Predictive validity of the HCR-20 for violent and non-violent sexual behaviour in a secure mental health service

Laura E. O'Shea, Dev-Kishan Thaker, Marco M. Picchioni, Fiona L. Mason, Caroline Knight, Geoffrey L. Dickens

Predictive validity of the HCR-20 for violent and non-violent sexual behaviour in a secure mental health service

Short Title: HCR-20 prediction of inappropriate sexual behaviour

Laura E. O'Shea[ab], Dev-Kishan Thaker[a], Marco M. Picchioni[ab], Fiona L. Mason[c], Caroline Knight[c], Geoffrey L. Dickens*[a,d]

[a] St Andrew's Academic Department, Billing Road, Northampton, NN1 5DG, United Kingdom

[b] Institute of Psychiatry, King's College London, 16 De Crespigny Park, London, SE5 8AF

[c] St Andrews, Billing Road, Northampton, NN1 5DG, United Kingdom

[d] School of Social and Health Sciences, Abertay University, Bell Street, Dundee, DD1 1HG, UK

*Correspondence: Geoffrey L. Dickens, School of Social and Health Sciences, Abertay University, Bell Street, Dundee, Scotland, DD1 1HG, UK. Email: g.dickens@abertay.ac.uk. Tel: +441382 308257

Abstract

Background: Violent and non-violent sexual behaviour is a fairly common problem among secure mental health service patients, but specialist sexual violence risk assessment is time consuming and so done infrequently.

Aims: We aimed to establish whether a commonly used violence risk assessment tool, the HCR-20, has predictive validity specifically for inappropriate sexual behaviour.

Methods: A pseudo-prospective cohort design was used for a study in the adult wards of a large provider of specialist secure mental health services. Routine clinical team HCR-20 assessments were extracted from records and incidents involving inappropriate sexual behaviour were recorded for the 3-months following assessment.

Results: 104 (17%) of 613 patients had engaged in at least one inappropriate sexual behaviour; in 65 (10.6%) the sexual act was violent. HCR-20 total score, clinical and risk management subscales predicted violent and non-violent sexual behaviour. The negative predictive value of the HCR-20 for inappropriate sexual behaviour was over 90%.

Conclusions: Prediction of violent sexual behaviour may be regarded as well within the scope of the HCR-20 as a structured professional judgement tool to aid violence risk prediction, but we found that it also predicts behaviours that may be of concern but fall below the violence threshold. High negative predictive values suggest that HCR-20 scores may have some utility for screening out patients who do not require more specialist assessment for inappropriate sexual behaviour.

Introduction

Inappropriate sexual behaviour is "a verbal or physical act of an explicit, or perceived, sexual nature, which is unacceptable within the social context in which it is carried out," (Johnson et al., 2006; p. 688). It encompasses acts which meet accepted definitions of sexual violence as actual, attempted or threatened non-consensual sexual contact (e.g., Boer et al., 1997; Hart et al., 2003) and those which may be of concern but fall below the threshold of this definition. Hughes and Hebb (2005) reported that around 40% of English medium-secure hospital inpatients had exhibited such behaviour at some time while in the unit. Inappropriate sexual behaviour by patients is associated with increased staff turnover, sickness-related absences, stress and low morale (Garcia et al., 2005; Hayward et al., 2013; Needham et al., 2005). It may contribute to the social exclusion of patients and delay their rehabilitation and recovery (Johnson et al., 2006). It is important, therefore, to be able to identify those most at risk so that relevant care and treatment strategies can be best focussed (Griffin et al., 2012; Walji et al., 2013). Sexual violence risk assessment instruments are, however, rather infrequently used in inpatient hospital units, where generic violence risk assessments are more commonly used (Khiroya et al., 2009), particularly the HCR-20 (Webster et al., 1997).

The HCR-20 has good predictive validity for future violence (Singh et al., 2011; Yang et al., 2010), including in the inpatient setting (Hogan et al., 2010; O'Shea et al., 2013). Investigation of its predictive validity for sexually violent behaviour has previously been subsumed under the broader category of aggression and has only been considered in a small number of studies (e.g., Nicholls et al., 2004; Tengström et al., 2006). We are not aware of any studies that have examined its predictive value with respect to non-violent but inappropriate sexual acts. There is emerging evidence

however, that the HCR-20 may be used with good effect to predict behaviour outside its specific target outcomes, notably self-harm (O'Shea et al., 2014b), possibly because of common risk factors. Our aim, therefore, was to investigate whether the HCR-20 can significantly predict both violent and non-violent inappropriate sexual behaviours by secure unit inpatients. Specifically, we set out to examine the predictive validity of the HCR-20 for such behaviours and to identify which individual HCR-20 items best predicted them. We also investigated the degree of overlap between inappropriate sexual behaviours and non-sexual aggression as this may have implications for the feasibility of using the HCR-20 to examine risk of both outcomes.

