

Appraising the risk of sexual and violent recidivism among intellectually disabled offenders

Joseph A. Camilleri^{a*} and Vernon L. Quinsey^b

^aDepartment of Psychology, Westfield College, Westfield, USA; ^bDepartment of Psychology, Queen's University, Kingston, Canada

The relatively high prevalence and recidivism rates of offenders with intellectual disabilities suggest research on appraising their risk is an important priority. Although research has found good predictive accuracy of available risk assessments with intellectually disabled (ID) offenders, we recommend several ways to improve on them: understanding the theoretical link between intellectual disability and offending may help to identify new risk items; avoiding assessments that require clinical judgment in risk appraisal; developing risk assessments using best practices; and accumulating studies with larger samples from all intellectual disability categories for the purposes of meta-analytic research. To demonstrate an approach to reaching the latter goal, we present new analyses that show the Violence Risk Appraisal Guide (VRAG) has good predictive accuracy with psychiatric patients of lower intelligence.

Keywords: risk assessment; intellectual disability

Introduction

Interest in using structured and actuarial assessments to assess the risk of intellectually disabled (ID) offenders has increased over the past 5 years. Much of the recent work has been addressed in Lindsay, Taylor, and Sturmey's (2004) edited book on *Offenders with Developmental Disabilities* and in a special issue on risk assessment in the December 2004 issue of the *Journal of Applied Research in Intellectual Disabilities*. These and other reviews describe how aspects of the general risk assessment literature, such as the distinction between static and dynamic predictors, the utility of clinical judgment, and the predictive accuracy of risk assessments, could be applied to offenders with intellectual disabilities (Anderson, 2005; Harris & Tough, 2004; Quinsey, Book, & Skilling, 2004; Turner, 2000). The general consensus is that research on risk assessments with ID offenders has been scarce, suggesting more work on applying already developed risk assessment instruments to ID offenders and proper development of new ones are required (Lindsay & Beail, 2004; Mohr & Gray, 2005; Turner, 2005). We extend these reviews by examining the latest research on using risk assessments with ID offenders and by presenting new data to demonstrate how to validate risk assessments with ID offenders in the interim while more thorough research programs are developed.

*Corresponding author. Email: jcamilleri@wsc.ma.edu

Defining intellectual disability

Intellectual disability is a term used to describe individuals who show marked cognitive impairments resulting from atypical neurodevelopment. An intellectual disability is ordinarily signified by an IQ score lower than 70 and evidence of deficits in adaptive social behaviour before the age of 18 (for clinical diagnostic criteria see the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2000); *International Statistical Classification of Diseases and Related Health Problems* (World Health Organization, 1994)). IQ scores are assessed from standardized measures, such as the Wechsler Adult Intelligence Scale (Wechsler, 1981), whereas adaptive behaviours are deficits in behavioural standards for one's age and cultural group. Explicit descriptions on how intellectual disability was assessed (e.g. IQ alone, or full clinical assessment) should be included in reports of risk assessment studies in order to ascertain sample equivalence when comparing studies or collapsing across studies in meta-analyses, and allows us to know exactly in which populations the risk assessments have been validated.

Intellectual disability and crime: understanding risk factors

Understanding the causal relationship between intellectual disability and crime provides insight into which items should be included and excluded from specialized risk assessments with ID offenders. Antisocial characteristics of offenders with an intellectual disability can be categorized into those that are likely a result of the intellectual disability and those that are not. Antisocial characteristics found among men with intellectual disability that are either not or unlikely to be a direct result of the intellectual disability include antisocial attitudes, familial criminality, personality disorder, psychopathy, mental health problems, prior offences, and being a young man (Keeling, Beech, & Rose, 2007; Lindsay et al., 2006; Turner, 2005). These factors are found among all offender populations and explain why many actuarial assessments designed with heterogeneous samples of offenders work just as well with exclusive samples of ID offenders (see 'Risk assessments' section).

Predictors of antisocial behaviour that may directly or indirectly result from intellectual disability¹ include various cognitive deficits, certain childhood behaviour problems, unemployment, lower socioeconomic status, deviant sexual preferences, low self-esteem, unassertiveness, and poor response to treatment (Holland, Clare, & Mukhopadhyay, 2002; Keeling et al., 2007; Lindsay, Elliot, & Astell, 2004; Turner, 2005). Although these factors correlate with crime in most offender groups (Ellis & Walsh, 2000), they are more pronounced in ID offenders, perhaps accounting for their higher recidivism rates. With further research, these variables may also be useful as dynamic risk factors in risk assessments for ID offenders.

To address specific criminogenic needs of ID offenders, risk assessments should therefore include unique characteristics of ID offenders that lead to offending and maintaining criminal behaviour (see 'Future research' section). Examples of these characteristics include slower information processing, learning difficulties, concrete thinking, language and communication problems, and sexual naiveté (Craig & Hutchinson, 2005; Day, 1993; Keeling et al., 2007). Because these characteristics are not established criminogenic needs (i.e. potentially modifiable aspects of the individual which are related to criminality), more research comparing ID offenders

with ID non-offenders is required to determine their relationship to criminal behaviour (Lindsay, 2002). Where no differences are found, these characteristics are useless in assessing risk, and provide no novel avenues for treatment designed to lower risk. If differences are found, the next step would be to assess whether treatment programs can address these needs, and whether their modification results in lower recidivism rates.

Much has been theorized about the causes of sexual offending by men with intellectual disabilities. Lindsay (2004) placed these theories into five categories: counterfeit deviance, mental illness, sexual abuse, sexual deviance, and impulsivity. Only counterfeit deviance accounts for characteristics of intellectual deficits that may give rise to antisocial behaviour. Initially, counterfeit deviance was used to account for ID individuals who behaved in a sexually deviant manner but tested normally in phallometric assessments (Hingsburger, Griffiths, & Quinsey, 1991). Their offending appears not to be due to antisociality or anomalous sexual preferences, but a byproduct of sexual naivety, underdeveloped social and heterosexual skills, and limited sexual opportunities (see Lindsay, 2004; Lunsy, Frijters, Griffiths, Watson, & Willinston, 2007). These putative causes of counterfeit deviance are criminogenic needs that could be used in specialized risk assessments for ID offenders.

