158
- Muheizen, J. (2014). The effect of risk assessment data and colleagues' consensus on clinical decisions to discharge insanity acquittees. ProQuest Information & Learning. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 74* (9). Retrieved from <http://proxy.lib.sfu.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psych&AN=2014-99051-114&site=ehost-live>
- National Institute for Health and Care Excellence [NICE] (2012). *Methods for the development of NICE public health guidance* (3rd ed.). Retrieved from <http://nice.org.uk/process/pmg4>
- Needham, I., Abderhalden, C., Meer, R., Dassen, T., Haug, H. J., Halfens, R. G., & Fischer, J. E. (2004). The effectiveness of two interventions in the management of patient violence in acute mental inpatient settings: Report on a pilot study. *Journal of Psychiatric and Mental Health Nursing, 11*(5), 595-601. doi:10.1111/j.1365-2850.2004.00767.x
- Nielsen, O., Ryan, C., & Large, M. (2011). Risk assessment and resource allocation. *Australasian Psychiatry: Bulletin of Royal Australian and New Zealand College of Psychiatrists, 19*(3), 270. doi:10.3109/10398562.2011.561849
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior, 36*(4),

- 329-353. doi:10.1177/0093854809331457
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, 26(1), 156-176. doi:10.1037/a0035080
- Oxford Centre for Evidence-Based Medicine [OCEBM] Levels of Evidence Working Group (2011). *The Oxford 2011 "Levels of Evidence"*. Oxford Centre for Evidence-Based Medicine. Retrieved from <http://www.cebm.net/index.aspx?o=5653>
- Peterson-Badali, M., Skilling, T., & Haqanee, Z. (2015). Examining implementation of risk assessment in case management for youth in the justice system. *Criminal Justice and Behavior*, 42(3), 304-320. doi:10.1177/0093854814549595
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ... Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews – A product from the ESRC Methods Programme*. Lancaster University. Retrieved from: <http://www.lancaster.ac.uk/shm/research/nssr/research/dissemination/publications.php>
- Prendergast, M. L., Pearson, F. S., Podus, D., Hamilton, Z. K., & Greenwell, L. (2013). The Andrews' principles of risk, needs, and responsivity as applied in drug treatment programs: Meta-analysis of crime and drug use outcomes. *Journal of Experimental Criminology*, 9(3), 275-300. doi:10.1007/s11292-013-9178-z
- Rice, M. E., Harris, G. T., & Hilton, N. Z. (2010). The Violence Risk Appraisal Guide and Sex Offender Risk Appraisal Guide for Violence Risk Assessment and the Ontario Domestic Assault Risk Assessment and Domestic Violence Risk Appraisal Guide for Wife Assault Risk Assessment. In R. K. Otto, K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 99-119). New York, NY: Routledge/Taylor & Francis Group.
- Richardson, C. L. (2009). *Structured professional judgement of risk in forensic psychiatric practice* (Order No. U504885). Available from ProQuest Dissertations & Theses A&I. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/899736093?accountid=13800>
- Rose, N. (1998). Governing risky individuals: The role of psychiatry in new regimes of control. *Psychiatry, Psychology and Law*, 5(2), 177-195. doi:10.1080/13218719809524933
- Royal College of Psychiatrists (2008). *Rethinking risk to others in mental health services: Final report of a scoping group* (CR150). Royal College of Psychiatrists.
- Roychowdhury, A., & Adshear, G. (2014). Violence risk assessment as a medical intervention: ethical tensions. *Psychiatric Bulletin*, 38(2), 75-82. doi:10.1192/pb.bp.113.043315
- Ryan, R. & Cochrane Consumers and Communication Review Group (2013, June). *Cochrane Consumers and Communication Review Group: Data synthesis and analysis*. Retrieved from <http://cccr.org>
- Ryan, C., Nielsen, O., Paton, M., & Large, M. (2010). Clinical decisions in psychiatry should not be based on risk assessment. *Australasian Psychiatry: Bulletin of Royal Australian and New Zealand College of Psychiatrists*, 18(5), 398-403. doi:10.3109/10398562.2010.507816
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., & ... Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7(10), 1-7. doi:10.1186/1471-2288-7-10
- Shook, J. J., & Sarri, R. C. (2007). Structured decision making in juvenile justice: Judges' and

- probation officers' perceptions and use. *Children and Youth Services Review*, 29, 1335–1351. doi:10.1016/j.chilyouth.2007.05.008
- Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., ... Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health*, 13(3), 193-206. doi:10.1080/14999013.2014.922141
- Singh, J. P., Desmarais, S. L., Sellers, B. G., Hylton, T., Tirotti, M., & Van Dorn, R. A. (2014). From risk assessment to risk management: Matching interventions to adolescent offenders' strengths and vulnerabilities. *Children and Youth Services Review*, 47(Part 1), 1-9. doi:10.1016/j.chilyouth.2013.09.015
- Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, 31(3), 499–513. doi:10.1016/j.cpr.2010.11.009
- Skeem, J. (2013). Risk technology in sentencing: Testing the promises and perils. *Justice Quarterly*, 30(2), 297-303. doi:10.1080/07418825.2012.687513
- Skeem, J. L., & Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology*, 54(4), 680-712. doi:10.1111/1745-9125.12123
- Skeem, J. L., & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, 20(1), 38-42. doi:10.1177/0963721410397271
- Starr, S. (2014). Evidence-based sentencing and the scientific rationalization of discrimination. *Stanford Law Review*, 66(4), 803-872.
- Storey, J. E., Gibas, A. L., Reeves, K. A., & Hart, S. D. (2011). Evaluation of a violence risk (threat) assessment training program for police and other criminal justice professionals. *Criminal Justice and Behavior*, 38(6), 554-564. doi:10.1177/0093854811403123
- Storey, J. E., & Strand, S. (2013). Assessing violence risk among female IPV perpetrators: An examination of the B-SAFER. *Journal of Aggression Maltreatment & Trauma*, 22(9), 964-980. doi:10.1080/10926771.2013.835015
- Szmukler, G., & Rose, N. (2013). Risk assessment in mental health care: Values and costs. *Behavioral Sciences & the Law*, 31(1), 125-140. doi:10.1002/bsl.2046
- Taxman, F. S., Cropsey, K. L., Young, D. W., & Wexler, H. (2007). Screening, assessment, and referral practices in adult correctional settings: A national perspective. *Criminal Justice and Behavior*, 34(9), 1216-1234. doi:10.1177/0093854807304431
- Troquete, N. A. C., van, d. B., Beintema, H., Mulder, T., van Os, T. W. D. P., Schoevers, R. A., & Wiersma, D. (2013). Risk assessment and shared care planning in out-patient forensic psychiatry: Cluster randomised controlled trial. *British Journal of Psychiatry*, 202(5), 365-371. doi:10.1192/bjp.bp.112.113043
- Undrill, G. (2007). The risks of risk assessment. *Advances in Psychiatric Treatment*, 13(4), 291–297. doi:10.1192/apt.bp.106.003160
- Urquhart, T. A., & Viljoen, J. L. (2014). The use of the SAVRY and YLS/CMI in adolescent court proceedings: A case law review. *International Journal of Forensic Mental Health*, 13(1), 47-61. doi:10.1080/14999013.2014.885470
- van de Sande, R., Nijman, H. I., Noorthoorn, E. O., Wierdsma, A. I., Hellendoorn, E., van der Staak, C., & Mulder, C. L. (2011). Aggression and seclusion on acute psychiatric wards: Effect of short-term risk assessment. *British Journal of Psychiatry*, 199(6), 473-478. doi:10.1192/bjp.bp.111.095141
- Vieira, T. A., Skilling, T. A., & Peterson-Badali, M. (2009). Matching court-ordered services

- with treatment needs: Predicting treatment success with young offenders. *Criminal Justice and Behavior*, 36(4), 385-401. doi:10.1177/0093854808331249
- Viglione, J., Rudes, D. S., & Taxman, F. S. (2015). Misalignment in supervision: Implementing risk/needs assessment instruments in probation. *Criminal Justice and Behavior*, 42(3), 263-285. doi:10.1177/0093854814548447
- Viljoen, J. L., McLachlan, K., & Vincent, G. M. (2010). Assessing violence risk and psychopathy in juvenile and adult offenders: A survey of clinical practices. *Assessment*, 17(3), 377-395. doi:10.1177/1073191109359587
- Vincent, G. M., Guy, L. S., Gershenson, B. G., & McCabe, P. (2012). Does risk assessment make a difference? Results of implementing the SAVRY in juvenile probation. *Behavioral Sciences & the Law*, 30(4), 384-405. doi:10.1002/bsl.2014
- Vincent, G. M., Guy, L. S., & Grisso, T. (2012). *Risk assessment in juvenile justice: A guidebook for implementation*. Retrieved from http://njn.org/uploads/digital-library/Risk_Assessment_in_Juvenile_Justice_A_Guidebook_for_Implementation.pdf
- Vincent, G. M., Guy, L. S., Perrault, R. T., & Gershenson, B. (2016). Risk assessment matters, but only when implemented well: A multisite study in juvenile probation. *Law and Human Behavior*, 40(6), 683-696. doi:10.1037/lhb0000214
- Vincent, G. M., Paiva-Salisbury, M. L., Cook, N. E., Guy, L. S., & Perrault, R. T. (2012). Impact of risk/needs assessment on juvenile probation officers' decision making: Importance of implementation. *Psychology, Public Policy, and Law*, 18(4), 549-576. doi:10.1037/a0027186
- Wachter, A. (2015). *Statewide risk assessment in juvenile probation: JJGPS StateScan*. Pittsburgh, PA: National Center for Juvenile Justice.
- Wand, T. (2012). Investigating the evidence for the effectiveness of risk assessment in mental health care. *Issues in Mental Health Nursing*, 33(1), 2-7. doi:10.3109/01612840.2011.616984
- Webster, C. D. (2011). From flipping the coin to seeing both its sides. *Journal of Psychiatric Intensive Care*, 7(2), 73-78. doi:10.1017/S174264641100015X
- Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. L. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START) (Version 1.1)*. Coquitlam, Canada: Forensic Psychiatric Services Commission and St. Joseph's Healthcare.
- Wong, S. P., & Olver, M. E. (2010). Two treatment- and change-oriented risk assessment tools: The Violence Risk Scale and Violence Risk Scale-Sexual Offender Version. In R. K. Otto, K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 121-146). New York, NY: Routledge.
- Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136(5), 740-767. doi:10.1037/a0020473

Table 1
Do Professionals Perceive Risk Assessment Tools to be Useful for Risk Management?