Method

The study was conducted as a service evaluation and approved by St Andrew's clinical audit and service evaluation committee.

Participants

St Andrew's provides secure mental health inpatient care at four sites in England. Eligible participants were adult inpatients between September 2010 and March 2011, who had had at least one HCR-20 risk assessment completed. Patients were excluded if their assessment was missing items in excess of prorating guidelines in the HCR-20 manual (Webster, et al., 1997).

Procedure

A pseudo-prospective design was used with routinely collected data. Risk assessment data and possibly inappropriate sexual behaviour occurring during the three months after risk assessment was identified from an anonymised version of patients' clinical records. We obtained all narrative entries that had been electronically flagged

for: verbal aggression, intimidation/bullying", sexual offending, physical aggression and self-harm/suicide". Rating was conducted by three trained researchers (LO, GD, DKT), blind to the HCR-20 risk assessments.

Measures

HCR-20 risk assessment

Assessment was conducted by clinical teams during routine practice using version 2 of the HCR-20 (Webster, et al., 1997; Version 3 Douglas et al., 2013) is now available but was not implemented in the study setting at the time of data collection. The risk-management policy for the study setting dictates that the HCR-20 should be completed by a clinician who has undertaken a two-day approved training in its use, covering theoretical and practical aspects and discussion of item scoring and report writing. Graduate psychology assistants may complete assessments after training, but only under the direct supervision or a psychiatrist or clinical psychologist. Assessments should be completed from multiple sources of information, in accordance with guidelines in the manual for version 2 of the HCR-20 (Webster, et al., 1997). Version 2 comprises 20 items across 3 subscales which cover Historical (H), Clinical (C) and Risk-management (R) factors (*see appendix, online version only*). Each item is rated as 0 (not present), 1 (possibly present) or 2 (definitely present) (Webster et al., 2001). "Psychopathy" (H item7) was omitted because most patients were not assessed in this regard; exclusion of this item has a negligible effect on the predictive effectiveness of the HCR-20 (Dolan et al., 2004; Wilson et al., 2013). The risk-management items are rated twice, once under the assumption that the individual is institutionalised (In) and once under the assumption that they are in the community (Out). As our study was concerned with inpatient behaviours, we used only the 'In' ratings. Clinicians also make

a summary judgement regarding the risk of future violence; classified as low, moderate, or high.

Inappropriate sexual behaviour

Risk outcomes were coded using the St. Andrews Sexual Behaviour Assessment (SASBA; Knight et al., 2008). Incidents are assigned to one of four categories: verbal comments, non-contact, exposure or touching others. Each category has a criterion-referenced severity scale from 1 (least severe) to 4 (most severe); for example, the verbal comments category ranges from intimate personal comments of mild severity (level 1) to explicit person-directed accounts of sexual intent, requests or activity (level 4). The SASBA has strong construct and content validity, inter-rater, and test-retest reliability (Knight et al., 2008). In our study, we tested inter-rater reliability on 100 incidents, with good categorical agreement (K 0.75) and excellent severity agreement (K 0.84–1.0, mean 0.96). SASBA recordings were further categorised into non-violent but inappropriate sexual behaviours (verbal comments levels 1-3, non-contact levels 1-2, exposure levels 1-2 and/or touching others level 1) or into violent and inappropriate sexual behaviours (verbal comments level 4, non-contact levels 3-4, exposure levels 3-4 and/or touching others levels 2-4). This categorisation is consistent with the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003) definition of sexual violence.

Non-sexual aggressive outcomes

Incidents of non-sexual aggressive behaviour were coded as part of a previous study (O'Shea et al., 2014a) using the Overt Aggression Scale (OAS; Yudofsky et al., 1986) . Inter-rater agreement was in the excellent range (K 0.81-1.0).

Demographic and clinical data

Information regarding age, gender, date of admission, date of discharge (if applicable), legal status, security level, ethnicity, and ICD-10 (World Health Organisation, 1992) psychiatric diagnoses, as recorded by each patient's responsible clinician, were obtained from patient records.

