The empirical evidence base for the use of the HCR-20: A narrative review of study designs and transferability of results to clinical practice

Alexander Challinor[bef] AbisolaOgundalu[a] Jason C.McIntyre[d] Vicky Bramwell[a] Rajan Nathan[abcd]

[a] Cheshire and Wirral Partnership NHS Trust, UK

[b] University of Liverpool, UK

[c] University of Chester, UK

[d] Liverpool John Moores University Liverpool, UK

[e] Health Education England North West, Liverpool, UK

[f] Mersey Care NHS Foundation Trust, UK

Abstract

The HCR-20, a widely used method of assessing and managing risk, relies on the structured professional judgement approach. This paper reports a narrative literature review of the HCR-20 studies to explore the applicability of the study results to the use of the HCR-20 in clinical practice. From a literature search using terms "HCR-20" and "HCR 20", 206 papers were included. Of studies using the HCR-20 version 2 ($n = 191$), 92% ($n = 176$) relied on variables based on scores derived by adding item scores, and 50% ($n = 95$) tested the HCR-20 using predictive validity methodology. Of the HCR-20 version 3 studies ($n = 21$), the "presence of risk factors" step was the most commonly examined ($n = 18$, 86%), but 2 of the 7 steps ("scenario planning" and "management") were not examined at all. Amongst those studies whose primary focus was on the HCR-20, 67% ($n = 64/95$) did so by assessing the predictive validity of the tool. Only one employed a design to test whether the use of the HCR-20 affected violence rates. The predominant study design provides support for the use of the HCR-20 as an actuarial tool, and there is limited empirical evidence in support of its effectiveness as a structured professional judgement approach to the assessment and management of the risk of violence.

Keywords

1. Introduction

1.1. Approaches to the assessment of risk to others in forensic mental health services

The term risk is a multifaceted construct, which is often used in mental health services to refer to the likelihood, and the predicted magnitude, of certain types of adverse outcomes. Within forensic mental health services, particular attention is paid to the potential of an adverse outcome in which the patient causes harm to others.

Formalised procedures for assessing the risk of harm to others emerged in light of the research into limitations of unstructured approaches (Scurich, 2016). Given the pragmatic and conceptual complexities of risk, the use of more structured, comprehensive risk decision models was explored. Many of the formal procedures that were developed fell into one of two distinct types of approach: the actuarial model and structured professional judgement (SPJ). The actuarial model of assessing risk involves identifying and quantifying those factors that have been empirically shown to predict a specific harm outcome and combining information about those factors in a pre-defined way to produce a quantitative estimate of the likelihood of that outcome (Scurich, 2016). Direct comparisons can be made between estimates for individual assessments and summary data for defined samples (Singh, Grann, & Fazel, 2011). Actuarial risk assessments tend to rely on static factors (e.g., age at first offence, the seriousness of offending and gender). The problem for the clinical use of static factors is that by definition they are immutable, they tend to lack applicably to real-life scenarios, and they have little rehabilitative value (Doyle & Dolan, 2002; Tully, 2017). Whilst these factors have predictive power, they are less useful to inform intervention plans or to measure the

success of risk-reduction strategies. Despite these criticisms, the greater predictive power of this statistical approach over unstructured clinical assessment has been known for over half a century (Meehl, 1954; Monahan & Skeem, 2014; Steadman et al., 2000). Additionally, it has been shown that the accuracy of actuarial approaches in the prediction of violence has improved in the 30 years between 1970 and 2000, with attempts to include more empirically validated factors that are intended to be clinically useful (i.e., factors that measure change and guide interventions) (Buchanan, 2008; Harris & Hanson, 2010). A notable example of an actuarial risk assessment tool is the Violence Risk Appraisal Guide (Quinsey, Harris, Rice, & Cormier, 1998; Rice, Harris, & Lang, 2013).

