

Northumbria Research Link

Citation: Lindsay, William, Hogue, Todd E., Taylor, John, Steptoe, Lesley, Mooney, Paul, O'Brien, Gregory, Johnston, Susan and Smith, Anne (2008) Risk assessment in offenders with intellectual disability: a comparison across three levels of security. *International Journal of Offender Therapy and Comparative Criminology*, 52 (1). pp. 90-111. ISSN 0306-624X

Published by: SAGE

URL: <http://dx.doi.org/10.1177/0306624X07308111>
<<http://dx.doi.org/10.1177/0306624X07308111>>

This version was downloaded from Northumbria Research Link:
<https://nrl.northumbria.ac.uk/id/eprint/2871/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

Risk Assessment in Offenders with Intellectual Disability:
A Comparison Across Three Levels of Security

William R Lindsay ^{1,2,3}

Todd Hogue ⁴

John L Taylor ⁵

Lesley Steptoe ^{1,2,3}

Paul Mooney ⁴

Gregory O'Brien ⁵

Susan Johnston ⁴

Anne H W Smith ²

International Journal of Offender Therapy and Comparative Criminology, 52, 90-111.

1. The State Hospital, Carstairs
2. NHS Tayside
3. University of Abertay, Dundee
4. Rampton Hospital
5. Northgate & Prudhoe NHS Trust & University of Northumbria

Keywords: Risk assessment, offenders and intellectual disability.

Correspondence to: Professor Bill Lindsay, Clinical Psychology Department, Wedderburn House,
1 Edward Street, Dundee, Scotland. DD1 5NS

Abstract

Background

In mainstream offender samples, several risk assessments have been evaluated for their relative predictive validity. This work has helped to advance the field of risk assessment in general criminological services. The present study extends this work to the field of intellectual disabilities.

Method

Seventy three participants from a high secure setting, 70 participants from medium/low secure and 69 participants from a community setting were compared employing a range of static and dynamic risk assessments. The total cohort was also combined in the evaluation of predictive validity for each assessment.

Results

The VRAG, HCR20 H Scale and the EPS Internalising Scale discriminated between groups with participants from high security having higher mean scores than those in medium security who in turn had higher mean scores than those in a community setting. The VRAG, all HCR20 Scales, the SDRS and both EPS Internalising and Externalising Scales showed significant AUCs in relation to the prediction of violent incidents.

Conclusions

The results are discussed in terms of the value of these various scales to intellectual disability services. The VRAG, HCR20, SDRS and EPS would appear to have some value in the evaluation of risk for future violent incidents. The EPS in particular may also be valuable for the evaluation of the severity of symptomatology.

Risk Assessment in Offenders with Intellectual Disability: A Comparison Across Three Levels of Security

The field of risk assessment has developed considerably in the last 15 years. A number of assessments have been produced for the prediction of both violent and sexual incidents and studies have compared the relative effectiveness of a range of instruments. The Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice & Cromier 1998) and the Sex Offender Risk Appraisal Guide (SORAG; Quinsey et al. 1998) are perhaps the most widely researched actuarial assessments in existence. Of the several studies comparing the predictive accuracy of different risk assessment instruments on a range of data bases, most employ either the VRAG or the SORAG as comparators. For example, Barbaree, Seto, Langton and Peacock (2001) compared the VRAG, SORAG, Rapid Risk Assessment of Sexual Offence Recidivism (RRASOR; Hanson 1997), Static-99 (Hanson & Thornton 1999), The Minnesota Sex Offender Screening Tool – Revised (MnSOST-R; Epperson, Kaul & Hesselton 1998) and the Multi-factorial Assessment of Sex Offender Risk for Recidivism (MASORR). The paper itself gives a detailed description of these various instruments. The instruments were compared on assessments of recidivism for 215 sex offenders in Canada. They had been released from prison for an average of 4.5 years. They found that the VRAG, SORAG, RRASOR and Static-99 successfully predicted general recidivism and sexual recidivism. They found that the Psychopathy Checklist – Revised (PCL-R; Hare 1991), which is incorporated into the VRAG and SORAG, predicted general and serious recidivism but was less successful with sexual recidivism. However the RRASOR, which includes only four easily scored items, was superior to the other instruments in predicting sexual recidivism.