Data Analysis

Demographic and clinical characteristics were summarised using descriptive statistics. Pearson's Chi squared tests were used to identify differences in rates of inappropriate sexual behaviours as a function of age and gender, and to investigate the overlap between perpetration of violent inappropriate sexual behaviour, non-violent inappropriate sexual behaviour and non-sexual aggression. Independent *t*-tests were conducted to compare mean HCR-20 scores as a function of age, gender, and perpetration of inappropriate sexual behaviours. Cohen's *d* values were also calculated to quantify the size of differences between mean scores; 0.2, 0.5, and 0.8 are typically considered the thresholds for small, moderate, and large effect sizes respectively (Cohen, 1992). Positive predictive values (PPV), negative predictive values (NPV), sensitivity and specificity for the HCR-20 summary judgment were calculated to indicate how accurately the tool was identifying true positive and true negative cases in this sample; for this analysis, patients rated as low risk were treated as test negative cases and those rated as moderate or high risk as test positive. Area under the curve (AUC) values derived from receiver operating characteristic (ROC) analysis were used to investigate the predictive validity of HCR-20 scores for inappropriate sexual behaviours; AUC values can range from 0–1 with 0.5 indicating prediction no better

than chance, and 0.75 suggesting a large effect size (Dolan et al., 2000). Predictive validity was calculated for the total HCR-20 score, H, C and R subscale scores separately, summary judgement and all individual items.

Results

Sample characteristics

Of 692 patients who met the inclusion criteria, 51 were excluded due to missing data and 28 because they did not remain in the service for at least three months post-assessment. This left a sample of 613 (89% of the total) for full analysis, 418 (68%) men and 195 (32%) women. Mean age was 38.5 years (standard deviation [SD] 15.48). Almost two thirds (387, 63%) identified themselves as Caucasian, 54 (9%) as Black, African, Caribbean, or Black British, 29 (5%) as mixed ethnic background, 29 (5%) as Asian or Asian British and 3 (<1%) as any other ethnic group, while 111 (18%) did not state their ethnicity. Over half (348, 57%) were located on low secure wards, 173 patients (28%) were on medium secure wards, 70 (11%) on locked wards and 22 (4%) on open wards. Most (329, 54%) were admitted under a forensic section of the Mental Health Act 1983 (amended 2007), 243 (40%) under a civil section and 31 (7%) were informal patients. Average time between admission and assessment was 829 days (SD 1253). The most common ICD-10 diagnoses were schizophrenia, schizotypal or delusional disorders (F20-F29; 334, 33%), disorders of adult personality and behaviour (F60-F69; 228, 22%), mental retardation (F70-F79; 108, 11%), mental and behavioural disorders due to psychoactive substance use (F10-F19; 98, 10%), disorders of psychological development (F80-F89; 94, 9%), organic, including symptomatic, mental disorders (F00-F09; 60, 6%), and mood [affective] disorders (F30-F39; 46, 5%).

HCR-20 scores

The mean HCR-20 score was 27.35 (SD 5.56, range 7–38); subscale averages were H 13.79 (SD 3.01, range 0–18), C 6.81 (SD 2.33, range 0–10) and R 6.75 (SD 2.48, range 0–10).

Mean HCR-20 scores differed across age groups; those aged under 40 had a higher HCR-20 score than those aged 40 or above (27.78 vs. 26.79; $t(611)$ -2.18, p=0.030; d= -0.35). HCR-20 scores also differed by gender, with women scoring more highly than men (28.19 vs. 26.97; t (611) -2.54, p=0.011; d= -0.41).

Clinical, risk-management and HCR-20 total scores differed significantly between those who had obtained SASBA scores and those who had not. Clinical scores differed significantly for all categories; risk-management and HCR-20 total scores were found to be significantly different for verbal comments, touching others, non-violent and violent inappropriate sexual behaviours (see Table 1).

[INSERT TABLE 1 ABOUT HERE]

Incidents of inappropriate sexual behaviour

During the 'time at risk' – the follow-up period of 3 months – there were 239 incidents of inappropriate sexual behaviour involving 104 patients. Overall, 52 (9%) had touched others, 47 (8%) had made verbal comments, 25 (4%) other non-contact behaviours and 23 (4%) had exposed themselves. Seventy patients (11%) had perpetrated at least one non-violent sexual act and 65 (11%) at least one violent one; 31 (30% of those involved in incidents) had done both, while 39 (37%) had only been non-

violent in this respect and 34 (32%) had committed only violent sexual acts. Those aged

under 40 were more likely to touch others sexually than those aged 40 or above (x^2

[N=613]7.44, p=0.006). Men were more likely to make comments of a sexual nature

than the women (x^2[N=613]15.17, p<0.001), but overall rates of non-violent sexual

behaviours were similar according to age and gender. Men, however, and those under

40 years old were more likely be sexually violent (see Tables 2 and 3).