Psychotic symptoms, such as hallucinations or delusions, are often thought to aggravate antisocial characteristics. In this view, having a psychiatric illness does not cause crime *per se*, but increases the rate of offending. Recidivism studies, however, have found the opposite to be true among offender samples. For example, although mental illness is a modest risk factor for violence in community samples, Harris, Rice, and Quinsey (1993) found that a diagnosis of schizophrenia was actually a protective factor for violent recidivism among offenders. Theorists have also suggested that ID sexual offenders were sexually victimized as children, particularly in institutions. Although most child victims of sexual assault are female and very unlikely to become sexual offenders themselves as adults, there are some data supporting the sexual abuse hypothesis for male offenders (e.g. Beail & Warden, 1995; Jespersen, Lalumière, & Seto, 2009). These examples highlight the importance of not only identifying the causes of offending, but testing whether they do, in fact, predict recidivism.

Verbal and physical aggression are more prevalent types of challenging behaviour among ID persons. Allen (2000) suggested that these causes of aggression in the general population are the causes of aggression in ID groups, but having an intellectual disability introduces more risk factors, such as the ones described by Gardner and Moffat (1990). These factors include neural and physiological correlates of crime and negative interactions with staff, for example. If true, then currently available risk assessments should work just as well with ID offenders (see 'Risk assessments').

Selectionist or Darwinian approaches to the aetiology of violent and sexual offending (Lalumière, Harris, Quinsey, & Rice, 2005; Quinsey, 2002; Quinsey & Lalumière, 1995) suggest that the characteristics of ID offenders, such as lower intelligence, health, attractiveness, education, and specialized skills result in competitive disadvantage in the mating market. If, for example, sexual offending results from such competitive disadvantage, one would expect it primarily to involve sexual coercion directed toward young adult women. Persistent sexual behaviours directed

toward pre-pubescent children or men, on the other hand, more likely indicate a malfunctioning of the male sexual preference mechanism. This latter hypothesis is consistent with finding lower IQ among sexual offenders who exhibit sexual preferences for children than men whose sexual preferences are for adults (Cantor et al., 2004) and the finding that ID offenders have more deviant sexual age and gender preferences than non-handicapped offenders (Blanchard et al., 1999). If intellectual disability is more likely to result in child molestation than sexual assault of adult women, accuracy of prediction may therefore be dependent on the outcome measure. That is, risk assessments developed for ID offenders may be more accurate when predicting sexual offences against children.²

Psychopathy, a major correlate of coercive sexual behaviour, is an unlikely candidate explanation of ID offending. A Darwinian hypothesis of psychopathy suggests it is a nonpathological but obligate (i.e. not conditional on environment) alternative life-history strategy that involves precocious and coercive sexual behaviour directed toward adult females, in addition to other antisocial acts. Psychopaths are characterized by having early and frequent sexual activity (Harris, Rice, Hilton, Lalumière, & Quinsey, 2007) while having fewer neurodevelopmental perturbations than non-psychopathic offenders (Lalumière, Harris, & Rice, 2001), suggesting that most ID offenders are unlikely to be psychopathic. Thus, the predictive accuracy of psychopathy among ID offenders still needs to be established (see 'Risk assessments').

Measuring risk

Historically, assessing a person's 'dangerousness' was used primarily among forensic psychiatric patients, then later to guide judicial decisions such as releasing offenders into the community or determining the applicability of various 'dangerous offender' statutes (Quinsey, Harris, Rice, & Cormier, 2006). In addition to guiding sentencing and probation decisions, risk assessments are used nowadays to address the principles of effective correctional programming, such as risk (treatments should target those at higher risk) and need (specific criminogenic needs requiring attention from treatment; Andrews et al., 1990). How risk is determined has changed over the past 40 years (Doyle & Dolan, 2002). Initially, psychiatrists and psychologists used clinical expertise in working with these groups to determine their dangerousness. Many studies, however, have shown that clinical judgments are no better than lay judgments (e.g. Quinsey & Ambtman, 1979) and that actuarial instruments – the second generation of risk assessment tools – had consistently higher accuracy (Ægisdóttir et al., 2006; Hilton, Harris, & Rice, 2006). Ideally, actuarial instruments include multiple risk factors (i.e. variables that correlate with recidivism) that together provide a more accurate prediction of recidivism than any single risk factor. Reasons for the superiority of actuarial assessments include errors in human judgment such as base rate neglect, and human inability to weight the relative importance of multiple risk factors (Quinsey, Harris et al., 2006). The latest generation, structured clinical judgment, has sought to reconcile the advantages of clinical and actuarial assessments (Doyle & Dolan, 2002). So far, however, structured clinical appraisals have not been observed to be more accurate than actuarial assessments (for a review, see Quinsey, Harris et al., 2006).

Sophisticated statistics have been adopted in this literature to ensure appropriate comparisons between different risk assessments and different populations. One of the more popular and useful statistics is the *receiver operating characteristic* (ROC). The ROC is used to generate a value that represents the area under the curve (AUC) from plotting the trade-off between a measure's hit rate by false alarm rate at various risk cut-offs (for more details see Quinsey, Harris et al., 2006). Values range from 0.00 to 1.00, where higher values represent more accurate predictions. AUCs have been useful because they are robust to fluctuations in base rates and are intuitive in their interpretation. A score of 0.5 means the assessment predicts at chance, whereas a score of 1.00 represents perfect prediction and errorless outcome data. Good actuarial risk assessments typically fall around 0.75 (Seto, 2005).