Other authors have compared the predictive accuracy of various instruments across a range of clients and cultures. Kroner and Mills (2001) compared the predictive accuracy of

the PCL-R, the VRAG and the HCR-20 (Webster, Eaves, Douglas & Wintrup 1995) with 97 male offenders in Canada. They found no statistically significant differences between the instruments although the VRAG had the highest prediction correlations for both minor and major misconducts. Sjostedt and Langstrom (2002) compared the predictive accuracy of the Sexual Violence Risk – 20 (SVR-20; Boer et al. 1997), the RRASOR, the PCL-R and the VRAG on prediction of recidivism amongst 51 men convicted of rape and followed up for 92 months in Sweden. Only the RRASOR showed predictive accuracy for sexual recidivism while the other assessments showed some predictive accuracy with violent non-sexual recidivism. Cooke, Michie and Ryan (2002) compared the VRAG, SORAG, Static-99 and HCR-20 with 250 violent offenders in Scotland. They found that each instrument predicted recidivism with similar accuracy to other studies, but that there was no significant differences between the instruments. Finally, Bartosh, Garby, Lewis and Gray (2003) compared the Static-99, RRASOR, MnSOST-R and SORAG in predicting recidivism in 251 sexual offenders in the USA. They categorised their participants in terms of index offence type and found that none of the four tests had consistent predictive validity across categories. However, the Static-99 and SORAG emerged as the most consistent instruments in terms of predictive accuracy.

Clearly, although different assessments are used in different comparative studies, it would appear that the VRAG/SORAG is often included as a comparison, presumably because of its extensive psychometric derivation and its long history. More recently, Harris et al. (2003) compared the VRAG, SORAG, RRASOR and Static-99 in the prediction of recidivism for 396 sexual offenders in Canada. All four instruments predicted recidivism with significantly greater accuracy than chance. Prediction of violent recidivism was consistently higher for the VRAG and SORAG with effect sizes large for violent recidivism and moderate for sexual recidivism.

With mainstream offenders, actuarial prediction has recently advanced in accuracy. Banks et al. (2004) have developed the work of the MacArthur studies (Monahan 2002) in which over 1,000 patients were assessed with a wide range of interview and psychometric techniques. These individuals developed an Iterative Classification Tree (ICT) based on variables which emerged from their study including gender, prior violence, childhood experience, neighbourhood context, psychiatric diagnosis, psychopathy as measured by the PCL-R, delusions, hallucinations, violent thoughts and anger as measured by the Novaco Anger Scale (Novaco & Taylor 2004). Banks et al. (2004) developed 10 different prediction models generated by the ICT each based on different risk factors. They found that each model produced reasonable accuracy in predicting violent incidents with Receiver Operator Characteristics (ROC), Area Under the Curve (AUC) of greater than 0.738. Combining the best five ICT models, they derived a ROC value of 0.878. This would suggest that a highly structured combination of variables conducted in their study produces more accurate risk predictions. Harris and Rice (2004) recently reported predictive accuracies maximising near ROC areas of 0.90. It would appear that risk prediction in mainstream criminogenic research has developed to a considerable degree of sophistication. Unfortunately, work on individuals with intellectual disabilities is at present in its infancy.

The first study applying the VRAG to a population of individuals with intellectual disability (ID) and histories of serious antisocial behaviour has been conducted by Quinsey, Book and Skilling (2004) they followed up 58 clients for an average of 16 months. Sixty seven percent of individuals exhibited antisocial behaviour during the follow-up period and the predictive accuracy of the VRAG for these individuals was a ROC, AUC of 0.69. This is somewhat lower than some other studies but considerably better than chance. The only other variable that significantly predicted antisocial incidents was previous inappropriate and antisocial behaviour. In another study, on a sample of 124 individuals with ID, MacMillan,

Hastings and Caldwell (2004) also found that an individual's history of violence predicted future violence with a ROC value of greater than 0.7. In the development of a structured clinical assessment for men with ID who have committed sex offences or sexual abuse, Boer, Tough and Haaven (2004) recommend the additional use of the Static-99 or RRASOR as an actuarial assessment. They also note that both have been used on samples of adult males with ID in an unpublished thesis (Tough 2001). However, there is clearly a paucity of work in this area for this client group.