[INSERT TABLES 2 and 3 ABOUT HERE]

Overlap between inappropriate sexual behaviours and aggression

A significantly larger proportion of those who fell into each SASBA outcome

category had been non-sexually aggressive than those who had not been sexually

inappropriate in the study period. This was most pronounced for verbal comments; *all*

of those who made inappropriate sexual verbal comments had also been non-sexually

aggressive to others, verbally and/or physically, compared with about two-thirds (62%)

of those who were not verbally inappropriate (χ^2[N=613]27.37, p<0.001). Of the 65

individuals who had engaged in violent inappropriate sexual behaviours, 64 (99%) had

also been non-sexually aggressive during the study period (χ^2[N=613]35.64, p<0.001).

Predictive validity

AUC values ranged from 0.451 to 0.688. HCR-20 total scores, C scores and R

scores significantly predicted sexual verbal comments and sexually touching others, as

well as the broader categories of violent and non-violent inappropriate sexual

behaviour; C scores also significantly predicted non-contact incidents. H scores and the

summary judgement were not predictive of any outcome. The largest AUC values were obtained for predictions based on C scores, with the exception of touching others, where the highest AUC value was for the total HCR-20 score (see Table 3).

Positive predictive values for violent and non-violent sexual behaviours according to the HCR-20 summary judgement were 12% and 13% respectively and negative predictive values 93% and 97%. Sensitivity of the summary judgement was high for both outcomes (non-violent inappropriate sexual behaviour 93%; violent 84%), whereas specificity was less so (27% non-violent inappropriate sexual behaviour; 26% violent). Inspection of co-ordinate points of the ROC curves revealed that sensitivity was perfect (1.0) for HCR-20 total scores up to 13.4 for prediction of both, and was greater than 0.9 for scores up to 23.1.

Item-outcome analysis

Inappropriate sexual verbal comments were best predicted by negative attitudes (item C2) and noncompliance with remediation attempts (R4); non-contact incidents were best predicted by lack of insight (C1) and noncompliance (R4); exposure was significantly predicted by impulsivity (C4) and, again, noncompliance (R4); touching others was best predicted by negative attitudes (C2) and unfeasible plans (R1). Non-violent and violent sexual behaviours were best predicted by negative attitudes and impulsivity. None of the AUC values reached the threshold for a large effect size.

Discussion

We have demonstrated that the HCR-20 total score and the Clinical (C) and Risk management scales can, separately, discriminate between patients who are and are not sexually aggressive during three months in an inpatient setting. C scores were higher

among those who had engaged in each category of sexual behaviour compared with those who had not, typically producing moderate to large effect sizes; HCR-20 total scores and R scores also differentiated between engagers and non-engagers, but typically with small to moderate effect sizes. This suggests that some dynamic risk factors are shared between general aggression against others and inappropriate sexual behaviours. This is supported by the fact that individuals showing sexually inappropriate behaviours were also more likely to have been non-sexually aggressive.

The finding that HCR-20 elements predict violent sexual behaviours might have been expected as, in the HCR-20 manual, Webster, et al. (1997; p. 25) state that "all sexual assaults should be considered violent behaviour". This could mean that the HCR-20 is simply predicting the aggressive element of the behaviour rather than the sexual element, but our finding that the HCR-20 also predicts non-violent but inappropriate sexual behaviour suggests that this is not the case. Further, AUC values for the prediction of non-violent and violent sexual behaviours were comparable, the non-violent having the largest number of significant relevant items of all the outcomes.

While HCR-20 total and C and R subscale scores were predictive of inappropriate sexual behaviour, none of the AUC values indicated a large effect (Dolan, et al., 2000) and they were smaller than those for aggression in the same sample in the same setting (O'Shea, et al., 2014a). Perhaps unsurprisingly, they were also smaller than AUC values obtained from tools designed to predict sexual violence (Fazel et al., 2012). The *summary judgement* did not predict any of the sexual outcomes, although it did indicate likelihood that any future violence would cause serious physical harm. This finding thus precludes any conclusions about the ability of clinicians to predict non-violent inappropriate sexual behaviour, but we think that the high sensitivity and

negative predictive value of the *summary judgement* suggests that, where the sexual offending status of an individual is unknown, a low risk summary judgement might indicate that an individual does not require more specialised sexual violence risk assessment. High HCR-20 scores, however, with the presence of additional risk factors for inappropriate sexual behaviour such as childhood abuse (Lee et al., 2002; Starzyk et al., 2003) and sexual deviancy (Hanson et al., 2005), may indicate the need for further assessment.