Although research shows that most actuarial risk assessments are applicable with a wide variety of specific populations, such as psychiatric patients and women (Harris, Rice, & Camilleri, 2004), researchers have created specialized actuarial instruments for different offender groups or for the prediction of particular outcomes. These specialized assessments incorporate unique risk items and some are tailored for use in particular settings. A good example of this type of assessment is the Ontario Domestic Assault Risk Assessment scale (ODARA), which includes unique items such as being a stepfather to one of the victim's children and having committed a domestic assault against a pregnant victim.³ The ODARA excludes items that are difficult to assess (e.g. PCL-R, childhood variables), making it easier for front-line workers to use. This approach to developing actuarial assessments with ID offenders has not been adopted (Harris & Tough, 2004; Lindsay, 2004), even though risk assessments of ID offenders are increasingly common (Turner, 2000, 2005). Some clinicians use their own risk assessments without properly establishing their predictive validity. The use of unvalidated measures also leads to inconsistent definitions of high risk men (Lindsay & Beail, 2004). In the following section we review papers that have either reviewed the use of clinical judgment or provided evidence for using scale-based assessment with ID offenders.

Clinical judgment

Meta-analytic research has shown that, even though clinical assessments of risk perform better than chance, they are consistently outperformed by actuarial assessments (Bonta, Law, & Hanson, 1998; Grove & Meehl, 1996; Mossman, 1994). There has only been one study that directly tested the predictive accuracy of clinical judgment relative to actuarial assessment among ID offenders. McMillan, Hastings, and Coldwell (2004) reported an AUC of 0.74 for the clinical prediction of recidivism over a 6-month period based on a single nine-point rating scale, whereas the AUC for the number of times each client was violent in the last 6 months (their one-item 'actuarial tool') was 0.77 (the predictive accuracy of risk assessments with ID populations is summarized in Table 1). It was not reported whether the assessors were blind to the actuarial tool results.

Other more elaborate procedures have been developed to guide clinical judgment of risk with ID offenders. For example, Anderson (2005) described the Risk Assessment Management and Audit System (RAMAS), which requires multiple assessments from multiple sources. Anderson has recommended the RAMAS for use in various settings with ID offenders, including community, institutional, and

Table 1. Predictive accuracy of risk assessments with ID offenders.

Study	Risk assessment	Group description (<i>N</i>)	Recidivism type	Follow-up years	AUC	<i>r</i>
Camilleri and Quinsey	VRAG-modified	Psychiatric patients 9th percentile ^b (51)	Violence	0.4 (fixed)	0.70***	–
		2nd percentile to 9th percentile ^b (40)			0.68*	–
Gray et al. (2007)		ID offenders	Violence	2–5		
	VRAG	(84)			0.73***	–
	PCL-SV	(100)			0.73***	–
	HCR-20	(107)			0.79***	–
	History	(106)			0.81***	–
	Clinical	(106)			0.71*	–
	Risk management	(106)			0.64	–
Lindsay et al. (2008)	VRAG	ID offenders in institutional and community settings (212)	Violence	1 (fixed)	0.71***	–
	HCR-20 (total)		Violence		0.72***	–
	Historical				0.68***	–
	Clinical				0.67***	–
	Risk management				0.62*	–
	SDRS		Violence		0.72***	–
	EPS-E		Violence		0.75***	–
	EPS-I		Violence		0.73***	–
	RM2000V		Violence		0.62*	–
	Static-99		Sexual		0.71***	–
McMillan, Hastings, and Coldwell (2004)	Clinical rating scale	ID offenders (124)	Institutional violence	0.5 (fixed)	0.74 ^a	–
	Past behaviour				0.77 ^a	–
Quinsey, Book, and Skilling (2004)	VRAG-modified	ID offenders (58)	Violence and sexual	1.25 (mean)	0.69*	–

Table 1 (Continued)

Study	Risk assessment	Group description (<i>N</i>)	Recidivism type	Follow-up years	AUC	<i>r</i>
Quinsey, Harris, Rice, and Cormier (2006)	VRAG	ID offenders (128)	Violence	10 (mean)	0.79 ^a	–
Tough (2001)	RRASOR Static-99	ID sexual offenders (76)	Sexual	0.2–19 (range)	– –	0.31** 0.08

Note. VRAG, Violence Risk Appraisal Guide; PCL-SV, Psychopathy Checklist Screening Version; HCR-20, Historical, Clinical, and Risk Management 20; SDRS, Short Dynamic Risk Scale; EPS-E, Emotional Problem Scale: Externalising; EPS-I, Emotional Problem Scale: Internalising; RM2000V, Risk Matrix 2000: Violence; RRASOR, Rapid Risk Assessment for Sexual Offender Recidivism.

^a*p*-value not reported. ^bPercentiles on WAIS-R vocabulary scores.

p* < 0.05. **p* < 0.001.

research settings, although there is at present little evidence for its predictive validity. Frameworks that guide clinical judgment by including more objective scores have also been developed. For example, the Structured Anchored Clinical Judgment Risk Assessment Protocol (SACJ) includes, as part of a three-step process, items such as offence characteristics, criminal history, marital status, and deviant sexual arousal, in addition to judgments about treatment progress (Grubin, 1998). The only peer-reviewed study on the SACJ, however, used a heterogeneous group of non-ID sexual offenders and excluded the only component that requires clinical judgment, referring the revised scale as SACJ-min. The SACJ-min had inconsistent predictive accuracy across four institutions and did not significantly predict recidivism in three out of the four institutions. It was outperformed by the Static-99 for both violent and sexual recidivism (Hanson & Thornton, 2000).