Risk assessments are split into those which focus on historical/actuarial factors and those which focus on dynamic/proximal factors (see Lindsay & Beail, 2004 for a description and discussion). The preceding paragraphs have outlined studies on historical/actuarial risk assessments which include factors that, in general, will not change in the person's psychological profile. Dynamic risk factors are psychological variables which are more amenable to change in the short or medium term. In relation to sex offence recidivism, Hanson and Harris (2000) review the comprehensive range of stable and acute dynamic risk factors in 208 sex offence recidivists and 201 non-recidivists. They found the most predictive variables associated with recidivism were poor social supports, antisocial lifestyle, poor self-management strategies, difficulties co-operating with supervision, anger and subjective distress just before the offence and attitudes tolerant of sexual assault. Quinsey, Book and Skilling (2004) also employed the Short Dynamic Risk Scale in their study and found that hostile attitude and non-compliance with current management/therapeutic procedures were significantly associated with future aggressive incidents. Lindsay et al. (2004) also found significant predictive results for the proximal variables of mood (inappropriate anger, anxiety, mania and sadness) and antisocial behaviour (verbal and non-verbal threats and violence). Therefore there is some evidence that certain dynamic variables predict offending incidents with individuals with ID.

Lindsay and Beail (2004) have written about the importance of developing this research for two reasons. With the massive development of community management for individuals with ID, there is a pressing need for risk assessments which can be used by clinical and community teams. This would help planning services from national strategic frameworks to individual case management. Secondly, in the absence of reliable valid instruments, local services tend to develop their own risk assessments with serious drawbacks. Often the items that may be used in these assessments may have no relevant research as to their predictive validity in relation to violent or sexual incidents. Secondly, since local risk assessments have been idiosyncratic by definition, communication between professionals from these different services is compromised. On the basis of their risk assessments, one service may feel it has a high percentage of clients with high risk while another service may feel it has a moderate percentage of individuals with high risk. Since each conclusion may be based on locally developed risk assessments, such comparisons may be meaningless. This clearly has serious repercussions for regional and national service planning.

For these reasons there is a pressing need to develop work on risk assessment in this client group and the current paper is the first of its kind to investigate the performance of a range of risk assessments with this client group. The VRAG, HCR20, Static-99 and RM 2000 (Thornton, 1998) have been employed with 213 men in offender services distributed across three levels of security: high security, medium/low security and a community forensic service. Broadly, it is hypothesised that the risk assessments should reflect the level of security within each cohort, i.e. those in high security should show significantly higher risk than those in medium security who in turn should show significantly higher risk than those in low security. The study also employs two assessments which incorporate the dynamic variables which have shown to be predictive of violent incidents. One, the Short Dynamic

Risk Scale, has already been mentioned and the other, the Emotional Problem Scale, was developed by Prout and Hommer (1983) specifically for use with individuals with ID. In addition to these assessment instruments, violent and sexual incidents were recorded over a period of 12 months and the predictive value of each instrument is reported.

Method

Study Sites/Locations

The three study sites/locations will be referred to as follows: L1 (high security), L2 (medium/low security), L3 (community). L1 is the site for the national centre for high secure ID services, receiving patients from England and Wales. Patients are generally referred from prison, secure hospitals and the courts. The service accepts patients deemed to present a grave and immediate danger to the public. All patients may be considered for escorted ground access within the perimeter, only a small number have singleton escort and a small percentage have unescorted status. All educational, therapy, vocational and leisure facilities are provided within the secure perimeter by staff employed or contracted to the organisation to meet the specific needs of the resident population.

L2 provides in-patient forensic services on a local, regional and national basis to people with ID. Patients are generally referred via health authorities, the courts and prisons. There are several forensic units and wards providing medium secure, low secure and rehabilitation facilities. There are comprehensive assessment and treatment services for all psychological, psychiatric and criminogenic presentations.

L3 caters for a range of offenders with ID through a small 10 bedded open unit with a large number of day places associated with the service. The service has an area responsibility and all offenders with ID are referred for assessment and most are engaged in subsequent treatment. Most clients are treated while maintaining their community placement. Opportunities exist for clients to move between a variety of treatment options, allowing in-

patient, day-patient and out-patient treatment and comprehensive assessment both within the unit and in the community.

All three services have a history of dealing with such clients. The high security setting has been a location for offenders with ID for many years and there have been several reports over the years including this sample (e.g. Butwell, Jamieson, Leese & Taylor, 2000), the medium/low secure setting has been the subject of investigation over the years (Day, 1994; Taylor et al., 2002, 2005) and the community setting is again a mature service (Lindsay et al., 2004, 2005a). Therefore, the comparison between the three settings is likely to be a valid indication of differences in these client groups. A detailed description of referral information, psychiatric information, previous convictions and imprisonment, offences across the lifespan and details of the index offence can be found in Lindsay et al. (2005b),

Ethical Approval

Applications for ethical approval were made and granted locally at each site.