Authors, Year (country)	Sample	Risk Tool	Design	Results	Study Appraisal	Summary of Results
Bonta et al., 2005 (CA)	482 judges, POs, lawyers	Tools in general	Mail survey, interviews	Judges/POs rated usefulness of risk assessments in presentencing reports as $M = 8.1-9.1$ (of 10)	**	High
Flores et al., 2004 (USA)	195 corrections staff	Y-LSI	Mail survey	Rated usefulness for placements as $M = 5.2$ (of 10) and usefulness for identifying treatment needs as $M = 5.5$	**	Mixed
Flores, 2013 (USA)	185 corrections staff	YLS/CMI	Mail survey	Rated usefulness for placement decisions, identifying treatment needs, and case planning as $M = 5.3 - 5.7$ (of 10)	**	Mixed
Guy et al., 2014 (USA)	71 YPOs	SAVRY YLS/ CMI	Interviews	63 – 71% rated SAVRY as very helpful for service, disposition, and supervision decisions; 41 – 50% rated YLS/CMI as very helpful for these decisions	**	Mixed-High
Mair et al., 2016 (UK)	180 POs	OASys	Interviews	36% stated that tool helps produce better supervision plans	*	Low
McKee et al., 2007 (CA)	157 forensic staff	VRAG	In-person survey	Rated importance of tool in transfer decisions as $M = 7.8$ (of 10); tools seen as less important than index offense, etc.	**	Mixed
Muheizen, 2014 (USA)	433 clinicians	VRAG	Online survey	78% rated tool as very helpful to discharge decisions	***	High
Richardson, 2009 (UK)	45 clinical staff	SCJ:Risk	In-person survey	72% rated tool as useful for informing clinical practice, 78% rated it as useful for risk management plan	***	High
Royal College of Psychiatrists, 2008 (UK)	1937 psychiatrists	Tools in general	Online survey	48% reported that tools lead to better decisions	*	Low
Schneider et al., 1996 (USA)	179 POs	WRNA	Mail survey	37% stated that tool helps supervision decisions/caseload management; 53% indicated it helps high risk offenders receive supervision	***	Low- Mixed
Shook & Saari, 2007 (USA)	665 POs and judges	Tools in general	Survey	Rated usefulness for pretrial detention, placement, and release decisions as $M = 3.06 - 3.71$ (of 5); 36% judges/41% POs stated that tools ensure appropriate placement	***	Mixed
Singh et al., 2014 (international)	2135 clinicians, 44 countries	Various tools	Online survey	Rated usefulness for risk management as $M = 3.2 - 4.5$ (of 6); results varied by tool	***	Mixed – High

Note. The summary ratings were: Low (participants rated helpfulness as being below the midpoint [e.g., less than half believed the tool aided in risk management]), Mixed (participants rated helpfulness as being just above the midpoint [e.g., 5-7 out of 10, with 10 being the most positive]), High (most participants rated tool as helpful [e.g., mean of >7 out of 10, with 10 being most positive]). The study appraisal ratings were: * (high limitations [<50% on the study appraisal form]), ** (medium limitations [51 – 79% on the study appraisal form]), *** (low limitations [>80% on the

study appraisal form]). CA = Canada; N.R. = not reported; PO = probation or parole officer; *M* = mean; U.K. = United Kingdom; USA = United States of America; YPO = youth probation officer. Sample size refers to the sample used for the relevant analyses.

Table 2

When Professionals Conduct Risk Assessments with Tools, Do These Assessments Guide Risk Management Efforts?

Authors, Year (country)	Sample	Risk Tool	Design	Results	Study Appraisal	Summary of Results
A. Self-Reported Use						
Flores et al., 2004 (USA)	195 corrections staff	Y-LSI	Mail survey	57% used tool to identify treatment goals, 71% used it to gauge treatment effectiveness, 80% used it to develop case plans, 86% used it to inform supervision decisions	**	Mixed – High
Flores, 2013 (USA)	185 corrections staff	YLS/CMI	Mail survey	82% used tool to develop case plans	**	High
Haas & DeTardo-Bora, 2009 (USA)	94 CMs, POs, counsellors	LSI-R	Mail survey	42% used tool to develop reentry plans, 67% to assess treatment progress, 78% to develop treatment plan, 79% to set supervision level, and 80% for referrals	**	Mixed – High
Hawley et al., 2010 (UK)	300 nurses, doctors, etc.	Tools in general	In-person survey	27% used tool less than half the time, 21% used it half the time, and 52% used it more than half the time	*	Mixed
Krysiak & LeCroy, 2002 (USA)	80 POs, judges	NCCD tool†	Survey	36% of POs used tool to make court recommendations, 42% of judges said tool played a significant role in decisions	*	Low
Miller & Maloney, 2013 (USA)	1087 community corrections staff	Tools in general	Email survey	48% completed the tool and applied it to decision-making, 40% completed the tool but didn't fully apply it to decisions, and 12% neither completed nor applied the tool	**	Low – Mixed
Shook & Sarri, 2007 (USA)	665 POs, judges, etc.	Tools in general	Survey	Judges: 45-48% routinely used risk and needs assessments in dispositions; POs: 64-72% routinely used risk and needs assessment in dispositions	***	Mixed
Singh et al., 2014 (international)	2135 clinicians, 44 countries	Various tools	Online survey	69% used tools to develop risk management plans; in 35% of known cases, plans were not implemented	***	Mixed
Vincent, Paiva-Salisbury et al., 2012 (USA)	88 YPOs, 23 managers	SAVRY, YLS/CMI (YPOs)	Interview	80% used tools in service decisions, 77% in disposition decisions, and 90% in supervision decisions	***	High
B. Observational Studies						
Boccaccini et al., 2013 (USA)	299 jurors	Static-99, MnSOST-R	In-person	No significant relationship between risk scores and jurors' perception of offenders' reoffense risk	***	Low
Kewley et al., 2015 (UK)	216 plans by POs	RM2000, OASys	File review	Many risk factors identified weren't mentioned in the plans; 8% of management plans referred to risk assessment	**	Low

RISK MANAGEMENT

42

Turner et al., 2015 (USA)	462 jurors	Static-99, MnSOST-R	In-person survey	Perceived influence of tools in decisions was $M = 4.32 - 4.98$ (of 6); tools had lower influence than most other factors (e.g., offenders' testimony perceived remorse)	***	Mixed
Urquhart & Viljoen, 2014 (international)	50 young offender cases with risk assessment tool	SAVRY, YLS/CMI	Case law review	46% of cases – judges directly referred to tool in decisions and 30% of cases – did not refer to tool but mentioned risk; 4% of cases – judges placed great weight on tool and 58% of cases – placed some weight on tool	***	Mixed
Viglione et al., 2015 (USA)	69 sessions with adult probationers	Tool N.R. † (POs)	Observation	POs referred to tool in 3% of interactions, and used tool for case planning in 1% of interactions	*	Low

Note. † No support for the predictive validity of the tool could be found. The summary ratings were: Low (less than half of participants used the tool for risk management), Mixed (just above half used the tool for risk management [e.g., 50 – 70%]), High (most participants used the tool for risk management [e.g., > 70%]). The study appraisal ratings were: * (high limitations [<50% on the study appraisal form]), ** (medium limitations [51 – 79% on the study appraisal form]), *** (low limitations [>80% on the study appraisal form]). CM = case manager; M = mean; N.R. = not reported; PO = probation or parole officer; UK = United Kingdom; USA = United States of America.

Table 3
Does the Use of Risk Assessment Tools Facilitate Match to the Risk Principle?

Authors, Year (country)	Sample (gender of evaluatees)	Risk Tool (assessor)	Design	Results	Study Appraisal	Summary of Results
A. Studies Comparing Tool vs. No Tool						
Bonta & Motiuk, 1990 (CA)	580 incarcerated adults (M)	LSI (staff)	Non- randomized comparison	In the jails that used tool, inmates with low risk totals were more likely to be transferred to a halfway house (51% vs. 16% at a control jail)	**	Improvement
Vincent, Paiva- Salisbury et al., 2012 (USA)	88 YPOs, 23 managers	SAVRY, YLS/CMI (YPOs)	Pre- post- interviews	After implementation, YPOs were more likely to spontaneously identify risk as important for their supervision decisions but not for their services or disposition decisions	***	Some Improvement
B. Studies with No Comparison Group						
1. Placement						
DeGue et al., 2008 (USA)	220 young offenders (M/F)	YLS/CMI (clinicians)	File review	Risk total predicted out-of-home placement recommendations ($t = 7.07$)	**	Moderate
Gebo, 2002 (USA)	46 incarcerated young offenders (M/F)	RAI† (N.R.)	File review	Risk score predicted dispositions (i.e., commitment) more strongly in district vs. family court (62% vs. 29%)	**	Low – Moderate
Guy et al. (2015)	539 youth on probation (M/F)	SAVRY, JAG (YPOs)	File review	SAVRY summary risk ratings and JAG total scores generally predicted out-of-home placements	***	Moderate
Jung et al., 2015 (CA)	165 adult forensic outpatients (M/F)	LSI, HCR-20 (psych.)	File review	HCR-20 and LSI total predicted incarceration ($\eta^2 = .10 - .11$) but not length of incarceration or probation	**	Moderate
Perrault et al., 2012 (USA)	64 YPOs	YLS/CMI, SAVRY (POs)	Interviews	90% of YPOS said they used risk level to inform placement decisions	**	High
Simpson, 2010 (USA)	202 young offenders (gender N.R.)	DSI† (police)	Pre-post	Risk total predicted detention decision ($r = .32$)	**	Moderate
Vincent et al., 2016 (USA)	847 youth on probation (M/F)	SAVRY (YPOs)	File review	Risk level predicted placements at 5 of 6 sites	***	High
2. Other Sentencing Decisions						
Blais & Forth, 2014 (CA)	120 adult inmates (gender N.R.)	VRAG, etc. (clinicians)	File review	Risk total (VRAG, Static-99, SORAG) did not predict dangerous offender designation (AUCs = .50 – .51)	*	Low
Hsieh, 2016 (USA)	2137 adults with sexual offenses (M)	Static-99 etc. (N.R.)	File review	Risk total <i>inversely</i> predicted sentence length but once other factors controlled, <i>positively</i> predicted sentence length	*	Moderate
3. Supervision						