Although these tools may be "prediction friendly", they are not necessarily "treatment friendly" (Lewis, Olver, & Wong, 2013). Since the mid-1990s, a new generation of tools were developed to meet what Douglas and Kropp (2002) argued was the purpose of violence risk assessment, risk management and violence prevention. The SPJ approach rejects the goal of quantified prediction and relies on clinical judgement guided by a framework of factors and decision-making (Cheng, Haag, & Olver, 2019). SPJ approaches include static factors, but they also involve examination of dynamic factors, which are potentially modifiable and can be used to quantify progress over time. These risk factors are more easily used to infer appropriate intervention and supervision strategies. Given these characteristics, the SPJ approach has become the preferred method for clinical risk assessment and management (Beazley, Carter, Stewart, & Renton, 2017). Of the SPJ tools, the Historical, Clinical, and Risk Management - 20 (HCR −20) is considered the most researched and the procedure of choice in many forensic mental health services (Judges, Egan, & Broad, 2016; Ramesh, Igoumenou, Vazquez Montes, & Fazel,

2018). The aim of this review is to examine the empirical evidence base for the use of the HCR-20.

1.2. The HCR-20

The HCR-20 was first published in 1993 and then updated in 1995. Due to the need for "periodic revisions" the HCR-20 Version 2 (V2) was published in 1997, (Douglas et al., 2014). The HCR-20 was further revised in 2013 with the publication of the HCR-20 Version 3 (V3) (Douglas et al., 2014). The HCR-20 aimed to provide a narrative categorical approach to risk communication and its widespread adoption is attributable to extensive empirical evaluation in support of this tool. Claims have been made about the content-related validity (based on the adequacy of the literature reviews of the relationship between components of the HCR-20 and violence), inter-rater reliability, concurrent validity (by examining associations between scores from V2 and V3), and predictive validity of the HCR-20. In order to comply with the recommendation that SPJ tools are evaluated "in the way it is recommended they be used" (Heilbrun, Yasuhara, & Shah, 2010), before scrutinising the literature supportive of the HCR-20, one needs to understand how the HCR-20 should be administered. The evidence base presented in favour of the continued use of the HCR-20 relies on studies of the previous, as well as the current, version of the HCR-20 and therefore the nature of both versions need to be understood.

The HCR-20 was developed as a "violence risk assessment scheme" (Douglas, Ogloff, Nicholls, & Grant, 1999) in the form of a "checklist of risk factors for violent behaviour" (Webster et al., 1997a, Webster et al., 1997b). The 20 risk factors were grouped by ten past ("Historical") factors, five present ("Clinical") variables, and five future ("Risk Management") issues. The V2 manual, which served as a "guide to

assessment," explained that coding the HCR-20 required two kinds of judgement. Firstly, there was an item level judgement in which the assessor, who would be familiar with brief guidance notes for each factor, would allocate one of three codes (0, 1, or 2 which correspond to absent, possibly/less seriously present, or present, respectively). The second type of judgement involved the integration of the item-level information to reach a final decision regarding the risk of violence. Although the precise approach was left to the discretion of the assessors, the recommendation was to make a final decision on a three-point scale of low, moderate, or high. However, the manual was clear that for clinical purposes this final decision should not be reached by a numerical operation. The authors stated that "for clinical purposes, it makes little sense to sum the numbers of risk factors in a given case, and then use fixed, arbitrary cut-offs to classify the individual as low, moderate, or high risk" (Webster et al., 1997a, Webster et al., 1997b).

As with actuarial approaches, the HCR-20 comprised a predefined list of risk factors. However, the critical distinction was that the HCR-20 required the assessor to "integrate" information relating to each item rather than applying the actuarial approach of turning the information into scores for each item and then feeding them into a formula to produce a probability estimate (Douglas et al., 2014). Thus, empirical studies relying only on the item codes (whether individually or combined) would not be a valid test of the performance of the overall HCR-20 scheme, at least as the V2 manual recommended for its use in clinical practice. To have direct applicability to the use of the HCR-20 in clinical practice, studies would have to include the key manualised steps of integration and final judgement. When studying the HCR-20 researchers have the option of using numerical (or actuarial) variables

(i.e., subscale and/or total scores), or adopting a non-numerical (or SPJ) method of integration and final risk judgement.