Risk assessments

Risk assessments include any scale-based measure that uses risk factors to predict recidivism. Some items are more objective (e.g. criminal history), whereas other items require clinical assessments (e.g. PCL-R). We do not use the term 'actuarial' to describe all scale-based risk assessments because only a few of them were developed from statistical procedures needed to make such a claim. To qualify as an actuarial instrument, a probability of recidivism must be generated from follow-up data. To date, measures that include clinical overrides are not actuarials because recidivism probabilities were not known. Also, those rating scales that use arbitrary cut-offs to identify individuals as high, moderate, and low risk (without specifying a specific probability) do not qualify. In this connection, there is no agreed upon definition of high, medium, or low risk (Hilton, Carter, Harris, & Sharpe, 2008; Hilton et al., 2006; Hilton & Simmons, 2001).

There are a few actuarial instruments with promising results among the ID population. The Violence Risk Appraisal Guide (VRAG) is a 12-item actuarial tool developed to predict the likelihood of violent or sexual reoffending. Although the scale was developed with a heterogeneous sample of psychiatric and non-psychiatric offenders, numerous studies have demonstrated its ability to predict recidivism in a variety of populations (Quinsey, Harris et al., 2006), including ID offenders. In the construction sample, the VRAG had an AUC of 0.79 among those with IQs less than 85 (Quinsey, Harris et al., 2006). Recently, a modified version of the scale was used to predict violence in a small group ($n = 58$) of ID offenders released to the community over the course of an average of 15 months. With this group, the VRAG had an AUC of 0.69 in predicting violent and sexual reoffending in supervised settings (Quinsey et al., 2004). Recent research has confirmed these expectations: Gray, Fitzgerald, Taylor, MacCullough, and Snowden (2007) found that the AUC for the VRAG was 0.73 for violent recidivism over a 5-year follow-up.

Tough (2001) tested the accuracy of two risk assessments with ID sexual offenders, the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997), and the Static-99 (created from combining items from the RRASOR and SACJ-min; Hanson & Thornton, 2000). Interestingly, Tough found that the RRASOR correlated with recidivism ($r = 0.31$, $p = 0.007$) whereas the Static-99 did not ($r = 0.08$, $p = 0.50$). After using modified versions of the scales due to differences in data collection, the results remained the same. Tough suggested that RRASOR

items may generalize to other populations, but to explain this finding correlations between Static-99 items and recidivism need to be observed among ID offenders. For a few years, the RRASOR was considered the best assessment tool for use with ID offenders due to its significant relationship with reoffending (Boer, Tough, & Haaven, 2004; Keeling et al., 2007).

More recently, the VRAG was confirmed as a useful predictor of violent incidents ($AUC = 0.71$, $p < 0.001$) and the Static-99 significantly predicted sexual recidivism ($AUC = 0.71$, $p = 0.08$) with offenders from low to high levels of security and from community settings (Lindsay et al., 2008). In the same study, a newer risk assessment, the Risk Matrix 2000 (RM 2000; Thornton et al., 2003), was not as successful in predicting violent ($AUC = 0.62$, $p = 0.07$) or sexual recidivism ($AUC = 0.61$, $p = 0.08$), although it approached significance.

Although not initially designed as a risk assessment measure, the Psychopathy Checklist: Revised (PCL-R; Hare, 1991) correlated highly with recidivism and has been included in actuarial assessments. Morrissey et al. (2005) studied psychopathy among people with an intellectual disability using the PCL-R. They found PCL-R scores correlated with aggressive incidents over a 6-month period with externalizing behaviour (from the Emotional Problems Scales: Behaviour Rating Scales), and with risk assessment scores, including the VRAG and Historical, Clinical, and Risk Management 20 (HCR-20; a tool with static and potentially dynamic risk items; Webster, Eaves, Douglas, & Wintrup, 1995). More recently, however, Morrissey and colleagues (2007) did not find a relationship between the PCL-R nor its two factors and institutional violence over a 12-month period. A possible explanation is that there is insufficient variance in psychopathy among ID offenders (Boer et al., 2004), and, as discussed earlier, these results are consistent with the hypothesis that psychopathy is a nonpathological obligate strategy.

Dynamic risk assessments

The VRAG, Static-99, RRASOR and PCL-R have been criticized for their lack of clinical utility because they are made up entirely of static items. Static risk factors are historical events or stable characteristics that are correlated with recidivism. Although clinicians can use static risk scales to identify offenders who are worth treating (i.e. those at high risk), static scales do not predict changing imminence of recidivism during a follow-up period (although they do predict how quickly individuals will reoffend in a long follow-up) nor do they identify treatment targets. Assessments that include acute dynamic risk items (characteristics that if altered also alter a person's risk) are capable of determining short term risk.⁴ Although many studies now show statistical relationships between dynamic risk and recidivism, a significant setback has been the lack of evidence showing that a change in the dynamic risk factors results in reductions in recidivism. To date, only one study has tested for this effect (Olver, Wong, Nicholaichuk, & Gordon, 2007).

Quinsey et al. (2004) conducted a field trial of dynamic risk predictors with ID offenders. A variety of staff rating scales were used to see if there was an increase in psychotic behaviours, inappropriate and antisocial behaviours, mood problems, social withdrawal, dynamic antisociality, psychotic symptoms, poor compliance, and medication compliance/dysphoria before new offences. A within-subjects design was used to look for changes in these ratings between the index month (the month in

which the new offence occurred), prior month, and previous months (average scores from 6 months before the prior month). There were increases in inappropriate/antisocial behaviours, mood problems, social withdrawal, and denial of all problems across the three time periods. The authors suggested that weekly ratings would improve accuracy in detecting imminent offences.