RISK MANAGEMENT

44

Auditor General of Ontario, 2012 (CA)	Young offenders (<i>n</i> and gender N.R.)	Tool N.R. (YPOs)	File review	“Almost half” of youth met with POs less than required based on risk level	*	Low
Bonta et al., 2008 (CA)	154 adults/youth on probation (M/F)	PRA (POs)	File review	Risk total predicted # of probation contacts for adults ($r = .22$) but not for youth ($r = .09$)	**	Low – Moderate
Bosker et al., 2013b (NL)	44 POs, 173 plans	RISc-2 (POs)	Vignettes, wrote plans	Low agreement about intensity of supervision when given risk assessment results ($\kappa = .20$)	**	Low
Guy et al. (2015)	248 youth on probation (M/F)	SAVRY, JAG (YPOs)	File review	JAG and SAVRY scores were used to assign initial level of supervision, but risk ratings did not consistently predict supervision contacts received (at Mississippi site)	***	Low-Moderate
Luong & Wormith, 2011 (CA)	192 youth on probation (M/F)	Youth LSI-SK (POs)	File review	Risk total predicted supervision frequency ($r = .66$)	**	High
Maupin, 1993 (USA)	280 youth on parole (gender N.R.)	AJADMS† (POs)	File review	Risk classification predicted contacts/month ($R^2 = .12$) but inconsistent, with some findings in the opposite direction	**	Low – Moderate
Vincent et al., 2016 (USA)	847 youth on probation (M/F)	SAVRY (YPOs)	File review	Risk level predicted supervision level at 4 of 5 sites	***	High

4. Release or Discharge

Folino et al., 2004 (AR)	65 adult inmates & forensic patients (M)	HCR-20, VRAG (N.R.)	File review	Tool ratings were not correlated with judicial release decisions	*	Low
Guy et al., 2015 (USA)	5187 adult offenders, life sentence (M/F)	HCR-20, LS/CMI (psyc.)	File review	HCR-20 and LS/CMI totals predicted parole release decisions ($d = 1.09$ and 0.97 , respectively)	***	High
Hilton & Simmons, 2001 (CA)	187 forensic inpatients (gender N.R.)	VRAG (staff)	File review	Risk total did not predict tribunal decisions ($r = .06$), team recommendations ($r = .01$), or clinician testimony ($r = -.02$)	**	Low
McKee et al., 2007 (CA)	104 forensic inpatients (gender N.R.)	VRAG (N.R.)	File review	Risk total predicted psychiatrist recommendations ($r = .23$) & team recommendations ($r = .21$) but not tribunal decisions	*	Low – Moderate
Morrissey et al., 2014 (UK)	532 adult psychiatric patients (M/F)	HCR-20 (team)	File review	Risk total did not predict discharge but clinical scale total did	*	Low – Moderate
Muheizen, 2014 (USA)	433 clinicians	VRAG	Vignettes	78% less likely to recommend discharge if received high risk vignette (compared to low risk or no tool)	***	High

5. Overall Level or Intensity of Services

Campbell, 2015 (USA)	2739 youth on probation (M/F)	YLS/CMI (court officers)	File review	Risk total predicted number of program referrals ($t = 12.20$)	*	Moderate
Guy et al. (2015)	539 youth on probation (M/F)	SAVRY, JAG (YPOs)	File review	SAVRY summary ratings and JAG totals sometimes predicted number of referrals but did not predict services completed	***	Low

McCormick et al. (2017)	232 presentencing youth (M/F)	VRAG	File review	Risk total did not predict intensity of services (i.e., number of needs addressed in services) while on probation	**	Low
Perrault et al., 2012 (USA)	64 YPOs	YLS/CMI, SAVRY (POs)	Interviews	73% of POS said they used risk level to inform service referrals	**	High
Vincent et al., 2016 (USA)	847 youth on probation (M/F)	SAVRY (YPOs)	File review	Risk level predicted number of service referrals at 4 out of 6 sites	***	Moderate
6. Overall Level or Intensity of Intervention or Case Management Plans						
Belfrage, Strand, et al., 2012 (SE)	429 adult IPV offenders (M)	SARA (police)	File review	Risk total predicted number of recommended risk management strategies ($r = .40$)	**	Moderate
Bosker et al., 2013a (NL)	300 adults on probation (M/F)	RISc-2 (POs)	File review	Risk total predicted intensity of planned interventions ($r = .22$)	*	Moderate
Dyck, 2016 (CA)	100 diverted youth (M/F)	YLS/CMI (staff)	File review	After using tool, 60% of cases adhered to risk principle in case plans	**	Moderate
Holloway, 2015 (USA)	147 YPOs	YLS/CMI (YPOs)	Vignettes	30% of plans were appropriate intensity, 59% had some limitations, and 11% had substantial limitations	**	Low – Moderate
Kewley et al., 2015 (UK)	216 adult probationers with sexual offenses (M)	RM 2000, OASys (practitioner)	File review	Plans for <i>very</i> high risk probationers included more support than plans for high risk probationers (83% vs. 53%)	**	Moderate
Singh et al., 2014 (USA)	120 detained youth (M/F)	START:AV (CM)	File review	Risk total did not predict number of planned interventions for full sample ($r = .12$) but it did for girls ($r = .44$)	**	Low – Moderate
Storey et al, 2014 (SE)	249 adult IPV offenders (M)	B-SAFER (police)	File review	Risk total predicted number of recommended management strategies ($r = .43$)	**	Moderate
Storey & Strand, 2013 (SE)	52 adult IPV offenders (F)	B-SAFER (police)	File review	Risk total did not predict number of recommended management strategies	**	Low
Storey & Strand, 2017 (SE)	867 adult IPV offenders (M/F)	B-SAFER (police)	File review	Victim vulnerability scores predicted number of risk management strategies recommended (if victim was female)	**	Moderate
7. Other Strategies						
Belfrage & Strand, 2012 (SE)	216 adult IPV offenders (M)	B-SAFER (police)	File review	Police reported that they took more protective actions for individuals they rated as an imminent risk ($\tau = .30$)	**	Moderate
Miller & Maloney, 2013 (USA)	1087 POs	Variety of tools (POs)	Survey	POs reported that the extent to which they make decisions that correspond with risk level is $M = 4.5$ (of 6)	**	High
Trujillo & Ross, 2008 (AU)	501 IPV offenders (M/F; age N.R.)	L17A† (police)	File review	Risk total predicted whether action taken and intervention ordered (OR = 1.64 – 2.21), but not whether charges laid	**	Moderate

Note. † No support for the predictive validity of the tool could be found. The summary ratings were: Low (nonsignificant associations between risk total and intervention intensity, or many cases showed inadequate match to the risk principle [e.g., nearly half]), Moderate (significant associations between risk total and intervention intensity which fell in the small or moderate range [$r < .50$], or most cases showed adequate match to the risk

principle [e.g., 60-75%]), High (large and significant associations between risk total and intervention intensity, or almost all cases [e.g., >75%] showed adequate match to the risk principle), Improvements (improvements in all or most indicators of match to the risk principle), Some Improvements (improvements in some but not all of the indicators of match to the risk principle). The study appraisal ratings were: * (high limitations [<50% on the study appraisal form]), ** (medium limitations [51 – 79% on the study appraisal form]), *** (low limitations [>80% on the study appraisal form]). AU = Australia; AR = Argentina; CA = Canada; CM = case manager; F = female; IPV = intimate partner violence; M = male; M/F = male/female; N.R. = not reported; POs = parole or probation officer; Psyc = psychologists; SE = Sweden; UK = United Kingdom; USA = United States of America; YPO = youth probation officer.

Table 4
Does the Use of Risk Assessment Tools Facilitate Match to the Need Principle?

Authors, Year (country)	Sample (gender of evaluatees)	Risk Tool (assessor)	Design	Results	Study Appraisal	Summary of Results
A. Studies Comparing Tool vs. No Tool						
None						
B. Studies with No Comparison Group						
1. Needs Mentioned in Intervention or Case Management Plans						
Auditor General of Ontario, 2012 (CA)	Young offenders (<i>n</i> and gender N.R.)	Tool N.R.† (YPOs)	File review	>50% of case plans were missing goals for at least one high risk factor; 20% did not have goals for any high risk factor	*	Low
Bosker et al., 2013a (NL) ^c	300 adults on probation (M/F)	RISc-2 (POs)	File review	46% ^a of needs were adequately addressed in plans; ranged from 0% (friends) to 72% (cognitive skills)	**	Low
Bosker et al., 2013b (NL)	44 POs	RISc-2 (POs)	Vignettes, wrote plans	Agreement about the relevance of criminogenic needs fell in moderate range (kappa = .43)	**	Moderate
Dyck, 2016 (CA)	100 diverted youth (M/F)	YLS/CMI (officer)	File review	33% of needs were addressed in plans; ranged from 22% (criminal friends) to 47% (personality/behavior)	**	Low
Holloway, 2015 (USA)	147 young offenders (M/F)	YLS/CMI (YPOs)	Vignettes, wrote plans	60% of needs were addressed in plans; ranged from 37% (leisure/recreation) to 82% (family/parenting)	**	Moderate
Kewley et al., 2015 (UK)	216 adult probationers with sexual offenses (M)	OASys, RM2000 (practitioner)	File review	An average of 3.60 needs were identified in risk assessment but only 1.52 needs were addressed in plans	**	Low
Luong & Wormith, 2011 (CA)	192 youth on probation (M/F)	Youth LSI-SK (POs)	File review	63% ^a of needs were addressed in plans; ranged from 30% (education/ employment) to 76% (family/parenting); needs in <i>some</i> domains, but not all, predicted domain-specific services	**	Moderate
Sen et al., 2015 (UK)	88 adult psychiatric patients (M/F)	HCR-20 (team)	File review	Case managements plans were rated as $M = 3.9$ (of 8) on extent to which they addressed risk factors	*	Moderate
Singh et al., 2014 (USA)	120 detained youth (M/F)	START:AV (CMs)	File review	52% of critical needs were addressed in plans; ranged from 0% (e.g., coping) to 95% (substance use)	**	Moderate
Young et al., 2006 (USA)	3910-4592 young offenders (gender N.R.)	Intake Screen† (intake staff)	File review	28% of those with education needs received referral; 82% of those with substance abuse needs received referral	*	Moderate
2. Needs Addressed via Services Received						
Flores et al., 2004 (USA)	1679 young offenders (M/F)	Y-LSI	File review	1 – 5% of youth with needs in family, peer, education, or drug domains received treatment in these areas	*	Low