The HCR-20 V2 had more than 200 disseminations based on more than 33,000 cases across 25 countries, becoming the most used violence risk assessment instrument both in terms of assessing risk and creating risk management plans (Douglas et al., 2014; Singh, Grann, & Fazel, 2013). Prior to the introduction of the current HCR-20 V3, the HCR-20 V2 underwent several years of development and revision work. The rationale for the revision was to reflect the contemporary scholarship on violence, to enhance the decision-making process through the increased use of scenario planning and formulation, to develop the link between risk assessment and risk management, and to improve the balance of both idiographic and nomothetic data in meaningful and clinically useful ways (Douglas et al., 2014).

In 2013 the new version of the HCR-20 (V3) was introduced. Although there were some changes in the nature of the individual items, the principle of the first step in the previous version was retained in the current version. However, it was parsed into two constituent steps, namely step 1 (gather information), and step 2 (code presence of risk factors on a scale of certainty). The remainder of the new HCR-20 was changed considerably, and it was now presented not just as a risk assessment scheme, but as an "approach to violence risk assessment and management" which aims to reduce risk (Douglas, Hart, Webster, & Belfrage, 2013).

To enhance the HCR-20's capacity for applicability to the individualised aspects of a case, a step was introduced to consider the relevance of factors in a given case (step 3). Step 4 draws on the principle of integration introduced in the previous version but stipulates the means by which this is done (i.e., using formulation). Step 5 entails the generation of specific risk scenarios and step 6 involves recommending

risk management strategies. The final step (step 7) requires the documentation of summary judgements termed as "final options" or "conclusory opinions" within the HCR-20 manual and within research studies (this paper has used the term "final options" to describe step 7). The final risk judgements are made on three dimensions (future violence/case prioritisation, serious physical harm, and imminent violence), which are coded on a three-point scale of low, moderate, and high. All seven steps in combination make up the HCR-20 V3, which the manual describes as "a comprehensive set of professional guidelines for the assessment and management of violence risk" (Douglas et al., 2013).

The purpose for which the HCR-20 has been designed is "decision-making about violence risk" (Douglas et al., 2013). Studies that set out to evaluate the current version of the HCR-20 in its entirety should consider incorporating each of the seven steps for their results to have validity and clinical utility. However, there may be merit in examining individual aspects of the process, for example, a study that focusses on the consistency between different assessors or clinical teams of "risk management strategies" made for the same patient (or similar types of patients). The change of function of the HCR-20 from a risk assessment scheme to an approach to risk assessment and management also merits a change in methodology of studies. Given that the ultimate aim of the HCR-20 is to guide risk management, it can be conceptualised as an intervention to reduce the likelihood of violence. To determine whether the HCR-20 is effective in achieving this objective it would require the evaluation of a change in the level and extent of the risk.

1.3. Predictive validity and the HCR-20

The most common technique used in the evaluation of risk assessment approaches is predictive validity. In the context of the clinical use of structured violence risk assessment, predictive validity is "the ability of total scores, probabilistic risk bins from actuarial instruments, or categorical risk judgements from SPJ instruments to correctly assess the likelihood of violence" (Singh, 2013).

As set out in the definition above, predictive validity can be used to study SPJ approaches. In particular, it is a way of testing the ability of SPJ categorical risk judgements to correctly assess the likelihood of violence. According to the manualised approach for the HCR-20 V2, the risk judgement arises from the integration of information gathered to code each of the 20 items. As advised by the manual, this should not be based on adding the scores from the items. Therefore, an examination of the ability of the risk judgements derived through a non-numerical integration of the item-level information to predict violent outcomes would be a more valid test of the HCR-20 V2. The use of item scores combined in a numerical way would not only remove the critical step of integration, but also it would contravene the recommended approach for the use of the HCR-20 in clinical settings. It is important to assess whether studies investigating the HCR-20 are combining item scores in a numerical way or categorising the final risk judgements through the non-numerical integration of item-level data.

In the HCR-20 V3, the applicability of the predictive validity methodology is further complicated by the introduction of three types of judgement. The manual's explanation of how to make the first type judgement ("case prioritisation") suggests that it is based on a view about likelihood of future violence and therefore this would seem to represent a variable that could be used for the predictive validity of the HCR-20 as a risk assessment tool. The second type of judgement ("risk for serious

physical harm") does not clearly lend itself to use as a predictor of likelihood, but it could be tested as a predictor of seriousness. The third judgement ("risk for imminent violence") could be taken as a predictor of likelihood but only over a specific time period (i.e., in the near future).