Assessment Process

All information reported in this study was available in clinical files. Where there was uncertainty as to the reliability or integrity of the information, further information was collected from relevant keyworkers and support workers. This further information was then incorporated and scored for the project. Studies such as this which employ casenote information for both historical and dynamic risk assessments have a reasonable pedigree (Quinsey et al., 1997; Hanson & Harris, 2000). However they do have limitations which will be discussed later in this paper.

Extensive training on reviewing, collecting and coding information through file review was conducted through a dedicated one week training course undertaken by all staff involved in the collection of study data. Appropriate training was conducted by individuals with competence and experience in relevant data collection. This training was done to ensure

consistency of data collection across sites. Research assistants then underwent a further two day exercise on ensuring reliability of data recording on a series of test cases.

Assessment Measures

Violence Risk Appraisal Guide (VRAG; Quinsey et al., 1998): The VRAG is the most widely researched actuarial risk assessment and its predictive validity has been demonstrated with individuals with ID (Quinsey, Book & Skilling, 2004). The VRAG has been tested extensively across client groups and cultures and has demonstrated significant predictive values in a range of settings. It includes items which are related to general antisociality such as childhood behaviour problems, history of non-violent offending, history of personality disorders and PCL-R scores.

HCR20 (Webster et al., 1995): The HCR20 is the most widely used structured clinical assessment for the prediction of violence. It contains 10 historical items such as history of mental illness and history of antisocial behaviour, five clinical items such as current impulsivity and current mental illness and five risk items such as access to stimulants or alcohol. Each item is scored on a three point scale from 0 indicating absence of the risk factor through 1 indicating possible presence of the risk factor to 2 indicating definite presence of the risk factor. The HCR20 has also been tested extensively across client groups and cultures and has shown significant predictive values in a range of settings. However, the authors do not recommend that it is used as an actuarial risk assessment but rather it is used as an instrument to structure clinical judgement. For the purposes of the present study, it has been used as an actuarial risk assessment with total scores for each section.

RM 2000 (Thornton, 2000): The RM 2000 is a risk assessment frequently used in the UK but with less validation than the VRAG. The predictive results reported by Thornton (2000) suggest that it is as valid a predictor of recidivism as the VRAG but it does not have such extensive testing and comparison. The RM 2000 provides values for risk of sexual

recidivism (RM 2000/S including items such as number of previous sexual charges, any conviction for sexual offence against a male, non contact sexual offences), risk of violent recidivism (RM 2000/V including items of age at commencement of risk, number of previous violent offences, history of burglary) and a combined risk (RM 2000/C). Craig, Browne and Stringer (2004) used the RM 2000 in a comparison between sex offenders receiving probation and those in a regional secure unit. There were significant differences in the risk scores of the two samples with the somewhat surprising finding that on both the RM 2000 and the Static-99, those in the probation service scored significantly higher.

Static-99 (Hanson & Thornton, 2000): The Static-99 is an instrument extensively employed in studies reviewing sexual offence recidivism. It has also been employed with clients with ID (Boer, Tough & Haaven, 2004) and includes items such as offending against males, offending against children, offending against strangers, non contact sexual offences, persistence of sexual offences and whether the offender has ever been married.

Short Dynamic Risk Scale (SDRS; Quinsey, 2004): The SDRS is an eight item rating scale which records a range of dynamic factors including hostile attitude, coping skills, self-care skills and consideration of others. It has been shown to predict violent and sexual incidents with significant accuracy (Quinsey, Book & Skilling, 2004) and is easily available (Quinsey, 2004).

Emotional Problem Scale (EPS; Prout and Strohmer, 1991): The EPS is a rating instrument developed for individuals with mild ID. It includes scales on thought disorder, physical aggression, non-compliance, anxiety, distractibility, depression, hyperactivity, withdrawal, self-esteem, verbal aggression, somatic concerns and sexual maladjustment. These scales factor analyse into two main factors, externalising behaviour problems including physical aggression, non-compliance, hyperactivity and verbal aggression; and internalising

behaviour problems including anxiety, depression and self-esteem. The composite on these two factors will be used in the present study.

Violent and sexual incidents were also recorded over a period of one year. In order to gather sufficient numbers of incidents to employ predictive statistics, all incidents were combined and analysed against the risk prediction assessments using receiver operator characteristics (ROC) analyses. The incidents were recorded independent of the study through the nursing casenotes. Any significant incident recorded in the casenotes was incorporated in the study. A significant incident was defined as recorded verbal aggression, recorded physical aggression, recorded destruction of property and recorded inappropriate sexual behaviour. There were no incidents of absconding.