Item-outcome analysis

Item-outcome analysis suggested item redundancy with regard to prediction of inappropriate sexual behaviour. None of the historical items significantly predicted any outcomes, with the exception of H item 2 (young age at first violent incident) for touching others. The C scale contained the largest number of relevant items for all outcomes except non-violent inappropriate sexual behaviours, for which R scale items were more relevant. These items share a degree of similarity with items described in the Risk of Sexual Violence Protocol (RSVP; Hart, et al., 2003), which tends to support our finding.

Limitations

The main limitation of our study is that the risk assessment was conducted as part of routine clinical practice; therefore, the same individuals are tasked with both completing the risk assessment and implementing management strategies. ROC analysis cannot distinguish between false prediction and successful prevention of incidents, hence, the findings may underestimate the predictive ability of the HCR-20 for inpatient

sexual behaviours. Given, however, that the HCR-20 was not designed to predict them, and yet it did, we suggest that it may help clinicians to manage inappropriate sexual activity as well as aggression. Further, the way in which the HCR-20 was used more closely parallels what would occur in routine clinical practice, and is therefore more generalisable. It would have been beneficial to investigate if clinicians had introduced intervention strategies as a result of their ratings, in order to disentangle whether the apparent lower predictive validity is due to effective management of high risk patients. The relatively short follow-up period may have limited the opportunity for events of interest to occur, reducing both the number of incidents and the predictive ability of the HCR-20, but if such clinical aids are to be useful in practice, they must be able to perform over short as well as longer periods, and the creators of the HCR-20 do recommend its repeated and regular use with inpatients (Webster, et al., 1997).

Implications and future considerations

We found that the HCR-20 could predict inappropriate sexual behaviours even when these were not overtly violent, but the low positive predictive values confirm that it lacks specificity in this area. Pragmatically, therefore, it is likely to be maximally helpful in this context where the risk of inappropriate sexual behaviour is not already known, when elevated HCR-20 score may have some utility in identifying individuals for specific sexual violence risk assessment. Once it is established whether there is a significant sexual risk, clinicians can make a decision as to whether to continue with repeated assessments for sexual risk or revert to generic violence risk assessment only. Future research should determine whether interventions targeting impulsivity, negative attitudes, and lack of insight are effective in reducing inappropriate sexual behaviour.

References

Boer, D. P., Wilson, R. J., Gauthier, C. M., & Hart, S. D. (1997). Assessing risk for sexual violence: Guidelines for clinical practice. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: theory, assessment and treatment*. New York: Guildford.

Cann, J., Falshaw, L., & Friendship, C. (2004). Sexual offenders discharged from prison in England and Wales: A 21-year reconviction study. *Legal and Criminological Psychology, 9*(1), 1-10.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Cortoni, F., Hanson, R. K., & Coache, M.-È. (2010). The recidivism rates of female sexual offenders are low: A meta-analysis. *Sexual Abuse: A Journal of Research and Treatment, 22*(4), 387-401.

Dolan, M., & Doyle, M. (2000). Violence risk prediction. *The British Journal of Psychiatry, 177*(4), 303-311.

Dolan, M., & Khawaja, A. (2004). The HCR-20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior, 30*(6), 469-483. doi: 10.1002/ab.20044

Douglas, K., Hart, S., Webster, C., & Belfrage, H. (2013). HCR-20: Assessing risk for violence (Version 3). *Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University*.

Douglas, K. S., Webster, C. D., Hart, S. D., Eaves, D., & Ogloff, J. R. P. (Eds.). (2001). *HCR-20: Violence risk management companion guide*. Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University, and Department of Mental Health Law & Policy, University of South Florida.

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. *BMJ: British Medical Journal, 345*.

Garcia, I., Kennett, C., Quraishi, M., & Durcan, G. (2005). Acute care 2004. *A national survey of adult psychiatric wards in England. Sainsbury Centre for Mental Health, London*.

Griffin, H. L., & Vettor, S. (2012). Predicting sexual re-offending in a UK sample of adolescents with intellectual disabilities. *Journal of sexual aggression, 18*(1), 64-80.