Haas et al. (2007) (USA)	348 incarcerated adults (M/F)	LSI-R (staff)	File review	53% of those with substance use needs received services; 35% with education/employment needs received services	**	Low – Moderate
Peterson-Badali et al, 2015 (CA)	148 youth on probation (M/F)	YLS/CMI (clinicians)	File review	Clinicians mentioned identified needs (e.g., family) in reports, but services weren't received (i.e., 40% of youth had no needs addressed); on average, 31% of needs were addressed	**	Low
Vieira et al., 2009 (CA)	122 young offenders (M/F)	YLS/CMI (clinicians)	File review	35% of youths' needs addressed while on probation	*	Low
3. Needs Discussed in Probation Sessions						
Bonta et al., 2008 (CA)	154 adults/youth probation sessions (M/F)	PRA (POs)	Audiotaped sessions	Mean % of identified needs discussed in session with adult offenders was 54% ^a ; ranged from 0% (academic/vocational) to 90% (family/marital)	**	Moderate
Bonta et al., 2011 (CA)	75 probation sessions (M/F) from 19 POs ^b	PRA (POs)	Audiotapes	POs spent 43% of session time talking about criminogenic needs and 49% of time on non-criminogenic needs	***	Low – Moderate
4. Needs Addressed Based on Professionals' Self-Report						
Miller & Maloney, 2013 (USA)	1087 POs	Tools in general (POs)	Survey	POs reported that the extent to which they target needs identified by tool is $M = 4.6$ (of 6)	**	High

Note. † No support for the predictive validity of the tool could be found. The summary ratings were: Low (overall, less than half of cases showed adequate match to the need principle, or less than half of needs were addressed), Moderate (overall, slightly more than half of cases [50-69%] showed adequate match to the need principle, or slightly more than half of needs were addressed), High (overall, most cases [>70%] showed adequate match to the need principle or most needs were addressed). The study appraisal ratings were: * (high limitations [<50% on the study appraisal form]), ** (medium limitations [51 – 79% on the study appraisal form]), *** (low limitations [>80% on the study appraisal form]). ^a Average match was calculated as follows: (total number of needs addressed in intervention plan/total number of needs identified by risk assessment tool); ^b This study also examined whether training improves match; those results are described in Table 6. CA = Canada; CM = case manager; M = male; M/F = male/female; NL = Netherlands; N.R. = not reported; PO = parole or probation officer; UK = United Kingdom; USA = United States of America; YPO = youth probation officer.

Table 5
Does the Implementation of Risk Assessment Tools Reduce Violence or Offending?

Authors, Year (country)	Sample (gender)	Risk Tool (assessors)	Design	Results	Study Appraisal	Summary of Results
A. RCTs						
Abderhalden et al., 2008 (CH)	2364 adult psyc. inpatients (M/F)	BVC + risk management (nurses)	Cluster RCT	Wards using tool showed greater decline in severe aggressive incidents (adjusted risk reduction = 41%) but they had higher initial rates of aggression	**	Decrease
Troquete et al., 2013 (NL)	632 adult forensic outpatients (M/F)	START + joint plan (CMs)	Cluster RCT	No significant difference in violence/criminal incidents in wards using tool than in control wards	**	No change
van de Sande et al., 2011 (NL)	597 adult psyc. inpatients (M/F)	BVC, etc. (nurses/drs.)	Cluster RCT	Wards using tool showed greater decline in aggressive incidents than controls (relative risk ratio = -68%)	**	Decrease
B. Non-Randomized Studies						
Belfrage et al., 2004 (SE)	47 adult inmates (M)	HCR-20 (N.R.)	Pre-post	64% decrease in violent incidents following implementation of tool, but risk scores did not decrease	**	Decrease
Bhui et al., 2001 (UK)	324 psyc. inpatients (age, gender N.R.)	CANVAS† (clinicians)	Pre-post	No significant reduction in violent incidents (after taking into account cohort effect)	*	No change
Daffern et al., 2009 (UK)	51 adult forensic inpatients (M)	DASA & HCR-20 (nurses)	Pre-post	No significant reduction in aggression after tool implemented	**	No change
Gunenc et al., 2017 (UK)	50 adult psyc. inpatients (M)	START (clinicians)	Pre-post	No significant reduction in physical or verbal aggression	**	No change
Guy et al., 2015 (MS, USA)	208 youth on probation (M/F)	SAVRY (service counselors)	Comparison (matched)	No significant differences in nonviolent, status, or probation violation offenses, but fewer violent offenses	***	Mixed
Kling et al., 2011 (CA)	807 adult psyc. patients (M/F)	Alert System (nurses)	Pre-post	Violence initially decreased in hospital overall but then increased (results were non-significant)	**	No change
Laura & John Arnold Foundation, 2014 (USA)	Defendants from 120 countries (M/F, n & age N.R.)	PSA† (judges)	Pre-post	After tool implemented, state had a 15% reduction in pretrial crime but did not control for cohort effect, report on statistical significance, or provide methodology	*	Decrease
Needham et al., 2004 (CH)	576 adult psyc. inpatients (M/F)	BVC-R (nurses)	Pre-post	No significant reduction in aggressive incidents or attacks against persons, but positive results on one ward	**	No change (overall)
Vincent et al., 2016 (USA)	1,694 youth on probation (M/F)	SAVRY & YLS/CMI (YPOs)	Pre-post (matched)	New petitions did not change at 5 of 6 sites; new petitions decreased at one site; new adjudications increased at one site	***	No change (for most sites)

Note. † No support for the predictive validity of the tool could be found. The summary ratings were: Decrease (violence or offending decreased in all or almost all analyses), No change (violence or offending did not significantly change in all or almost all analyses), Mixed (no clear overall pattern of results). The study appraisal ratings were: * (high limitations/risk of bias on RoB tool or study appraisal form for non-randomized studies), ** (medium limitations/risk of bias), *** (low limitations/risk of bias). AU = Australia; CA = Canada; CH = Switzerland; CM = case manager; drs. = doctors; M = male; M/F = male/female; MS = Mississippi, NL = Netherlands; N.R. = not reported; psyc. = psychiatric; SE = Sweden; UK = United Kingdom; USA = United States of America; YPO = youth probation officer.

Table 6

What Strategies Help to Improve the Utility of Risk Assessment Tools for Risk Management?

Authors, Year (country)	Sample (gender)	Risk Tool (assessor)	Design	Outcome(s)	Results	Study Appraisal	Summary of Results
Training (i.e., Tool Alone vs. Tool Plus Training)							
Bonta et al., 2011 (CA)	143 adults on probation (M/F)	LSI-R (POs)	RCT	Need principle, offending	POs in STICS program were more likely than controls to discuss needs & attitudes, but less likely to discuss education/employment; clients were less likely to reoffend	***	Improvement
Needham et al. (2004)	576 psyc. inpatients	BVC-R (nurses)	Pre-post	Violence	No significant reduction in aggressive incidents or attacks in the tool + training condition vs. the tool alone condition	**	No change
Storey et al., 2011 (CA)	73 police, etc.	SARA, SAM, HCR-20 (police, etc.)	Pre-post, vignettes	Risk/need principles	After training, match between needs and planned management strategies increased; no change in number of strategies matched to risk level	***	Some improvement
Risk Management Guidelines (i.e., Tool Alone vs. Tool Plus Guidelines)							
Bosker et al., 2015 (NL)	29 POs	RISc2, RISc3 (POs)	Pre-post, vignettes	Risk/need principles	When tool + guidelines, agreement about planned intensity of supervision increased but was still poor; agreement about needs to target increased	**	Some improvement
Bosker & Witteman, 2016 (NL)	579 adult probationers (M/F)	RISc2, RISc3 (POs)	Pre-post	Risk/need principles	When tool + guidelines, plans were more likely to be appropriate intensity, but mean match to needs did not improve	**	Some improvement
Daffern et al., 2009 (AU)	51 adult forensic inpatients (M)	DASA, HCR-20 (nurses)	Pre-post	Violence	No change in rates of aggression for tool + guidelines vs. tool alone	**	No change
Quality of Implementation							
Vincent, Guy, et al., 2012 (USA)	410 young offenders, (M/F)	SAVRY (YPOs)	Pre-post, files	Risk principle	After full implementation (i.e., policies, protocols) vs. standard SAVRY training, match to risk principle increased	***	Improvement
Vincent et al., 2016 (USA)	1,694 youth on probation (M/F)	SAVRY, YLS/CMI (YPOs)	Pre-post, files	Risk principle	Sites with better implementation (i.e., completed risk assessments) had better match to risk principle	***	Improvement

Note. The summary ratings were: Improvement (analyses indicated that the strategy was associated with improvement); Some Improvement (some but not all analyses indicated that the strategy was associated with improvement); No change (all or almost all of analyses indicated that the strategy was not associated with improvement). The study appraisal ratings were: * (high limitations [e.g., <50% on the study appraisal form]), ** (medium limitations [e.g., 51 – 79% on the study appraisal form]), *** (low limitations [e.g., >80% on the study appraisal form]). AU = Australia; CA = Canada; M = male; M/F = male/female; NL = Netherlands; N.R. = not reported; POs = parole or probation officer; STICS = Strategic Training in

Community Supervision; USA = United States of America; YPO = youth probation officer.

Table 7
Overall Summary of Findings

	Study Appraisal						Summary of Findings					
	Low Limitations		Medium Limitations		High Limitations		Low		Mixed		High	
	<i>k</i>	%	<i>k</i>	%	<i>k</i>	%	<i>k</i>	%	<i>k</i>	%	<i>k</i>	%
1. Perceived Utility (<i>k</i> = 12)	5	41.7	5	41.7	2	16.7	2	16.7	7	58.3	3	25.0
2. Use for Risk Management (<i>k</i> = 14)	6	42.9	5	35.7	3	21.4	4	28.6	8	57.1	2	14.3
3. Match to Risk Principle (<i>k</i> = 36)							No Change		Some Improvement		Improvement	
a. Comparison group (<i>k</i> = 2)	1	50.0	1	50.0	–	–	–	–	1	50.0	1	50.0
b. No comparison group (<i>k</i> = 34)	4	11.8	23	67.6	7	20.6	7	20.6	22	64.7	5	14.7
4. Match to Need Principle (<i>k</i> = 17)												
a. Comparison group (<i>k</i> = 0)	–	–	–	–	–	–	–	–	–	–	–	–
b. No comparison group (<i>k</i> = 17)	1	5.9	11	64.7	29.4	5	7	41.2	9	52.9	1	5.9
5. Violence/Reoffending (<i>k</i> = 12)							No Change		Mixed		Decrease	
a. RCTs (<i>k</i> = 3)	–	–	3	100	–	–	1	33.3	–	–	2	66.7
b. Non-Randomized (<i>k</i> = 9)	2	22.5	5	55.6	2	22.0	6	66.7	1	11.1	2	22.2
6. Strategies to Improve Utility (<i>k</i> = 8)							No Change		Some Improvement		Improvement	
a. Training (<i>k</i> = 3)	1	33.3	2	66.7	–	–	1	33.3	2	66.7	–	–
b. Guidelines (<i>k</i> = 3)	–	–	3	100	–	–	1	33.3	2	66.7	–	–
c. Implementation Quality (<i>k</i> = 2)	2	100	–	–	–	–	–	–	–	–	2	100

Note. If the results were mixed, low-mixed, or high-mixed, they were coded as mixed. If the results were moderate, low-moderate, or high-moderate, they were coded as moderate.