Participants

A total of 212 adult males participated in the study. Parallel sets of data were collected across the three sites. For L1, the total cohort of individuals with ID in the hospital were included in the study ($n = 73$). For L2, a random sample of individuals with ID were drawn from the seven forensic units to make up a corresponding cohort to those in L1 ($n = 70$). For L3, the sample consisted of consecutive referrals counting backwards from the most recent until a corresponding cohort to L1 was reached ($n = 69$). In this way, three roughly equal cohorts were employed with no obvious source of bias in any of the samples. A table of basic characteristics is presented in Table I including mean age, mean IQ, the percentage of each cohort who have been diagnosed with mental illness, the percentage with a history of violent offences and the percentage with a history of sexual offences. There was no difference between the groups on IQ and participants from L3 were significantly younger than those in the other two settings. There were no significant differences between the groups on the percentage who had committed a sexual offence. A significantly higher percentage of participants in L1 had committed violent offences and a significantly lower

percentage of participants in L2 were diagnosed with major mental illness (psychotic disorders or major mood disorders). The actuarial assessments were completed for all participants. Unfortunately because of insufficient information in the casenotes, participants having left the services, inability to contact carers and difficulty in contacting psychiatrists/psychologists who had left the services, the SDRS was completed on 145 participants (48, 49 and 48 from L1, L2 and L3 respectively) and the EPS was completed for 169 participants (59, 55 and 55 from L1, L2 and L3 respectively). Details on characteristics can be found in Lindsay et al. (2005b).

Results

Reliability

For each assessment, reliability was calculated as percentage agreement between trained, independent raters using the formula number of agreed ratings divided by the number of agreements plus disagreements, expressed as a percentage. For the VRAG, on risk categories 1-9, agreement was 92.2%. For the HCR20 H Scale, agreement was 89.4%, for the C Scale, 93.1% and for the R Scale, 82.7%. For the RM 2000 reliability was 90.7% for the V Scale and 92.1% for the S Scale and for the Static-99, reliability for risk levels 1-4 was 97.2%. Independent raters rated 30 cases.

For the SDRS, average reliability on each scale across 30 participants was 88.6%. Reliability was not calculated for the EPS since this assessment had been standardised for use with people with intellectual disabilities. Therefore its reliability and validity for the population has already been established.

Group Comparisons

Table II. Table II shows the mean scores on each assessment for participants in L1, L2 and L3. The mean scores and standard deviations for the whole sample are also inserted. Of the actuarial assessments, the VRAG, HCR20 H Scale and RM 2000/C showed a

significant difference between groups. Both the VRAG and the HCR20 H show similar lawful patterns with average risk scores for L1 higher than L2 and L2 higher than L3. For the VRAG there was a significant difference between the risk score for L1 and L3. For the HCR20 there were significant differences between L1 and L3 and L2 and L3. There were no significant differences between L1 and L2. On the RM 2000/C, the average score for L1 is significantly higher than L2 and L3.

For the Dynamic Risk Assessments only the EPS internalising scale showed significant differences across groups with the average score for L1 higher than L2 which is in turn higher than L3. The difference between L1 and L3 is significant.

Prediction of incidents. The data from all participants was combined in order to compute predictive accuracy. Predictive accuracy was computed for each assessment using the area under the receiver operator characteristic curve (AUC). For a succinct explanation of AUCs see MacMillan, Hastings and Caldwell (2004).

Table III

Table III shows the test, the AUC and the significance of the predictive value for each assessment. As can be seen, of the static risk assessments, the VRAG and all categories of the HCR20 show significant AUC values. Of these significant values, the VRAG and HCR20 total score have the highest AUC. All three dynamic predictors, the SDRS, the EPS externalising and EPS internalising all had significant predictive value as indicated by the significant AUC.

Discussion

The present study assesses the discriminative validity and the predictive validity of several historical/actuarial risk assessments and dynamic/proximal risk assessments. Three groups of offenders with ID who have been administratively defined as requiring maximum security, medium/low security and community services respectively have been employed to

investigate discriminative validity and the data from all samples have been combined to investigate predictive validity.

The VRAG , HCR20 H Scale and the EPS Internalising Scale demonstrated significant discriminative validity and predictive validity. For the VRAG and the HCR20 H Scale, the mean risk scores across groups showed a lawful pattern with L1 having the highest mean, L2 an intermediate mean score and L3 the lowest. Both scales also had a significant AUC of 0.71 for the VRAG and 0.68 for the HCR20 H. There was also an AUC of 0.72 for the HCR20 total score. All of these results are consistent with the original standardisation of the VRAG (Quinsey et al., 1998) and the other comparative studies reported in the introduction. The VRAG results are also consistent with a cohort of individuals with ID in Canada who show disruptive and challenging behaviour (Quinsey, Book & Skilling, 2004). These results support findings across settings and cultures and from the only previous study on individuals with ID.