Recently, Lindsay et al. (2008) looked at the predictive accuracy of the HCR-20 along with three dynamic risk scales, the Short Dynamic Risk Scale (SDRS; Quinsey, 2004), and the Emotional Problem Scale for both externalizing (EPS-E) and internalizing (EPS-I) behaviour problems, designed specifically for individuals with intellectual disability (Prout & Strohmer, 1991).⁵ The predictive accuracy of the dynamic subscales of the HCR-20 and both dynamic risk scales was comparable to the static risk scales. The HCR-20 C (AUC = 0.67, $p < 0.001$), HCR-20 R (AUC = 0.62, $p = 0.02$), SDRS (AUC = 0.72, $p < 0.001$), EPS-E (AUC = 0.75, $p < 0.001$), and EPS-I (AUC = 0.73, $p < 0.001$) all significantly predicted violent incidents. It remains to be demonstrated whether this result can be obtained with raters blind to the static items.

Special populations

Despite the moderate to good accuracy of actuarial risk assessments with all criminal subgroups studied thus far, there may be different risk factors for different types of offenders (Quinsey, Harris et al., 2006). Some more recent developments, therefore, include the construction of risk scales for particular populations. A good example is the ODARA (Hilton et al., 2004), developed to predict domestic violence. The scale includes items that are unique to domestic offending, such as having assaulted the victim while pregnant or whether offender is a stepfather, and predicted remarkably well (AUC = 0.77). Only one scale, the Dynamic Risk Assessment and Management System (DRAMS; Lindsay, Murphy et al., 2004), was designed specifically for ID offenders. Although its items were drawn from the dynamic risk literature and are comprehensive in terms of important criminogenic needs, the scale was not developed using statistics that identify the most accurate combination of predictors⁶ and items are not specific to ID offenders. Even though the total DRAMS score was found to have an AUC of 0.73 ($p < 0.001$), two of the eight items were not related to recidivism, and four items did not have significant AUCs. This poses a significant problem for targeting criminogenic needs – there is little reason to believe that targeting items that do not predict recidivism would reduce risk, although it is still plausible that change scores relate to recidivism while absolute values do not. Research evaluating the relationship between change scores and recidivism are required to justify the inclusion of these items.

Future research

Having reviewed the literature on appraising the risk of recidivism with ID offenders, we foresee research taking two routes. In the short term, databases could easily be reanalysed to see if a risk assessment predicts with ID offenders. In the long term, risk assessments should be developed with criminogenic needs in mind and more research on treatment effects on recidivism.

Considering the relative lack of risk assessment research with ID offenders, clinicians and researchers need to know which scales can accurately predict reoffending with this group. Although we outlined a few such studies, most had relatively small samples (i.e. less than 100). By accumulating a larger number of risk assessment studies using conventional risk assessments and by including many descriptive statistics, useful meta-analytic research could be conducted. We therefore suggest reporting AUCs for each category of intellectual disability and including those with a borderline diagnosis. These categories are also important because research indicates that most offenders with intellectual disability are higher functioning – whether IQ category also influences accuracy of prediction remains to be tested. It should be noted that individuals with severe intellectual disabilities usually do not pose a risk to the community because they require close supervision and are housed in group home or institutional settings. Managing risk with this group involves primarily contingency management and related behavioural interventions.

The construction of matched control groups, such as ID non-offenders or non-ID offenders, may also be informative. Other categories worth reporting are the source of intellectual disability (developmental disability vs brain injury), type of intellectual disability (e.g. autism vs Downs syndrome), any dual diagnoses, source of recidivism reports (ID offenders are less likely to be charged with crimes, staff reports may be a better source; see Doyle, 2004; Keeling et al., 2007), all of which may have idiosyncratic effects on predictive accuracy.

One way to provide new results in the short term is to reanalyse risk prediction data sets if measures of intellectual disability were also collected. To demonstrate the ease to which databases could be reanalysed to test for predictive accuracy, we used a dataset created from the publicly available MacArthur Violence Risk Assessment Study (macarthur.virginia.edu/risk.html) to test the application of the VRAG to non-forensic psychiatric patients (Harris et al., 2004). WAIS-R vocabulary scores were also collected in the original data file, so we had a means (albeit an imperfect means) of identifying a sample of men and women who are likely ID, or at the very least have linguistic impairments.⁷ We used two cut-offs to ensure both liberal and conservative estimates of intellectual disability. Typically, borderline intellectual disability falls with the 9th percentile of scaled WAIS-R distributions, and intellectual disability within the 2nd percentile (Sattler, 1988). We therefore calculated AUCs for individuals who were within the 9th percentile when including ($n=51$) and excluding ($n=40$) those within the 2nd percentile (ID groups), and for individuals who scored above the 9th ($n=677$) percentile (non-ID group) on WAIS-R vocabulary scores. Thirteen participants were excluded because they did not have WAIS-R vocabulary scores. When using the 9th percentile, the VRAG predicted just as well among ID ($AUC=0.70$, $p=0.015$, $95\% CI \pm 0.14$) as non-ID groups ($AUC=0.70$, $p<0.001$, $95\% CI \pm 0.04$) over a 20 week follow-up. Excluding individuals under the 2nd percentile slightly weakened the predictive accuracy among the ID group ($AUC=0.68$, $p=0.05$, $95\% CI \pm 0.17$), probably due to the smaller sample size. These results not only suggest the VRAG can be used with people of lower intelligence, but highlights the importance of using large samples from all categories of intellectual disability.

In the long term, greater time and monetary investments are required to develop actuarial assessments that are specific to ID offenders. Optimal methods for scale development requires the use of two large samples (one for development and another

for validation), measured by multiple sources on a comprehensive pool of risk factors, with reliability checks on coding and a reasonable amount of follow-up time (at least 4 years). Using these rigorous procedures may result in enhanced predictive accuracy, and identification of criminogenic needs that are specific to ID offenders.⁸ Even if including variables that are causally related to recidivism do not improve overall actuarial accuracy, they may be useful for designing interventions to reduce recidivism. Depending on their purpose, actuarial instruments can focus on predicting different forms of recidivism, include (either inclusively or exclusively) clinically relevant variables, or use variables that are readily available to front line workers.⁹ Normative data for ID offenders with available risk assessments should also be generated. Until each of these developments have occurred, we can not be certain that current assessments are optimal with this population. Having said that, assessments with established psychometric properties and validated with ID offenders, particularly the VRAG and RRASOR, should be used in the interim.