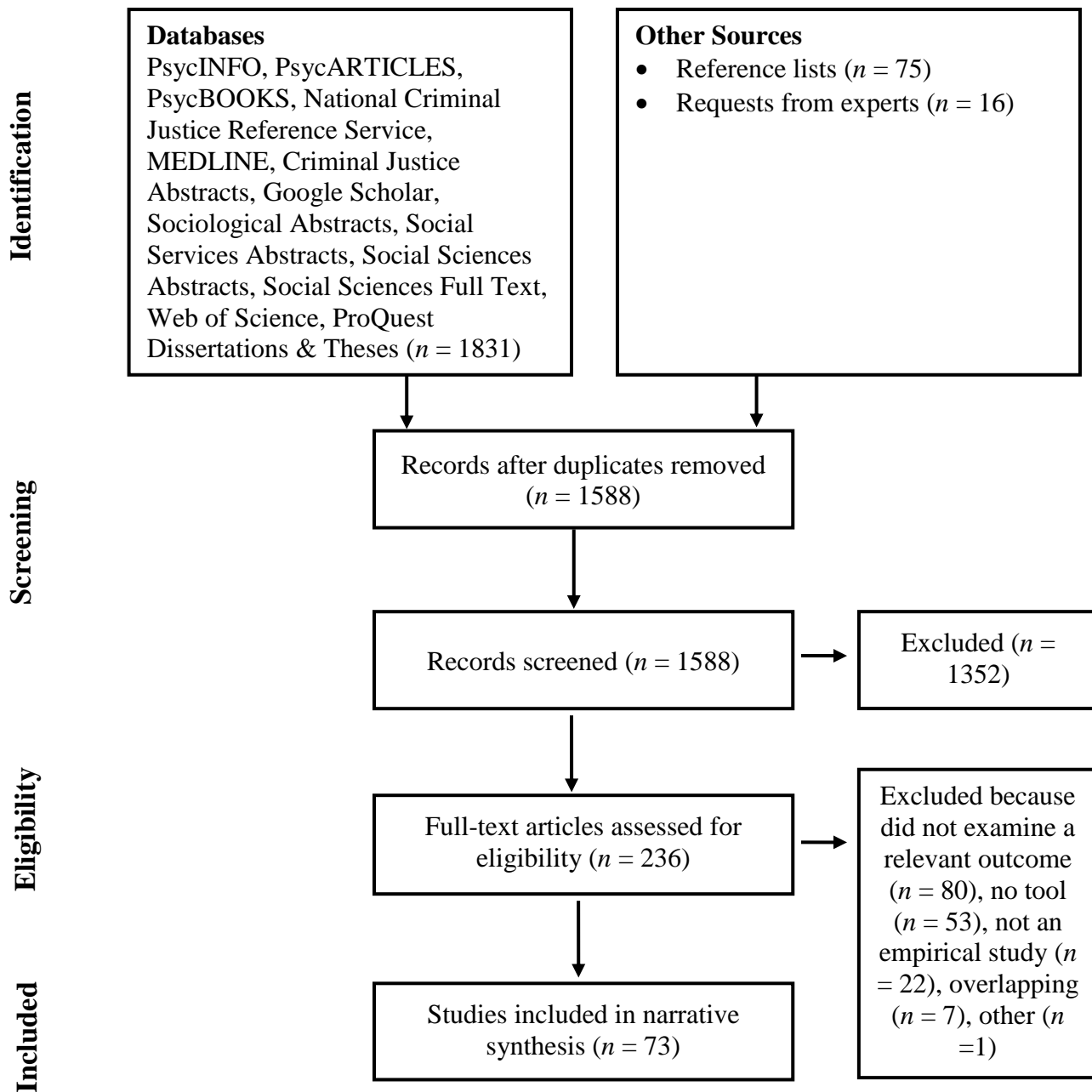


Figure 1. Search strategy and phases of review.

Appendix A

List of Included Studies

1. Abderhalden, C., Needham, I., Dassen, T., Halfens, R., Haug, H., & Fischer, J. E. (2008). Structured risk assessment and violence in acute psychiatric wards: Randomised controlled trial. *The British Journal of Psychiatry*, 193(1), 44-50. doi:10.1192/bjp.bp.107.045534
2. Auditor General of Ontario. (2012). *Annual report: Section 3.13. Youth Justice Services Program*. Retrieved from http://www.auditor.on.ca/en/reports_2012_en.htm
3. Belfrage, H., Fransson, G., & Strand, S. (2004). Management of violent behaviour in the correctional system using qualified risk assessments. *Legal and Criminological Psychology*, 9(1), 11-22. doi:10.1348/135532504322776825
4. Belfrage, H., & Strand, S. (2012). Measuring the outcome of structured spousal violence risk assessments using the B-SAFER: Risk in relation to recidivism and intervention. *Behavioral Sciences and the Law*, 30(4), 420-430. doi:10.1002/bsl.2019
5. Belfrage, H., Strand, S., Storey, J. E., Gibas, A. L., Kropp, P. R., & Hart, S. D. (2012). Assessment and management of risk for intimate partner violence by police officers using the spousal assault risk assessment guide. *Law and Human Behavior*, 36(1), 60-67. doi:10.1037/h0093948
6. Bhui, K., Outhwaite, J., Adzinku, F., Dixon, P., McGabhann, L., Pereira, S., et al. (2001). Implementing clinical practice guidelines on the management of imminent violence on two acute psychiatric in-patient units. *Journal of Mental Health*, 10(5), 559-569. doi:10.1080/09638230120041317
7. Blais, J., & Forth, A. E. (2014). Prosecution-retained versus court-appointed experts: Comparing and contrasting risk assessment reports in preventative detention hearings. *Law and Human Behavior*, 38(6), 531-543. doi:http://dx.doi.org/10.1037/lhb0000082
8. Boccaccini, M. T., Turner, D. B., Murrie, D. C., Henderson, C. E., & Chevalier, C. (2013). Do scores from risk measures matter to jurors? *Psychology, Public Policy, and Law*, 19(2), 259-269. doi:10.1037/a0031354
9. Bonta, J., Bourgon, G., Jesseman, R., & Yessine, A. K. (2005). *Presentence reports in Canada*. (User Report 2005-03). Ottawa, Canada: Public Canada. Retrieved from: <http://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/prsntnc-rprts-cnd/index-en.aspx>
10. Bonta, J., Bourgon, G., Rugge, T., Scott, T., Yessine, A. K., Gutierrez, L., et al. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior*, 38(11), 1127-1148. doi:10.1177/0093854811420678
11. Bonta, J., & Motiuk, L. L. (1990). Classification to half-way houses: Quasi-experimental evaluation. *Criminology*, 28(3), 497-506. doi:10.1111/j.1745-9125.1990.tb01336.x
12. Bonta, J., Rugge, T., Scott, T., Bourgon, G., & Yessine, A. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation*, 47, 248-270. doi:10.1080/10509670802134085
13. Bosker, J., & Witteman, C. (2016). Finding the right focus: Improving the link between risk/needs assessment and case management in probation. *Psychology, Public Policy, and Law*, 22(2), 221-233. doi:10.1037/law0000075
14. Bosker, J., Witteman, C., & Hermanns, J. (2013a). Do intervention plans meet criteria for effective practice to reduce recidivism? How probation officers forget about social capital

- and basic needs. *European Journal of Probation*, 5, 65-85.
doi:10.1177/206622031300500105
15. Bosker, J., Witteman, C., & Hermans, J. (2013b). Agreement about intervention plans by probation officers. *Criminal Justice and Behavior*, 40(5), 569-581.
doi:10.1177/0093854812464220
 16. Bosker, J., Witteman, C., Hermans, J., & Heij, D. (2015). Improving agreement about intervention plans in probation by decision support. *International Journal of Offender Therapy and Comparative Criminology*, 59(13), 1459-1473.
doi:10.1177/0306624X14539127
 17. Campbell, N. A. (2015). Disproportionate minority contact in the juvenile justice system: An investigation of racial disparity in program referral at disposition. (M.A., Michigan State University). *ProQuest Dissertations and Theses*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/1749306194?accountid=13800>
 18. Daffern, M., Howells, K., Hamilton, L., Mannion, A., Howard, R., & Lilly, M. (2009). The impact of structured risk assessments followed by management recommendations on aggression in patients with personality disorder. *Journal of Forensic Psychiatry and Psychology*, 20(5), 661-679. doi:10.1080/14789940903173990
 19. DeGue, S., Scalora, M., Ullman, D., & Gallavan, D. (2008). In-home or out-of-home?: Predicting long-term placement recommendations for juvenile offenders. *International Journal of Forensic Mental Health*, 7(1), 15-27. doi:10.1080/14999013.2008.9914400
 20. Dyck, H. (2016). *An empirical evaluation of the RCMP's youth intervention and diversion strategy in New Brunswick*. Ph.D. Dissertation in Psychology, University of New Brunswick.
 21. Flores, A. W., Travis, L. F., & Latessa, E. J. (2004). *Case classification for juvenile corrections: An assessment of the Youth Level of Service/Case Management Inventory (YLS/CMI): Final report*. Washington, DC: National Institute of Justice.
 22. Flores, A. W. (2013). *Examining the Youth Level of Service/Case Management Inventory in the context of reliability, validity, equity, and utility: A six-year evaluation*. Unpublished doctoral dissertation, University of Cincinnati.
 23. Folino, J. O., Marengo, C. M., Marchiano, S. E., & Ascazibar, M. (2004). The Risk Assessment Program and the Court of Penal Execution in the Province of Buenos Aires, Argentina. *International Journal of Offender Therapy and Comparative Criminology*, 48(1), 49-58. doi:10.1177/0306624X03257526
 24. Gebo, E. (2002). Juvenile justice reform: A tale of two systems. (Ph.D., University of New Hampshire). *ProQuest Dissertations and Theses*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/305500435?accountid=13800>
 25. Gunenc, C., O'Shea, L. E., & Dickens, G. L. (2017). Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: Use of the START. *Criminal Behaviour and Mental Health*. doi:10.1002/cbm.2036
 26. Guy, L. S., Kusaj, C., Packer, I. K., & Douglas, K. S. (2015). Influence of the HCR-20, LS/CMI, and PCL-R on decisions about parole suitability among lifers. *Law and Human Behavior*, 39(3), 232-243. doi:10.1037/lhb0000111
 27. Guy, L. S., Nelson, R. J., Fusco-Morin, S., & Vincent, G. M. (2014). What do juvenile probation officers think of using the SAVRY and YLS/CMI for case management, and do they use the instruments properly? *International Journal of Forensic Mental Health*, 13(3), 227-241. doi:10.1080/14999013.2014.939789