Interestingly, the EPS Internalising Scale also shows discriminative validity and predictive validity. The EPS Internalising Scale again discriminates between groups in a lawful fashion with L1 showing a higher mean score than L2 which in turn has a higher mean score than L3. In addition, both scales on the EPS have significant predictive validity with the Internalising Scale showing an AUC of 0.73 and the Externalising Scale an AUC of 0.75. Therefore the EPS shows good discriminative validity and the Externalising Scale has the highest AUC of all the assessments in this study which suggests that it may be a significant predictor of violent or sexual incidents. This finding alone may be important for intellectual disability services. The EPS has been developed for people with ID and the present results validate the utility of both scales for discriminating between groups who are administratively defined as being of higher or lower risk and therefore evaluating the severity of symptomatology and for predicting challenging and violent behaviour. These two subscales

may be useful for the general assessment of individuals with ID and for the assessment of those showing forensic and challenging behaviour in particular. In addition, the VRAG and HCR20 also show significant predictive results in relation to violent incidents.

Other scales show either discriminative or predictive validity. The RM 2000/C shows significant differences between groups with L1 having higher average risk scores. However, when these scores are computed against actual incidents, the predictive value was little better than chance. Therefore, the use of the RM 2000 for the prediction of incidents in people with ID is not supported. The RM 2000/S, the RM 2000/V and the Static-99 did not discriminate significantly between groups nor did they predict recorded incidents. Since these assessments are specific to violent or sexual incidents it may be that by combining all recorded incidents, we have diluted the predictive ability of these scales, however, it has to be said that the VRAG is also designed to predict violent offending and does show significant predictive results with the combined incident recording.

The results from the SDRS are consistent with those reported by Quinsey (2004). He found that scores recorded for participants using the SDRS were significantly higher in the month preceding an incident than in a prior month which did not precede an incident. The present study provides further validation for this finding. With an AUC of 0.71, the SDRS is likely to predict incidents with an accuracy at least equal to other valid risk predictors. Given its ease of use and availability it warrants further investigation.

Interestingly, the SDRS does not discriminate between groups. This would suggest that in day to day matters, participants in each of these settings show similar levels of dynamic risk. The most likely reason for this is that these services are relatively mature and staff are highly experienced in responding to and managing individuals who are showing dynamic risk factors. When staff are sensitive to the various manifestations of risk variables, they may pre-empt them or de-escalate them as they begin. Alternatively, it may be that

these individuals indeed show similar levels of antisociality, aggression, verbal threats and so on. This would suggest that there may be some individuals who are currently retained in high secure settings who could be managed under conditions of lesser security or even in community settings. While the current study would suggest this, further studies on pathways into forensic services would be required to provide more definitive information. For example, it may be that the availability of community forensic services at a local level is a major determinant of whether or not an individual is referred to medium or high security services. Alternatively, it may be the severity of the index crime rather than an individual's dynamic risk or manageability which determines referral to higher secure services. Under both these pathway conditions, there is the possibility that there would not be any significant difference between the dynamic risk of those placed across levels of security.

One particular issue should promote caution in the interpretation of the results of the current study. Many of the results depend on the quality of the casenotes and the accuracy of information recorded in the casenotes. In the present study, the casenotes were generally extensive and any ambiguous information was cross checked with relevant staff members. However, it remains the case that there may well be biases in those who have made up the casenote information and retrospective recall biases contained in the notes. All of these are common to studies such as the present one, but it undoubtedly remains a cautionary note.