Hanson, R. K. (2002). Recidivism and age: Follow-up data from 4,673 sexual offenders. *Journal of interpersonal violence, 17*(10), 1046-1062.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: a meta-analysis of recidivism studies. *Journal of consulting and clinical psychology, 73*(6), 1154.

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised*. Toronto: Multi-Health Systems.

Hart, S. D. (2001). Assessing and managing violence risk. In K. S. Douglas, C. D. Webster, S. D. Hart, D. Eaves & J. R. P. Ogloff (Eds.), *HCR-20 violence risk management companion guide* (pp. 13-26): Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University, and Department of Mental Health Law & Policy, University of South Florida.

Hart, S. D., Laws, D. R., Simon Fraser University. Mental Health, L., Institute, P., Kropp, P. R., Violence, B. C. I. A. F., . . . Corporation, P. P. A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*: Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

Hayward, L. E., Robertson, N., & Knight, C. (2013). Inappropriate sexual behaviour and dementia: An exploration of staff experiences. *Dementia, 12*(4), 463-480. doi: 10.1177/1471301211434673

Hogan, N., Ennis, L., & Assessment, F. A. (2010). Assessing Risk for Forensic Psychiatric Inpatient Violence: A Meta-analysis. *Open Access Journal of Forensic psychology*(2), 137-147.

Hughes, G. V., & Hebb, J. (2005). Problematic sexual behaviour in a secure psychiatric setting: Challenges and developing solutions. *Journal of sexual aggression, 11*(1), 95-102.

Johnson, C., Knight, C., & Alderman, N. (2006). Challenges associated with the definition and assessment of inappropriate sexual behaviour amongst individuals with an acquired neurological impairment. *Brain Injury, 20*(7), 687-693.

Khiroya, R., Weaver, T., & Maden, T. (2009). Use and perceived utility of structured violence risk assessments in English medium secure forensic units. *Psychiatric Bulletin, 33*, 129-132.

Langevin, R., Curnoe, S., Fedoroff, P., Bennett, R., Langevin, M., Peever, C., . . . Sandhu, S. (2004). Lifetime sex offender recidivism: A 25-year follow-up study. *Canadian Journal of Criminology and Criminal Justice/La Revue canadienne de criminologie et de justice pénale, 46*(5), 531-552.

Lee, J. K., Jackson, H. J., Pattison, P., & Ward, T. (2002). Developmental risk factors for sexual offending. *Child abuse & neglect, 26*(1), 73-92.

Needham, I., Abderhalden, C., Halfens, R. J., Fischer, J. E., & Dassen, T. (2005). Non-somatic effects of patient aggression on nurses: a systematic review. *Journal of advanced nursing, 49*(3), 283-296.

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing Risk for Violence among Male and Female Civil Psychiatric Patients: The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law, 22*(1), 127-158.

O'Shea, L. E., Mitchell, A. E., Picchioni, M. M., & Dickens, G. L. (2013). Moderators of the predictive efficacy of the historical, clinical and risk management-20 for aggression in psychiatric facilities: Systematic review and meta-analysis. *Aggression and Violent Behavior, 18*, 255-270.

O'Shea, L. E., Picchioni, M. M., Mason, F. L., Sugarman, P. A., & Dickens, G. L. (2014a). Differential predictive validity of HCR-20 for inpatient aggression. *Psychiatry Research, 220*, 669-678.

O'Shea, L. E., Picchioni, M. M., Mason, F. L., Sugarman, P. A., & Dickens, G. L. (2014b). Predictive validity of the HCR-20 for inpatient self-harm. *Comprehensive Psychiatry, 55*(8), 1937-1949.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: a systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical psychology review, 31*(3), 499-513.

Starzyk, K. B., & Marshall, W. L. (2003). Childhood family and personological risk factors for sexual offending. *Aggression and Violent Behavior, 8*(1), 93-105. doi: http://dx.doi.org/10.1016/S1359-1789(01)00053-2

Tengström, A., Hodgins, S., Müller-Isberner, R., Jöckel, D., Freese, R., Özokyay, K., & Sommer, J. (2006). Predicting Violent and Antisocial Behavior in Hospital Using the HCR-20: The Effect of Diagnoses on Predictive Accuracy. *International Journal of Forensic Mental Health, 5*(1), 39-53. doi: 10.1080/14999013.2006.10471229

Walji, I., Simpson, J., & Weatherhead, S. (2013). Experiences of engaging in psychotherapeutic interventions for sexual offending behaviours: A meta-synthesis. *Journal of Sexual Aggression*(ahead-of-print), 1-23.