28. Guy, L. S., Vincent, G. M., Grisso, T., & Perrault, R (2015). *Advancing use of risk assessment in juvenile probation*. Technical Report submitted to the Office of Juvenile Justice and Delinquency Prevention. University of Massachusetts Medical School, Department of Psychiatry.
29. Haas, S. M., & DeTardo-Bora, K. A. (2009, Spring). Inmate reentry and the utility of the LSI-R in case planning. *Corrections Compendium, 11-16*, 49-52.
30. Haas, S. M., Hamilton, C. A., & Hanley, D. (2007). *Preparing prisoners for returning home: A process evaluation of West Virginia's offender reentry initiative*. Charleston, WV: Mountain State Criminal Justice Research Services.
31. Hawley, C. J., Gale, T. M., Sivakumaran, T., & Littlechild, B. (2010). Risk assessment in mental health: Staff attitudes and an estimate of time cost. *Journal of Mental Health, 19*, 88–98. doi: 10.3109/09638230802523005
32. Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior, 25*(4), 393-408. doi:10.1023/A:1010607719239
33. Holloway, E. D. (2015). Risk assessment and case planning by juvenile probation officers: The role of trauma. (M.A., Fordham University). *ProQuest Dissertations and Theses*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/1733682593?accountid=13800>
34. Hsieh, M. (2016). Rational discretionary risk: The judicial risk x sentence model for sex offenders. (Ph.D., Washington State University). *ProQuest Dissertations and Theses*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/1805053291?accountid=13800>
35. Jung, S., Brown, K., Ennis, L., & Ledi, D. (2015). The association between presentence risk evaluations and sentencing outcome. *Applied Psychology in Criminal Justice, 11*(2), 111-125. Retrieved from <http://proxy.lib.sfu.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2016-07720-004&site=ehost-live>
36. Kewley, S., Beech, A., Harkins, L., & Bonsall, H. (2015). Effective risk management planning for those convicted of sexual offending. *Journal of Aggression, Conflict and Peace Research, 7*(4), 237-257. doi:10.1108/JACPR-05-2015-0171
37. Kling, R. N., Yassi, A., Smailes, E., Lovato, C. Y., & Koehoorn, M. (2011). Evaluation of a violence risk assessment system (the alert system) for reducing violence in an acute hospital: A before and after study. *International Journal of Nursing Studies, 48*(5), 534-539. doi:10.1016/j.ijnurstu.2010.10.006
38. Krysik, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12*(1), 71-81. doi:10.1177/104973150201200106
39. Laura and John Arnold Foundation (2014). *Results from the first six months of the public safety assessment – Court™ in Kentucky*. Laura and John Arnold Foundation.
40. Luong, D., & Wormith, J. S. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior, 38*(12), 1177-1199. doi:10.1177/0093854811421596
41. Mair, G., Burke, L., & Taylor, S. (2006). 'The worst tax form you've ever seen'? Probation officers' views about OASys. *Probation Journal, 53*(1), 7-23. doi:10.1177/0264550506060861

42. Maupin, J. R. (1993). Risk classification systems and the provision of juvenile aftercare. *Crime and Delinquency*, 39, 90–107. doi:10.1177/0011128793039001006
43. McCormick, S., Peterson-Badali, M., & Skilling, T. (2017). Mental health and specific responsivity: Its role in juvenile justice rehabilitation. *Law & Human Behavior*, 41(1), 55-67. <http://dx.doi.org/10.1037/lhb0000228>
44. McKee, S. A., Harris, G. T., & Rice, M. E. (2007). Improving forensic tribunal decisions: The role of the clinician. *Behavioral Sciences and the Law*, 25(4), 485-506. doi:10.1002/bsl.768
45. Miller, J., & Maloney, C. (2013). Practitioner compliance with risk/needs assessment tools: A theoretical and empirical assessment. *Criminal Justice and Behavior*, 40(7), 716-736. doi:10.1177/0093854812468883
46. Morrissey, C., Beeley, C., & Milton, J. (2014). Longitudinal HCR-20 scores in a high-secure psychiatric hospital. *Criminal Behaviour and Mental Health*, 24(3), 169-180. doi:10.1002/cbm.1893
47. Muheizen, J. (2014). The effect of risk assessment data and colleagues' consensus on clinical decisions to discharge insanity acquittees. ProQuest Information & Learning. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 74 (9). Retrieved from <http://proxy.lib.sfu.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2014-99051-114&site=ehost-live>
48. Needham, I., Abderhalden, C., Meer, R., Dassen, T., Haug, H. J., Halfens, R. G., & Fischer, J. E. (2004). The effectiveness of two interventions in the management of patient violence in acute mental inpatient settings: Report on a pilot study. *Journal of Psychiatric and Mental Health Nursing*, 11(5), 595-601. doi:10.1111/j.1365-2850.2004.00767.x
49. Perrault, R. T., Paiva-Salisbury, M., & Vincent, G. M. (2012). Probation officers' perceptions of youths' risk of reoffending and use of risk assessment in case management. *Behavioral Sciences and the Law*, 30(4), 487-505. doi:10.1002/bsl.2015
50. Peterson-Badali, M., Skilling, T., & Haqanee, Z. (2015). Examining implementation of risk assessment in case management for youth in the justice system. *Criminal Justice and Behavior*, 42(3), 304-320. doi:10.1177/0093854814549595
51. Richardson, C. L. (2009). Structured professional judgement of risk in forensic psychiatric practice. (Ph.D., The University of Nottingham (United Kingdom)). *PQDT - UK & Ireland*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/899736093?accountid=13800>
52. Royal College of Psychiatrists (2008). *Rethinking risk to others in mental health services: Final report of a scoping group* (CR150). Royal College of Psychiatrists.
53. Schneider, A. L., Ervin, L., & Snyder-Joy, Z. (1996). Further exploration of the flight from discretion: The role of risk/need instruments in probation supervision decisions. *Journal of Criminal Justice*, 24(2), 109-121. doi:10.1016/0047-2352(95)00059-3
54. Sen, P., Lindsey, S., Chatterjee, N., Rama-Iyer, R., & Picchioni, M. (2015). An audit of the quality of HCR-20 violence risk assessments in a low secure service. *Journal of Psychiatric Intensive Care*, 11, 1-9. doi:10.1017/S1742646415000096
55. Shook, J. J., & Sarri, R. C. (2007). Structured decision making in juvenile justice: Judges' and probation officers' perceptions and use. *Children and Youth Services Review*, 29(10), 1335-1351. doi:10.1016/j.childyouth.2007.05.008
56. Simpson, T. P. (2010). Do objective measures reduce the disproportionate rates of minority youth placed in detention: Validation of a risk assessment instrument? (Ph.D., University of

- New Orleans). *ProQuest Dissertations and Theses*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/733025086?accountid=13800>
57. Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., et al. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health, 13*(3), 193-206. doi:10.1080/14999013.2014.922141
 58. Singh, J. P., Desmarais, S. L., Sellers, B. G., Hylton, T., Tirotti, M., & Van Dorn, R. A. (2014). From risk assessment to risk management: Matching interventions to adolescent offenders' strengths and vulnerabilities. *Children and Youth Services Review, 47*, 1-9. doi:10.1016/j.chidyouth.2013.09.015
 59. Storey, J. E., & Strand, S. (2013). Assessing violence risk among female IPV perpetrators: An examination of the B-SAFER. *Journal of Aggression Maltreatment & Trauma, 22*(9), 964-980. doi:10.1080/10926771.2013.835015
 60. Storey, J. E., Gibas, A. L., Reeves, K. A., & Hart, S. D. (2011). Evaluation of a violence risk (threat) assessment training program for police and other criminal justice professionals. *Criminal Justice and Behavior, 38*(6), 554-564. doi:10.1177/0093854811403123
 61. Storey, J. E., Kropp, P. R., Hart, S. D., Belfrage, H., & Strand, S. (2014). Assessment and management of risk for intimate partner violence by police officers using the brief spousal assault form for the evaluation of risk. *Criminal Justice and Behavior, 41*(2), 256-271. doi:10.1177/0093854813503960
 62. Storey, J. E., & Strand, S. (2017). The influence of victim vulnerability and gender on police officers' assessment of intimate partner violence risk. *Journal of Family Violence, 32*(1), 125-134. doi:10.1007/s10896-016-9905-0
 63. Troquete, N. A. C., van, d. B., Beintema, H., Mulder, T., van Os, T. W. D. P., Schoevers, R. A., et al. (2013). Risk assessment and shared care planning in out-patient forensic psychiatry: Cluster randomised controlled trial. *British Journal of Psychiatry, 202*(5), 365-371. doi:10.1192/bjp.bp.112.113043
 64. Trujillo, M. P., & Ross, S. (2008). Police response to domestic violence: Making decisions about risk and risk management. *Journal of Interpersonal Violence, 23*(4), 454-473. doi:10.1177/0886260507312943
 65. Turner, D. B., Boccaccini, M. T., Murrie, D. C., & Harris, P. B. (2015). Jurors report that risk measure scores matter in sexually violent predator trials, but that other factors matter more. *Behavioral Sciences and the Law, 33*(1), 56-73. doi:10.1002/bsl.2154
 66. Urquhart, T. A., & Viljoen, J. L. (2014). The use of the SAVRY and YLS/CMI in adolescent court proceedings: A case law review. *International Journal of Forensic Mental Health, 13*(1), 47-61. doi:10.1080/14999013.2014.885470
 67. van de Sande, R., Nijman, H. I., Noorthoorn, E. O., Wierdsma, A. I., Hellendoorn, E., van der Staak, C., & Mulder, C. L. (2011). Aggression and seclusion on acute psychiatric wards: Effect of short-term risk assessment. *British Journal of Psychiatry, 199*(6), 473-478. doi:10.1192/bjp.bp.111.095141
 68. Vieira, T. A., Skilling, T. A., & Peterson-Badali, M. (2009). Matching court-ordered services with treatment needs: Predicting treatment success with young offenders. *Criminal Justice and Behavior, 36*(4), 385-401. doi:10.1177/0093854808331249
 69. Viglione, J., Rudes, D. S., & Taxman, F. S. (2015). Misalignment in supervision: Implementing risk/needs assessment instruments in probation. *Criminal Justice and Behavior, 42*(3), 263-285. doi:10.1177/0093854814548447