References

- Banks, S., Robbins, P.C., Silver, E., Besselinov, R., Steadman, H.J., Monahan, J., Mulvey, E.P., Appelbaum, P.S., Grisso, T., & Roth, L.H. (2004). A multiple-models approach to violence risk assessment among people with mental disorder. *Criminal Justice & Behaviour, 31*, 324-340.
- Barbaree, H.E., Seto, M.C., Langton, C.M., & Peacock, E.J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice & Behaviour, 28*, 490-521.
- Bartosh, D.L., Garby, T., Lewis, D., & Gray, S. (2003). Differences in the predictive validity of actuarial risk assessments in relation to sex offender type. *International Journal of Offender Therapy & Comparative Criminology, 47*, 422-438.
- Boer, D.P., Hart, S.D., Kropp, P.R., & Webster, C.D. (1997). *Manual for the sexual violence risk – 20: professional guidelines for assessing risk of sexual violence*. Vancouver, British Columbia: British Columbia Institute on Family Violence & Mental Health, Law & Policy Institute, Simon Fraser University.
- Boer, D.P., Tough, S., & Haaven, J. (2004). Assessment of risk manageability of developmentally disabled sex offenders. *Journal of Applied Research in Intellectual Disabilities, 17*, 275-284.
- Buttwell, M., Jamieson, E., Leese, M., & Taylor, P. (2000). Trends in special (high security) hospitals: 2: residency and discharge episodes, 1986-1995. *British Journal of Psychiatry, 176*, 260-265.
- Cooke, D.J., Michie, S., & Ryan, J. (2002). *Evaluating risk for violence: A preliminary study of the HCR-20, PCL-R and VRAG on a Scottish prison population*. Edinburgh, UK: Scottish Prison Service.

- Craig, L.A., Browne, K.D., & Stringer, I. (2004). Comparing sex offender risk assessment measures on a UK sample. *International Journal of Offender Therapy & Comparative Criminology*, 48, 7-27.
- Day, K. (1994). Male mentally handicapped sex offenders. *British Journal of Psychiatry*, 165, 630-639.
- Epperson, D.L., Kaul, J.D., & Hesselton, D. (1998). *Final report on the development of the Minnesota Sex Offender Screening Tool – Revised (MnSOST-R)*. Paper presented at the 17th Annual Conference of the Association for the Treatment of Sexual Abusers, Vancouver, Canada.
- Hanson, R.K. (1997). *The development of a brief actuarial risk scale for sexual offence recidivism*. (User report 1997-04). Ottawa: Department of the Solicitor General of Canada.
- Hanson, R.K., & Harris, A.J.R. (2000). Where should we intervene? Dynamic predictors of sexual offence recidivism. *Criminal Justice and Behaviour*, 27, 6-35.
- Hanson, R.K., & Thornton, D. (1999). *Static-99: improving actuarial risk assessments for sex offenders*. (User report 1999-02). Ottawa: Department of the Solicitor General of Canada.
- Hare, R.D. (1991). *The Hare Psychopathy Checklist – Revised*. Toronto, Ontario: Multi Health Systems.
- Harris, G.T., & Rice, M.E. (2003). Actuarial assessment of risk among sex offenders. Sexually coercive behaviour: Understanding and management. *Annals of the New York Academy of Sciences*, 989, 198-210.
- Kroner, D.G., & Mills, J.F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behaviour*, 28, 471-489.

- 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., Hogue, T., Taylor, J., Steptoe, L., Mooney, P., Pinkney, L., Johnston, S., Smith, A.H.W., & O'Brien, G. (2005b). A comparison of offenders with intellectual disability across three levels of security. *Criminal Behaviour and Mental Health, in press*.
- Lindsay, W.R., Smith, A.H.W., Law, J., Quinn, K., Anderson, A., Smith, A., & Allan, R. (2004). Sexual and non-sexual offenders with intellectual and learning disabilities: A comparison of characteristics, referral patterns and outcome. *Journal of Interpersonal Violence, 19*, 875-890.
- Lindsay, W.R., Steele, L., Smith, A.H.W., Quinn, K., & Allan, R. (2005a). A community forensic intellectual disability service: Twelve year follow-up of referrals, analysis of referral patterns and assessment of harm reduction. *Legal and Criminological Psychology, in press*.
- MacMillan, D., Hastings, R., & Caldwell, 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*, 255-266.
- Monahan, J. (2002). The MacArthur studies of violence risk. *Criminal Behaviour and Mental Health, 12*, S67-S72.
- Novaco, R.W., & Taylor, J.L. (2004). Assessment of anger and aggression in offenders with developmental disabilities. *Psychological Assessment, 16*, 42-50.
- Prout, T.H., & Strohmmer, D.C. (1991). *Emotional Problem Scale*. Lutz, FL: Psychological Assessment Resources, Inc.