Webster, C. D., & Douglas, K. S. (2001). Purpose of the companion guide. In K. S. Douglas, C. D. Webster, S. D. Hart, D. Eaves & J. R. P. Ogloff (Eds.), *HCR-20: Violence risk management companion guide*. Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University, and Department of Mental Health Law & Policy, University of South Florida.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence (version 2)*. Burnaby, British Columbia: Mental Health, Law and Policy Institute, Simon Fraser University.

Wilson, C. M., Desmarais, S. L., Nicholls, T. L., Hart, S. D., & Brink, J. (2013). Predictive validity of dynamic factors: assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior 37*(6), 377-388.

World Health Organisation. (1992). *The ICD-10 classification of mental and behavioural disorders*. Geneva: WHO.

Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools. *Psychological bulletin, 136*(5), 740.

Yudofsky, S. C., Silver, J. M., Jackson, W., Endicott, J., & Williams, D. (1986). The Overt Aggression Scale for the objective rating of verbal and physical aggression. *The American Journal of Psychiatry, 143*(1), 35-39.

Online appendix: HCR-20 items

Historical Items

	H1	Previous Violence
	H2	Young Age at First Violent Incident
	H3	Relationship Instability
	H4	Employment Problems
	H5	Substance Use Problems
	H6	Major Mental Illness
	H7	Psychopathy
	H8	Early Maladjustment
	H9	Personality Disorder
	H10	Prior Supervision Failure

Clinical Items

	C1	Lack of Insight
	C2	Negative Attitudes
	C3	Active Symptoms of Major Mental Illness
	C4	Impulsivity
	C5	Unresponsive to Treatment

Risk Management Items

	R1	Plan Lacks Feasibility
	R2	Exposure to Destabilizers
	R3	Lack of Personal Support
	R4	Noncompliance with Remediation Attempts
	R5	Stress

Table 1: Mean HCR-20 scores as a function of SASBA recordings

	Yes	No	Test	Cohen's d
VC				
HCR-20 Total	29.82 (SD=4.90)	27.15 (SD=5.58)	$t(611)=3.18, p=.002$	0.51
Historical	13.81 (SD=3.31)	13.79 (SD=2.99)	$t(611)=0.05, p=0.957$	0.01
Clinical	8.10 (SD=1.58)	6.71 (SD=2.35)	$t(64.307)=5.55, p<0.001$	0.69
Risk-management	7.91 (SD=1.99)	6.66 (SD=2.49)	$t(58.641)=4.05, p<0.001$	0.55
NC				
HCR-20 Total	29.41 (SD=4.44)	27.27 (SD=5.60)	$t(611)=1.89, p=0.060$	0.42
Historical	13.73 (SD=2.75)	13.79 (SD=3.03)	$t(611)=-0.10, p=0.920$	-0.02
Clinical	8.20 (SD=1.68)	6.78 (SD=2.34)	$t(28)=4.13, p<0.001$	0.70
Risk-management	7.48 (SD=1.87)	6.72 (SD=2.50)	$t(28)=1.95, p=0.062$	0.34
E				
HCR-20 Total	28.70 (SD=5.79)	27.30 (SD=5.56)	$t(611)=1.18, p=0.238$	0.25
Historical	13.07 (SD=3.92)	13.82 (SD=2.97)	$t(611)=-1.170, p=0.243$	-0.22
Clinical	7.89 (SD=1.91)	6.77 (SD=2.34)	$t(611)=2.27, p=0.023$	0.52
Risk-management	7.74 (SD=2.30)	6.72 (SD=2.48)	$t(611)=1.95, p=0.052$	0.43
TO				
HCR-20 Total	29.73 (SD=4.53)	27.14 (SD=5.61)	$t(611)=3.24, p=0.001$	0.51
Historical	14.44 (SD=2.46)	13.73 (SD=3.06)	$t(611)=1.64, p=0.102$	0.26
Clinical	7.62 (SD=1.86)	6.74 (SD=2.36)	$t(67)=3.20, p=0.002$	0.41
Risk-management	7.67 (SD=2.10)	6.67 (SD=2.50)	$t(611)=2.80, p=0.005$	0.43
NV-ISB				
HCR-20 Total	29.62 (SD=4.68)	27.05 (SD=5.60)	$t(611)=-3.67, p<0.001$	0.50
Historical	14.12 (SD=2.91)	13.75 (SD=3.03)	$t(611)=-0.97, p=0.330$	0.12
Clinical	7.87 (SD=1.78)	6.68 (SD=2.36)	$t(103)=-5.05, p=0.005$	0.57
Risk-	7.63	6.64	$t(611)=-3.15,$	0.42