70. Vincent, G. M., Guy, L. S., Gershenson, B. G., & McCabe, P. (2012). Does risk assessment make a difference? Results of implementing the SAVRY in juvenile probation. *Behavioral Sciences and the Law*, 30(4), 384-405. doi:10.1002/bsl.2014
71. Vincent, G. M., Guy, L. S., Perrault, R. T., & Gershenson, B. (2016). Risk assessment matters, but only when implemented well: A multisite study in juvenile probation. *Law and Human Behavior*, 40(6), 683-696. doi:10.1037/lhb0000214
72. Vincent, G. M., Paiva-Salisbury, M. L., Cook, N. E., Guy, L. S., & Perrault, R. T. (2012). Impact of risk/needs assessment on juvenile probation officers' decision making: Importance of implementation. *Psychology, Public Policy, and Law*, 18(4), 549-576. doi:10.1037/a0027186
73. Young, D., Moline, K., Farrell, J., & Bierie, D. (2006). Best implementation practices: Disseminating new assessment technologies in a juvenile justice agency. *Crime and Delinquency*, 52(1), 135-158. doi:10.1177/0011128705281752

Appendix B

Description of Included Tools

Tool	Risk Prediction	Population	Focus on Modifiable or Dynamic Items (i.e., Needs)	Support for Predictive Validity?	Description of Predictive Validity
1. AJADMS (Arizona Juvenile Aftercare Decision-Making System; Ashford et al., 1987)	General recidivism	Adolescents	No	No	No information could be identified
2. Alert System (authors N.R.)	Inpatient violence	Adults	Yes	Yes	Sensitivity = .71, specificity = .94 for violent incidents (Kling et al., 2006)
3. B-SAFER (Brief Spousal Assault Form for the Evaluation of Risk; Kropp et al., 2005, 2010)	Intimate partner violence	Adults	Yes	Yes	AUC = .70 and .65 for total scores & summary risk estimates respectively (Storey, Kropp, Hart, Belfrage, & Strand, 2014)
4. BVC (Brøset Violence Checklist; Woods & Almvik, 2002)	Short-term inpatient violence	Adults	Yes	Yes	AUC = .77 for any inpatient aggression (Chu, Daffern, & Ogloff, 2013)
5. CANVAS (Clinical Assessment of Need, Violence Appraisal System; Guite et al., 1998)	Short-term inpatient violence	Not clear	Yes	No	No information could be identified
6. DASA (Dynamic Appraisal of Situational Aggression; Ogloff & Daffern 2006)	Short-term inpatient violence	Adults	Yes	Yes	AUC = .74 for all aggressive incidents (Vojt, Marshall, & Thomson, 2010)
7. DSI (Rapides Parish Juvenile Detention Screening Instrument; authors N.R.)	General recidivism	Adolescents	No	No	DSI did not significantly predict offending (Simpson, 2010)
8. FOTRES (Forensisch Operationalisiertes Therapie- und Risiko-Evaluations-System; Urbaniok, 2007)	General recidivism	Adults	Yes	Yes	AUC = .81 and .76 for total score and summary risk rating respectively for repeat offending (Rossegger et al., 2011)
9. HCR-20 (Historical, Clinical,	Violence	Adults	Yes	Yes	AUC = .71 for violent recidivism

	Risk Management-20; Webster et al., 1997)					(Yang, Wong, & Coid, 2010 – meta-analysis)
10.	HKT-30 (Historical Clinical Future-30; Taskforce Risk Assessment Forensic Psychiatry, 2002)	Violence	Adults	Yes	Yes	AUC = .77 for violent or criminal behavior (van den Brink, van Os, Savenike, & Wiersma, 2010)
11.	Intake Screen (Intake Screen for Risk and Needs; authors N.R.)	General recidivism	Adolescents	Yes	No	No information could be identified
12.	JAG (Juvenile Assessment Generic; authors N.R.)	General recidivism	Adolescents	Yes	Yes	Total risk scores significantly predicted all offending outcomes (Guy et al., 2015)
13.	L17A (Family Violence Risk Assessment and Management Report; authors N.R.)	Domestic violence	Adults	Yes	No	No information could be identified
14.	LSI-R (Level of Service Inventory-Revised; Andrews & Bonta, 1995)	General recidivism	Adults	Yes	Yes	Mean weighted $r = .30$ for general recidivism (Olver, Stockdale, & Wormith, 2014 – meta-analysis)
15.	MnSOST-R (Minnesota Sex Offender Screening Tool-Revised; Epperson et al., 1998)	Sexual reoffense risk	Adult sex offenders	No (mainly static)	Yes	Mean weighted $d = .76$ for sexual recidivism (Hanson & Morton-Bourgon, 2009 – meta-analysis)
16.	NCCD Instrument (National Council on Crime and Delinquency; Wagner et al., 1994)	General recidivism	Adolescents	No	Yes (some)	High-risk youth were more likely to reoffend than low-risk youth, but clinical judgment outperformed tool (Krysiak & LeCroy, 2002)
17.	OASys (Offender Assessment System; Home Office, 2002)	General recidivism & violence	Adults	Yes	Yes	AUC = .79 and .74 for general and violence prediction, respectively (Howard, 2015)
18.	PRA (Primary Risk Assessment; Bonta et al., 2008)	General recidivism	Adults/Adolescents	Yes	Yes	$r_s = .22-.32$ (adults) and non-significant-.50 (youth) for 2-year convictions (Bonta et al., 1994)
19.	PSA (Public Safety Assessment; Laura & John Arnold Foundation, 2014)	General recidivism & violence	Adults	No	No	No information could be identified

20. RISC (Recidivism Risk Assessment Scale, Version 2 and 3; Bosker et al., 2013)	General recidivism	Adults	Yes	Yes	AUC = .70 for general recidivism (van der Knaap & Alberda, 2009)
21. RM2000 (Risk Matrix, 2000; Thornton et al., 2003)	Sexual & violent reoffense risk	Adult sex offenders	No	Yes	AUC = .70 for violent recidivism (Yang et al., 2010 – meta-analysis)
22. RAI (Risk Assessment Instrument; authors N.R.)	Not specified	Adolescents	No	No	No information could be identified
23. SAM (Stalking Assessment and Management; Kropp et al., 2008)	Stalking & violence	Adults	Yes	Yes	Total scores predicted stalking recidivism (hazard ratio = 1.11), but not violence (Foellmi et al., 2016)
24. SARA (Spousal Assault Risk Assessment; Kropp et al., 1994)	Intimate partner violence	Adults	Yes	Yes	Mean weighted <i>d</i> (total score) = .43 for spousal assault recidivism (Hanson et al., 2007 – meta-analysis)
25. SAPROF (Structured Assessment of Protective Factors; de Vogel et al., 2007)	Protective factors for violence	Adults	Yes	Yes	AUC = .78 for actual/threatened violence (Persson, Belfrage, Fredriksson, & Kristiansson, 2017)
26. SAVRY (Structured Assessment of Violence Risk in Youth; Borum et al., 2006)	Violent recidivism	Adolescents	Yes	Yes	Weighted <i>r</i> = .32 and .30 for general and violent recidivism (Olver, Stockdale, & Wormith, 2009)
27. SCJ:Risk (Structured Clinical Judgement: Risk Assessment Scheme; Hogue & Allen, 2006)	Inpatient violence to self and others	Adults	Yes	Yes	AUC = .68 and .66 for any and violent recidivism (Guy, 2008 – meta-analysis)
28. START (Short-Term Assessment of Risk and Treatability; Webster et al., 2009)	Violence, self-harm, etc.	Adults	Yes	Yes	Weighted AUC = .74 to .76 for predicting physical, verbal, or any aggression using the violence risk estimate (O'Shea & Dickens, 2014)
29. START:AV (Short-Term Assessment of Risk and Treatability: Adolescent Version; Nicholls et al., 2010)	Violence, victimization, etc.	Adolescents	Yes	Yes	AUC = .70 and .74 for physical and verbal aggression using risk total (Sher et al., 2014)
30. Static-99 (Hanson & Thornton, 1999)	Sexual & violent reconviction	Adult male sex offenders	No	Yes	AUC = .70 for sexual recidivism (Helmus et al., 2012 – meta-analysis)
31. VRAG (Violence Risk Appraisal Guide; Quinsey et	Violent recidivism	Adults	No	Yes	AUC = .68 for violent recidivism (Yang et al., 2010 – meta-analysis)

al., 2006)						
32. VRS (Violence Risk Scale; Wong & Gordon, 2009)	Violent recidivism	Adults	Yes	Yes	AUC = .65 for violent recidivism (Yang et al., 2010 – meta-analysis)	
33. WRNA (Wisconsin Risk/Needs Assessment Tools; Baird et al., 1979)	General recidivism	Adults	Yes	Yes	Median AUC = .67 and median r = .19 for any recidivism (Desmarais et al., 2016 – systematic review)	
34. YLS/CMI (Youth Level of Service/Case Management Inventory; Hoge & Andrews, 2002), Y-LSI (Youth Level of Service Inventory; Andrews et al., 1984), Youth LSI-SK (LSI Saskatchewan Youth Edition; Andrews et al., 2001)	General recidivism	Adolescents	Yes	Yes	Mean weighted r = .28 and .23 for general and violent recidivism (Olver et al., 2014 – meta-analysis)	

Note. AUC = area under the receiver operating curve, r = correlation, d = Cohen's d . In describing predictive validity, we first searched for meta-analyses. If meta-analyses were not available, we examined studies by independent researchers who were not authors of the tool. If independent studies were not available, we described studies by the tools' authors. Although Singh et al. (2014) surveyed professionals about a variety of different tools, we only referenced the tools described in the text or tables (i.e., FORTES, HKT-30, HCR-20, SAPROF, START, VRS, and VRAG). When tools had multiple versions, we grouped our description together. Interpretative guidelines for predictive validity of tools (Rice & Harris, 2005): d = < .50 (small), .50-.80 (medium), >.80 (high); AUCs: < .64 (small), .64-.71 (medium), > .71 (high); r (assuming a 50% base rate) = < .24 (small), .24-.37 (medium), > .37 (high).

Appendix C

Reference List for Tools

Note. This reference list includes all of the references and studies that are cited in Appendix B.

- Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory-Revised*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2001). *Level of Service Inventory – Saskatchewan Youth Edition*. Toronto, ON: Multi-Health Systems.
- Andrews, D. A., Robinson, D., & Hoge, R. D. (1984). *Manual for the Youth Level of Service Inventory*. Ottawa, Canada: Carleton University, Department of Psychology.
- Ashford, J. B., LeCroy, C. W., & Bond-Maupin, L. (1987). *The Arizona Juvenile Aftercare Decision Making System*. Arizona Department of Corrections.
- Baird, C. S., Heinz, R. C., & Bemus, B. J. (1979). *The Wisconsin Case Classification/Staff Deployment Project: A two-year follow-up report*. Madison, WI: Wisconsin Department of Corrections.
- Bonta, J., Rugge, T., Scott, T., Bourgon, G., & Yessine, A. K. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation, 47*(3), 248-270. doi:10.1080/10509670802134085
- Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY)*. Odessa, FL: Psychological Assessment Resources.
- Bosker, J., Witteman, C., & Hermanns, J. (2013). Do intervention plans meet criteria for effective practice to reduce recidivism? How probation officers forget about social capital and basic needs. *European Journal of Probation, 5*(1), 65-85. doi:10.1177/206622031300500105
- Chu, C. M., Daffern, M., & Ogloff, J. R. P. (2013). Predicting aggression in acute inpatient psychiatric setting using BVC, DASA, and HCR-20 Clinical scale. *Journal of Forensic Psychiatry & Psychology, 24*(2), 269-285. doi:10.1080/14789949.2013.773456
- Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in U.S. correctional settings. *Psychological Services, 13*(3), 206-222. doi:10.1037/ser0000075
- de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2007). *SAPROF Manual. Structured Assessment of Protective Factors for Violence Risk. Version 1*. Utrecht, The Netherlands: Forum Educatief.
- Epperson, D.L., Kaul, J.D., Huot, S.J., Hesselton, D., Alexander, W., & Goldman, R. (1998). *Minnesota Sex Offender Screening Tool-Revised (MnSOST-R)*. St. Paul, Minnesota: Department of Corrections.
- Foellmi, M. C., Rosenfeld, B., & Galietta, M. (2016). Assessing risk for recidivism in individuals convicted of stalking offenses: Predictive validity of the guidelines for stalking assessment and management. *Criminal Justice and Behavior, 43*(5), 600-616. doi:10.1177/0093854815610612
- Guite, H., Thornicroft, G., Kumar, S., Wright, S. & Nwosu, C. (1998). *The development and evaluation of a system to appraise the risk of violence in general adult mental health clients. The CANVAS (Clinical Assessment of Need: Violence Appraisal System)*. London: Institute of Psychiatry.

- Guy, L. S. (2008). Performance indicators of the structured professional judgment approach for assessing risk for violence to others: A meta-analytic survey. *Dissertation Abstracts International*, 71, 2671.
- Guy, L. S., Vincent, G. M., Grisso, T., & Perrault, R. (2015). *Advancing use of risk assessment in juvenile probation*. Technical Report submitted to the Office of Juvenile Justice and Delinquency Prevention. Worcester, MA: University of Massachusetts Medical School.
- Hanson, R. K., Helmus, L. & Bourgon, G. (2007). *The validity of risk assessments for intimate partner violence: A meta-analysis*. (User Report No. 2007-07). Ottawa, Canada: Public Safety Canada.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, 21(1), 1-21. doi:10.1037/a0014421
- Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders* (User Rep. No. 1999-02). Ottawa, Canada: Department of the Solicitor General of Canada.
- Helmus, L., Hanson, R. K., Thornton, D., Babchishin, K. M., & Harris, A. R. (2012). Absolute recidivism rates predicted by Static-99R and Static-2002R sex offender risk assessment tools vary across samples: A meta-analysis. *Criminal Justice and Behavior*, 39(9), 1148-1171. doi:10.1177/0093854812443648
- Hoge, R. D., & Andrews, D. A. (2002). *The Youth Level of Service/Case Management Inventory*. Toronto, Canada: Multi-Health Systems.
- Hogue, T. E., & Allen, C. (2006). *Structured Clinical Judgement: Risk. Multidisciplinary assessment of risk. A system for structuring clinical judgements of risk related to high security hospital needs*. Nottinghamshire, United Kingdom: Nottinghamshire Healthcare NHS Trust.
- Home Office (2002) *Offender Assessment System: OASys manual, Version 2*. London: Home Office.
- Howard, P. (2015). The prediction of reoffending by age, gender and ethnicity. In R. Moore (Ed.), *A compendium of research and analysis on the Offender Assessment System (OASys)* (pp. 45-74). National Offender Management Service, Ministry of Justice Analytical Series.
- Kling, R., Corbière, M., Milord, R., Morrison, J. G., Craib, K., Yassi, A., & ... Saunders, S. (2006). Use of a violence risk assessment tool in an acute care hospital: Effectiveness in identifying violent patients. *AAOHN Journal: Official Journal of the American Association of Occupational Health Nurses*, 54(11), 481-487. doi:10.1177/216507990605401102
- Kropp P. R., Hart S. D., & Belfrage H. (2005). *Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER): User manual*. Vancouver, Canada: ProActive ReSolutions.
- Kropp P. R., Hart S. D., & Belfrage H. (2010). *Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER), Version 2: User manual*. Vancouver, Canada: ProActive ReSolutions.
- Kropp, P. R., Hart, S. D., & Lyon, D. R. (2008). *Stalking Assessment and Management (SAM)*. Vancouver, British Columbia, Canada: ProActive ReSolutions Inc.
- Kropp, P. R., Hart, S. D., Webster, C. W., & Eaves, D. (1994). *Manual for the Spousal Assault Risk Assessment Guide*. Vancouver, Canada: British Columbia Institute on Family Violence.

- Krysiak, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12*(1), 71-81. doi:10.1177/104973150201200106
- Nicholls, T. L., Viljoen, J. L., Cruise, K. R., Desmarais, S. L., & Webster, C. D. (2010). *Short-Term Assessment of Risk and Treatability: Adolescent Version (START: AV) (Abbreviated Manual)*. Coquitlam, Canada: BC Mental Health and Addiction Services.
- Ogloff, J. R. P., & Daffern, M. (2006). The dynamic appraisal of situational aggression: an instrument to assess risk for imminent aggression in psychiatric inpatients. *Behavioral Science and Law, 24*(6), 799-813. doi:10.1002/bsl.741
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior, 36*(4), 329-353. doi:10.1177/0093854809331457
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment, 26*(1), 156-176. doi:10.1037/a0035080
- O'Shea, L. E., & Dickens, G. L. (2014). Short-Term Assessment of Risk and Treatability (START): Systematic review and meta-analysis. *Psychological Assessment, 26*(3), 990-1002. doi:10.1037/a0036794
- Persson, M., Belfrage, H., Fredriksson, B., & Kristiansson, M. (2017). Violence during imprisonment, forensic psychiatric care, and probation: Correlations and predictive validity of the risk assessment instruments COVR, LSI-R, HCR-20V3, and SAPROF. *International Journal of Forensic Mental Health, 16*(2), 117-129. doi:10.1080/14999013.2016.1266420
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.
- Rossegger, A., Laubacher, A., Moskvitin, K., Villmar, T., Palermo, G. B., & Endrass, J. (2011). Risk assessment instruments in repeat offending: The usefulness of FOTRES. *International Journal of Offender Therapy and Comparative Criminology, 55*(5), 716-731. doi:10.1177/0306624X09360662
- Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29*(5), 615-620. doi:10.1007/s10979-005-6832-7
- Sher, M. A., & Gralton, E. (2014). Implementation of the START:AV in a secure adolescent service. *Journal of Forensic Practice, 16*(3), 184-193. doi:10.1108/JFP-04-2013-0021
- Simpson, T. P. (2010). Do objective measures reduce the disproportionate rates of minority youth placed in detention: Validation of a risk assessment instrument? (Ph.D., University of New Orleans). *ProQuest Dissertations and Theses*. Retrieved from <http://search.proquest.com.proxy.lib.sfu.ca/docview/733025086?accountid=13800>
- Storey, J. E., Kropp, P. R., Hart, S. D., Belfrage, H., & Strand, S. (2014). Assessment and management of risk for intimate partner violence by police officers using the brief spousal assault form for the evaluation of risk. *Criminal Justice and Behavior, 41*(2), 256-271. doi:10.1177/0093854813503960
- Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing and combining risks for sexual and violent recidivism. *Annals of the New York Academy of Sciences, 989*(1), 225-235. doi:10.1111/j.1749-6632.2003.tb07308.x

- Urbaniok, F. (2007). FOTRES: *Forensisches Operationalisiertes Therapie-Risiko-Evaluations-System* [FOTRES: Forensic Operationalized Therapy/Risk Evaluation-System]. Bern, Switzerland: Zytglogge.
- van den Brink, R. H. S., Hooijschuur, A., van Os, T. W. D. P., Savenije, W., & Wiersma, D. (2010). Routine violence risk assessment in community forensic mental healthcare. *Behavioral Sciences and the Law*, 28(3), 396-410. doi:10.1002/bsl.904
- van der Knaap, L. M., & Alberda, D. L. (2009). *De predictieve validiteit van de Recidive Inschattingsschalen (RISc)* [Predictive validity of the Recidivism Risk Assessment Scales (RISc)]. Den Haag, the Netherlands: Ministerie van Justitie, WODC.
- Vojt, G., Marshall, L. A., & Thomson, L. D. G. (2010). The assessment of imminent inpatient aggression: A validation study of the DASA-IV in Scotland. *Journal of Forensic Psychiatry and Psychology*, 21(5), 789-800. doi:10.1080/14789949.2010.489952
- Wagner, D., DeComo, R. E., & Wiebush, R. (1994). *Development of a risk/needs instrument for first referrals to the Arizona Juvenile Courts: Structuring decisions about prevention services*. San Francisco, CA: National Council on Crime and Delinquency.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence, Version 2*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. L. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START)* (Version 1.1). Coquitlam, Canada: Forensic Psychiatric Services Commission and St. Joseph's Healthcare.
- Werkgroep Risicotaxatie Forensische Psychiatrie [Taskforce Risk Assessment Forensic Psychiatry]. (2002). *Handleiding HKT-30 versie 2. Risicotaxatie in de forensische psychiatrie*. Den Haag, The Netherlands: Ministerie van Justitie, Dienst Justitiële Inrichtingen.
- Wong, S., & Gordon, A. (2009). *Manual for the Violence Risk Scale*. Saskatoon, Canada: University of Saskatchewan.
- Woods, P., & Almvik R. (2002). The Brøset Violence Checklist (BVC). *Acta Psychiatrica Scandinavica*, 106, 103-5. doi:10.1034/j.1600-0447.106.s412.22.x
- Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136(5), 740-767. doi:10.1037/a0020473