Conclusions

More rigorous research on using sexual and violent recidivism risk assessments with ID offenders is required to keep up with the progress being made with other offender groups. Because risk assessments consist of variables that predict antisocial behaviour, we reviewed theoretical explanations that link intellectual disabilities with criminal behaviour. Using characteristics of intellectual disabilities that predict antisocial behaviour in risk assessments may lead to more accurate predictions, and use of dynamic risk factors may also assist with treatment programming. We used this theoretically guided approach to identify possible constraints in prediction, but such constraints still require empirical tests. Until specialized risk assessments with ID offenders are developed and validated, a few risk assessments can be used to predict recidivism with this group, such as the VRAG and RRASOR. We summarized the most current research, including our reanalysis of the modified VRAG's effectiveness with low verbal-IQ psychiatric patients, in Table 1. By accumulating a large number of studies that vary in their sample characteristics (e.g. type of intellectual disability) and outcome variable (e.g. type of sexual recidivism), researchers can then use meta-analytic techniques to fully understand the conditions that result in accurate prediction of recidivism among ID offenders.

Acknowledgements

The authors would like to thank the MacArthur Foundation and the MacArthur Risk Network members for allowing access to data from the MacArthur Violence Risk Assessment study, and Grant Harris for his feedback and assistance with the MacArthur database. We also acknowledge the support of the Social Sciences and Humanities Research Council of Canada, and Grant Harris, Martin Lalumière, Michael Seto, Jessica Jones, Patricia Minnes, Bill Lindsay, and two anonymous reviewers for their helpful comments.

Notes

1. How these risk factors are related to intellectual disabilities requires further research.
2. It should also be noted that accuracy in predicting sexual reoffending may be less accurate than predicting sexual or violent reoffending because sexual offences are sometimes

- misclassified and severity is underestimated on police 'rapsheets' (Rice, Harris, Lang, & Cormier, 2006).
3. Although these items are unique in the sense that they take place in relationships, they are likely proxies for general antisociality (G. T. Harris, personal communication, 4 September 2007).
 4. Dynamic risk factors can also be stable (Hanson & Harris, 2000) or temporally fixed, and so are treated like static variables (Quinsey, Jones, Book, & Barr, 2006).
 5. The EPS was designed specifically for people with intellectual disabilities, although not as a risk assessment tool.
 6. Hierarchical regression is typically used identify the group of variables that best predict recidivism.
 7. Most sources suggest an estimate of IQ requires at the very least WAIS-R scores from the vocabulary and block design tests (Silverstein, 1982). Still, the vocabulary test has the highest correlation with the full score (House, 1996), and differences between offenders and non-offenders are more pronounced on verbal IQ scores (Herrnstein, 1989).
 8. Researchers are encouraged to exhaust all possible factors unique to ID offenders in addition to established risk factors. Some of these unique characteristics are outlined in the section on 'Intellectual Disability and Crime'. Recommended dynamic risk characteristics include socioaffective function (emotion in social interactions), distorted attitudes, self-management, sexual preferences, noncompliance, and many others (Lindsay & Beail, 2004; Lindsay, Elliot, & Astell, 2004). Researchers then need to demonstrate that modifying these dynamic risk factors leads to reductions in recidivism.
 9. If and when such research programs start, we direct researchers to the procedures used to develop the ODARA, described earlier, because it targets a specific offender group.

References

- Ægisdóttir, S., White, M.J., Spengler, P.M., Maugherman, A.S., Anderson, L.A., Cook, R.S. et al. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*, 341–382.
- Allen, D. (2000). Recent research on physical aggression in persons with intellectual disability: An overview. *Journal of Intellectual & Developmental Disability*, *25*, 41–57.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th-TR ed.). Washington, DC: APA.
- Anderson, G. (2005). People with intellectual disabilities who offend or are at risk of offending: Assessing need and risk. In J. Hogg & A. Langa (Eds.), *Assessing adults with intellectual disabilities: A service provider's guide* (pp. 192–206). Malden: Blackwell.
- Andrews, D.A., Zinger, I., Hoge, R.D., Bonta, J., Gendreau, P., & Cullen, F.T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369–404.
- Beail, N., & Warden, S. (1995). Sexual abuse of adults with learning disabilities. *Journal of Intellectual Disability Research*, *39*, 382–387.
- Blanchard, R., Watson, M.S., Choy, A., Dickey, R., Klassen, P., Kuban, M. et al. (1999). Pedophiles: Mental retardation, maternal age, and sexual orientation. *Archives of Sexual Behavior*, *28*, 111–127.
- Boer, D.P., Tough, S., & Haaven, J. (2004). Assessment of risk manageability of intellectually disabled sex offenders. *Journal of Applied Research in Intellectual Disabilities*, *17*, 275–283.
- Bonta, J., Law, M., & Hanson, K. (1998). The prediction of violent recidivism among mentally disorder offenders: A meta-analysis. *Psychological Bulletin*, *123*, 123–142.
- Cantor, J.M., Blanchard, R., Christenson, B.K., Dickey, R., Klassen, P., Beckstead, A.L. et al. (2004). Intelligence, memory, and handedness in pedophilia. *Neuropsychology*, *18*, 3–14.
- Craig, L., & Hutchinson, R.B. (2005). Sexual offenders with learning disabilities: Risk, recidivism and treatment. *Journal of Sexual Aggression*, *11*, 289–304.
- Day, K. (1993). Crime and mental retardation: A review. In K. Howells, & C. Hollin (Eds.), *Clinical approaches to the mentally disordered offender* (pp. 111–144). Chichester, UK: Wiley.