V-ISB	**management**	(SD=2.22)	(SD=2.49)	*p*=0.002	
	HCR-20 Total	29.57 (SD=4.71)	27.08 (SD=5.60)	*t*(611)=-3.45, *p*<0.001	0.48
	Historical	13.73 (SD=3.11)	13.80 (SD=3.01)	*t*(611)=0.17, *p*=0.866	-0.02
	Clinical	8.10 (SD=1.58)	6.66 (SD=2.36)	*t*(101)=-6.53, *p*<0.001	0.72
	Risk-management	7.72 (SD=1.92)	6.64 (SD=5.52)	*t*(92)=-4.13, *p*<0.001	0.26

VC, verbal comments; NC, non-contact; E, exposure; TO, touching others; NV-ISB, non-violent inappropriate sexual behaviour; V-ISB, violent inappropriate sexual behaviour

Table 2: Rates of SASBA recordings as a function of age and gender

	Age			Gender		
	<40	**≥40**	**Test**	**Male**	**Female**	**Test**
VC	26 (7.4%)	21 (8%)	x^2(1, N=613)=0.07, p=0.798	44 (10.5%)	3 (1.5%)	x^2(1, N=613)=15.17, p<0.001
NC	19 (5.4%)	6 (2.3%)	x^2(1, N=613)=3.80, p=0.051	21 (5%)	4 (2.1%)	x^2(1, N=613)=3.00, p=0.083
E	15 (4.3%)	8 (3%)	x^2(1, N=613)=0.64, p=0.422	15 (3.6%)	8 (4.1%)	x^2(1, N=613)=0.10, p=0.755
TO	39 (11.1%)	13 (4.9%)	x^2(1, N=613)=7.44, p=0.006	38 (9.1%)	14 (7.2%)	x^2(1, N=613)=0.63, p=0.429
NV-ISB	40 (11.4%)	25 (9.5%)	x^2(1, N=613)=0.59, p=0.444	54 (12.9%)	16 (8.2%)	x^2(1, N=613)=2.92, p=0.087
V-ISB	52 (14.9%)	18 (6.8%)	x^2(1, N=613)=9.53, p=0.002	53 (12.7%)	12 (6.2%)	x^2(1, N=613)=5.97, p=0.015

VC, verbal comments; NC, non-contact; E, exposure; TO, touching others; NV-ISB, non-violent inappropriate sexual behaviour; V-ISB, violent inappropriate sexual behaviour

Table 3: Predictive validity of the HCR-20 for SASBA recordings

	HCR-20 Total			Historical			Clinical			Risk-management			SJ		
	AUC	99%CI	*p*	AUC	99%CI	*p*	AUC	99%CI	*p*	AUC	99%CI	*p*	AUC	99%CI	*p*
VC	0.652	[0.546-0.757]	0.001	0.507	[0.393-0.621]	0.872	0.674	[0.584-0.764]	<0.001	0.646	[0.547-0.745]	0.001	0.513	[0.403-0.622]	0.814
NC	0.600	[0.462-0.738	0.090	0.471	[0.319-0.623]	0.627	0.688	[0.561-0.814]	0.001	0.581	[0.458-0.705]	0.168	0.537	[0.370-0.704]	0.616
E	0.590	[0.426-0.755]	0.141	0.451	[0.294-0.609]	0.429	0.640	[0.494-0.785]	0.023	0.622	[0.473-0.771]	0.047	0.488	[0.331-0.645]	0.876
TO	0.642	[0.543-0.741]	0.001	0.563	[0.461-0.664]	0.134	0.606	[0.510-0.702]	0.011	0.612	[0.513-0.710]	0.008	0.512	[0.398-0.626]	0.816
NV-ISB	0.640	[0.552-0.727]	0001	0.529	[0.440-0.618]	0.426	0.649	[0.567-0.731]	<0.001	0.613	[0.524-0.702]	0.002	0.555	[0.455-0.654]	0.240
V-ISB	0.636	[0.545-0.727]	<0.001	0.491	[0.394-0.588]	0.816	0.679	[0.600-0.758]	<0.001	0.623	[0.541-0.706]	0.001	0.488	[0.387-0.589]	0.784

VC, verbal comments; NC, non-contact; E, exposure; TO, touching others; NV-ISB, non-violent inappropriate sexual behaviour; V-ISB, violent inappropriate sexual behaviour