- 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: John Wiley & Sons.
- Quinsey, V.L., Book, A., & Skilling, T.A. (2004). A follow-up of deinstitutionalised developmentally handicapped men with histories of antisocial behaviour. *Journal of Applied Research in Intellectual Disabilities*, 17, 243.
- Quinsey, V.L., Coleman, G., Jones, B., & Altrows, J. (1997). Proximal antecedents of eloping and re-offending among supervised mentally disordered offenders. *Journal of Interpersonal Violence*, 12, 794-813.
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Cromier, C.A. (1998). *Violent offenders: appraising and managing risk*. Washington DC: American Psychological Association.
- Sjostedt, G., & Langstrom, N. (2002). Assessment of risk for criminal recidivism among rapists: a comparison of four different measures. *Psychology, Crime & Law*, 8, 25-40.
- Taylor, J.L., Novaco, R.W., Gillmer, B., & Thorne, I. (2002). Cognitive behavioural treatment of anger intensity among offenders with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 15, 151-165.
- Taylor, J.L., Novaco, R.W., Gillmer, B.T., Robertson, A., & Thorne, I. (2005). Individual cognitive behavioural anger treatment for people with mild-borderline intellectual disabilities and histories of aggression: A controlled trial. *British Journal of Clinical Psychology*, in press.
- Thornton, D. (2000). *Risk Matrix 2000*. Unpublished.
- Tough, S. (2001). *Validation of two standard assessments (RRASOR, 1997; Static-99, 1999) on a sample of adult males who are intellectually disabled with significant cognitive*

deficits. Unpublished Masters Thesis, University of Toronto, Toronto, Ontario, Canada.

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.

Table I: Characteristics of participants in the three groups: High security (L1), medium/low security (L2), and the community (L3)

Characteristics	(L1)	(L2)	(L3)
Mean Age*	38.7	39	34.3
Mean IQ	66.6	66.7	64.7
Mental Illness (% of total)**	37	10	30.3
Violent Offences (% of total)**	69.9	35.7	24.6
Sexual Offences (% of total)	56.2	58.8	55.2

**p < 0.01

* p < 0.05

Table II: The means and standard deviations on each assessment for all three groups and the whole sample, f values, degrees of freedom and probability levels of the difference between the groups. Superscripts indicate that a is significantly greater than b at the 1% level. Where no superscript is inserted between two superscripts, the value is intermediate between the two significant values but significantly different from neither

	L1, mean (sd)	L2, mean (sd)	L3, mean (sd)	Whole Sample	f	df	p
VRAG	8.67 (8.35) ^{a**}	7.03 (8.04)	4.39 (9.64) ^{b**}	6.73 (8.84)	4.35	2,209	0.014
HCR H	14.61 ^{a**} (2.69)	13.75 ^{a**} (2.72)	8.30 ^{b**} (4.37)	12.09 (4.43)	59.17	2,207	0.000
HCR C	4.45 (2.54)	4.84 (2.34)	3.96 (2.26)	4.41 (2.40)	0.79	2,207	0.375
HCR R	3.55 (1.88)	2.63 (1.42)	2.96 (1.91)	3.04 (1.78)	3.621	2,207	0.06
RM 2000/S	1.97 (1.20)	1.58 (1.11)	1.78 (1.58)	1.77 (1.32)	0.81	2,207	0.45
RM 2000/V	3.23 (1.88)	2.60 (1.88)	2.60 (1.72)	2.83 (1.85)	2.79	2,207	0.64
RM 2000/C	3.57 (1.48) ^{a**}	2.52 (1.34) ^{b**}	2.66 (1.53) ^{b**}	2.92 (1.52)	4.73	2,207	0.011
Static-99	4.22 (2.58)	4.14 (2.16)	3.74 (2.39)	4.03 (2.39)	1.78	2,207	0.17
SDRS	12.18 (7.71)	9.65 (8.68)	10.86 (8.83)	10.92 (8.41)	1.76	2,144	0.17
EPS Externalising	48.97 (32.06)	44.85 (30.82)	38.87 (29.91)	44.34 (31.06)	1.523	2,168	0.22
EPS Internalising	32.08 (23.25) ^{a**}	26.56 (19.92)	17.93 (18.95) ^{b**}	25.68 (21.53)	6.64	2,168	0.002

Table III: The AUC and probability level for each assessment calculated against recorded incidents.

Assessment	AUC	p
VRAG	0.71	0.000
HCR20 Total	0.72	0.000
HCR20 H	0.68	0.000
HCR20 C	0.67	0.000
HCR20 R	0.62	0.02
RM 2000/S	0.54	0.58
RM 2000/V	0.62	0.07
RM 2000/C	0.54	0.58
Static-99	0.53	0.64
SDRS	0.72	0.000
EPS Externalising	0.75	0.000
EPS Internalising	0.73	0.000