- Doyle, D.M. (2004). The difference between sex offending and challenging behaviour in people with an intellectual disability. *Journal of Intellectual & Developmental Disability, 29*, 107–118.
- Doyle, M., & Dolan, M. (2002). Violence risk assessment: Combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing, 9*, 649–657.
- Ellis, L., & Walsh, A. (2000). *Criminology: A global perspective*. Toronto: Allyn & Bacon.
- Gardner, W.I., & Moffat, C.W. (1990). Aggressive behavior: Definition, assessment, treatment. *International Review of Psychiatry, 2*, 91–100.
- Gray, N.S., Fitzgerald, S., Taylor, J., MacCullough, M.J., & Snowden, R.J. (2007). Predicting future reconviction in offenders with intellectual disabilities: The predictive efficacy of VRAG, PCL–SV, and the HCR–20. *Psychological Assessment, 19*, 474–479.
- Grove, W.M., & Meehl, P.E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law, 2*, 293–323.
- Grubin, D. (1998). *Sex offending against children: Understanding the risk* (No. Police Research Series, Paper 99). London: Policing and Reducing Crime Unit, Research, Development and Statistics Directorate.
- Hanson, R.K. (1997). *The development of a brief actuarial risk scale for sexual offense recidivism*. Ottawa, Ontario: Department of the Solicitor General of Canada.
- Hanson, R.K., & Harris, A.J.R. (2000). Where should we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior, 27*, 6–35.
- Hanson, R.K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial assessments. *Law and Human Behavior, 24*, 119–136.
- Hare, R.D. (1991). *Manual for the Hare Psychopathy Checklist – Revised*. Toronto, ON: Multi-Health Systems.
- Harris, A.J.R., & Tough, S. (2004). Should actuarial risk assessments be used with sex offenders who are intellectually disabled? *Journal of Applied Research in Intellectual Disabilities, 17*, 235–241.
- Harris, G.T., Rice, M.E., & Quinsey, V.L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*, 315–355.
- Harris, G.T., Rice, M.E., & Camilleri, J.A. (2004). Applying a forensic actuarial assessment (the Violence Risk Appraisal Guide) to nonforensic patients. *Journal of Interpersonal Violence, 19*, 1063–1074.
- Harris, G.T., Rice, M.E., Hilton, N.Z., Lalumière, M.L., & Quinsey, V.L. (2007). Coercive and precocious sexuality as a fundamental aspect of psychopathy. *Journal of Personality Disorders, 21*(1), 1–27.
- Herrnstein, R. (1989). *Biology and crime. National Institute of Justice Crime File* (No. NCJ 97216). Washington, DC: US Department of Justice.
- 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*, 393–408.
- Hilton, N.Z., Harris, G.T., Rice, M.E., Lang, C., Cormier, A.C., & Lines, K.J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The ODARA. *Psychological Assessment, 16*, 267–275.
- Hilton, N.Z., Harris, G.T., & Rice, M.E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *The Counseling Psychologist, 34*, 400–409.
- Hilton, N.Z., Carter, A.M., Harris, G.T., & Sharpe, A.J.B. (2008). Does using nonnumerical terms to describe risk aid violence risk communication? Clinician agreement and decision making. *Journal of Interpersonal Violence, 23*, 171–188.
- Hingsburger, D., Griffiths, D., & Quinsey, V.L. (1991). Detecting counterfeit deviance: Differentiating sexual deviance from sexual inappropriateness. *The Habilitative Mental Healthcare Newsletter, 10*, 51–54.
- Holland, T., Clare, I.C.H., & Mukhopadhyay, T. (2002). Prevalence of ‘criminal offending’ by men and women with intellectual disability and the characteristics of ‘offenders’:

- Implications for research and service development. *Journal of Intellectual Disability Research*, 46, 6–20.
- House, A.E. (1996). The Wechsler Adult Intelligence Scale-Revised (WAIS-R). In C.E. Newmark (Ed.), *Major psychological assessment instruments* (2nd ed., pp. 320–347). Boston, MA: Allyn & Bacon.
- Jespersen, A.F., Lalumière, M.L., & Seto, M.C. (2009). Sexual abuse history among adult sex offenders and non-sex offenders: A meta-analysis. *Child Abuse & Neglect*, 33, 179–192.
- Keeling, J.A., Beech, A.R., & Rose, J.R. (2007). Assessment of intellectually disabled sexual offenders: The current position. *Aggression and Violent Behavior*, 12, 229–241.
- Lalumière, M.L., Harris, G.T., & Rice, M.E. (2001). Psychopathy and developmental instability. *Evolution and Human Behaviour*, 22, 75–92.
- Lalumière, M.L., Harris, G.T., Quinsey, V.L., & Rice, M.E. (2005). *The causes of rape: Understanding individual differences in male propensity for sexual aggression*. Washington, DC: American Psychological Association.
- Lindsay, W.R. (2002). Research and literature on sex offenders with intellectual and developmental disabilities. *Journal of Intellectual Disability Research*, 46, 74–85.
- Lindsay, W.R. (2004). Sex offenders: Conceptualisation of the issues, services, treatment and management. In W.R. Lindsay, J.L. Taylor & P. Sturmey (Eds), *Offenders with developmental disabilities* (pp. 163–185). Hoboken: Wiley.
- Lindsay, W.R., & Beail, N. (2004). Risk assessment: Actuarial prediction and clinical judgement of offending incidents and behaviour for intellectual disability services. *Journal of Applied Research in Intellectual Disabilities*, 17, 229–234.
- Lindsay, W.R., Elliot, S.F., & Astell, A. (2004). Predictors of sexual offence recidivism in offenders with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 17, 299–305.
- Lindsay, W.R., Murphy, L., Smith, G., Murphy, D., Edwards, Z., & Chittock, C. (2004). The Dynamic Risk Assessment and Management System: An assessment of immediate risk of violence for individuals with offending and challenging behaviour. *Journal of Applied Research in Intellectual Disabilities*, 17, 267–274.
- Lindsay, W.R., Taylor, J.L., & Sturmey, P. (2004). *Offenders with developmental disabilities*. Chichester, UK: Wiley.
- Lindsay, W.R., Hogue, T., Taylor, J.L., Mooney, P., Steptoe, L., Johnston, S. et al. (2006). Two studies on the prevalence and validity of personality disorder in three forensic intellectual disability samples. *The Journal of Forensic Psychiatry & Psychology*, 17, 485–506.
- Lindsay, W.R., Hogue, T., Taylor, J.L., Steptoe, L., Mooney, P., O'Brien, G. et al. (2008). Risk assessment in offenders with intellectual disability: A comparison across three levels of security. *International Journal of Offender Therapy and Comparative Criminology*, 52, 90–111.
- McMillan, D., Hastings, R.P., & Coldwell, J. (2004). Clinical and actuarial prediction of physical violence in a forensic intellectual disability hospital: A longitudinal study. *Journal of Applied Research in Intellectual Disabilities*, 17, 255–265.
- Mohr, C., & Gray, K.M. (2005). Assessment in intellectual disability. *Current Opinion in Psychiatry*, 18, 476–483.
- Morrissey, C., Hogue, T., Mooney, P., Lindsay, W.R., Steptoe, L., Taylor, J. et al. (2005). Applicability, reliability and validity of the Psychopathy Checklist – Revised in offenders with intellectual disabilities: Some initial findings. *International Journal of Forensic Mental Health*, 4, 207–220.
- Morrissey, C., Hogue, T., Mooney, P., Allen, C., Johnston, S., Hollin, C. et al. (2007). Predictive validity of the PCL-R in offenders with intellectual disability in a high secure hospital setting: Institutional aggression. *Journal of Forensic Psychiatry & Psychology*, 18, 1–15.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62, 783–792.
- Olver, M., Wong, S.C.P., Nicholaichuk, T.P., & Gordon, A.E. (2007). The validity and reliability of the Violence Risk Scale – Sexual Offender Version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment*, 19, 318–329.

- Prout, T.H., & Strohmer, D.C. (1991). *Emotional problem scale*. Lutz, FL: Psychological Assessment Resources.
- Quinsey, V.L. (2002). Evolutionary theory and criminal behavior. *Legal and Criminological Psychology*, 7, 1–13.
- Quinsey, V.L. (2004). Risk assessment and management in community settings. In W.R. Lindsay, J.L. Taylor, & P. Sturmey (Eds.), *Offenders with developmental disabilities* (pp. 131–142). Chichester, UK: Wiley.
- Quinsey, V.L., & Ambtman, R. (1979). Variables affecting psychiatrists' and teachers' assessments of the dangerousness of mentally ill offenders. *Journal of Consulting and Clinical Psychology*, 47, 353–362.
- Quinsey, V.L., & Lalumière, M.L. (1995). Evolutionary perspectives on sexual offending. *Sexual Abuse: A Journal of Research and Treatment*, 7, 301–315.
- Quinsey, V.L., Book, A.S., & Skilling, T.A. (2004). A follow-up of deinstitutionalized men with intellectual disabilities and histories of antisocial behavior. *Journal of Applied Research in Intellectual Disabilities*, 17, 243–253.
- 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.
- Quinsey, V.L., Jones, G.B., Book, A.S., & Barr, K.N. (2006). The dynamic prediction of antisocial behavior among forensic psychiatric patients: A prospective field study. *Journal of Interpersonal Violence*, 21, 1539–1565.
- Rice, M.E., Harris, G.T., Lang, C., & Cormier, A.C. (2006). Violent sex offenses: How are they best measured from official records? *Law and Human Behavior*, 30, 525–541.
- Sattler, J.M. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.
- Seto, M.C. (2005). Is more better? Combining actuarial risk scales to predict recidivism among adult sex offenders. *Psychological Assessment*, 17, 156–167.
- Silverstein, A.B. (1982). Two- and Four-Subtest Short Forms of the Wechsler Adult Intelligence Scale – Revised. *Journal of Consulting and Clinical Psychology*, 50, 415–418.
- Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C. et al. (2003). Distinguishing and combining risks for sexual and violent recidivism. *Annals of the New York Academy of Sciences*, 989, 225–235.
- Tough, S. (2001). *Validation of two standard risk assessments (RRASOR, 1997; STATIC-99, 1999) on a sample of adult males who are intellectually disabled with significant cognitive deficits*. Unpublished Master's Thesis, University of Toronto, Toronto.
- Turner, S. (2000). Forensic risk assessment in intellectual disabilities: The evidence base and current practice in one English region. *Journal of Applied Research in Intellectual Disabilities*, 13, 239–255.
- Turner, S. (2005). Risk assessment management. In T. Riding, C. Swann, & B. Swann (Eds.), *The handbook of forensic learning disabilities* (pp. 121–135). Oxford, U.K.: Radcliffe.
- Webster, C.D., Eaves, D., Douglas, K.S., & Wintrup, A. (1995). *The HCR-20: the assessment of dangerousness and risk*. Vancouver, Canada: Simon Fraser University and British Columbia Forensic Psychiatric Services Commission.
- Wechsler, D. (1981). *WAIS-R Manual*. New York: Psychological Corporation.
- World Health Organization. (1994). *International Statistical Classification of Diseases and Related Health Problems* (10th ed.). Geneva, Switzerland.