Even if it is accepted that the risk judgements arising from the seven steps of the HCR-20 can be used in predictive validity testing, there remains a question about whether this methodology is the correct one to empirically test whether the HCR-20 achieves its objectives. The stated goals of guiding decision-making about violence risk and risk management are activities that have distinct outcomes from risk assessment. Predictive validity testing would not tell us whether these goals are achieved. As already noted, a methodology to test the effectiveness of an intervention would be more appropriate.

Questions arise not only about whether predictive validity methodology is the correct methodology to empirically evaluate the HCR-20 or other SPJ approaches, but also about whether validity as a notion is one that aligns with the principles of these approaches. The general definition of validity is the extent to which something can be trusted. As well as predictive validity (defined above), other types include content validity (the degree to which the items of an assessment instrument are representative of the entirety of the what the test purports to assess) (Koller, Levenson, & Glück, 2017; Salkind, 2010) and construct validity (the extent to which the measurements are consistent with relevant underlying theory) (Ginty, 2013). Criterion-related validity refers to the degree to which the outcome of a test corresponds to an established representation of the construct and is operationalised as the strength of association between the results obtained by the test and by a

criterion measure (Karras, 1997). Although it is possible to use the HCR-20 as a measure (by isolating a step of the HCR-20, converting the non-numerical codes from that step into scores, and then combining those scores in several ways) this involves moving away significantly from the recommended clinical application of the HCR-20. Even if the focus is just on evaluating the second step (coding the presence of individual factors), caution needs to be exercised in making assumptions about how good that step is at influencing the crucial "taking action" step on the basis of scores derived using a process outside the rules of the HCR-20. Further, when the HCR-20 V3 is taken as a whole, it becomes clear that it is not a measure, it is a decision-making tool. The HCR-20 does not aim to predict future violence but purports to guide the clinician through the risk assessment process and inform decisions around violence risk. This tool for "decision-making about violence risk" guides risk management interventions and plans, with the ultimate objective to reduce the level and extent of risk (Douglas et al., 2013).

There has not previously been a thorough review of the designs employed in the empirical testing of the HCR-20 to determine the extent of research evidence that specifically supports the clinical use of the HCR-20. In particular, the case for the ability of the HCR-20 to reduce risk and violent outcomes has not been convincedly made.

2. Aim of the current review

The broad aim of this review was to explore the nature of the empirical support for the use of the HCR-20 in clinical practice. The objective was to review the designs used in published studies testing the HCR-20 so that the implications of the results of the studies can be understood for the clinical use of the HCR-20. It is acknowledged

that the use of the HCR-20 in clinical practice encompasses many different facets such as assessing risk, formulating risk, developing scenarios and risk management plans, and informing decisions about detention versus liberty. The focus of this review is the empirical grounds for concluding that the HCR-20, a non-actuarial approach, achieves the stated objective of improving the assessment and management of risk of violence in clinical practice.

The specific elements that we aimed to explore in the narrative review included:

-

    The extent of the HCR-20 literature that has a primary focus on the empirical evaluation of the HCR-20.

  -

    The extent to which the component steps of the HRC-20 (i.e., the 2-step process of HCR-20 V2 and the 7-step process of V3) have been examined.

-

    The extent to which the studies of the HCR-20 have adhered to the approach stipulated in the HCR-20 manual.

   -

    The proportion of the empirical literature that has investigated the predictive validity of the HCR-20.

    -

    The extent to which the HCR-20 has been studied as an intervention (i.e., examining the effect of the clinical use of the HCR-20 on the rate of future violence).

3. Methods

3.1. Search

A literature search was carried out on three databases in June 2018: EMBASE, PsychINFO, and CINAHL. The search was performed using the terms "HCR-20" and "HCR 20". The research papers were also searched using snowballing via citation tracking within the databases. The reference lists of relevant papers known to investigators were checked. There was no limit of date or specification of language within the literature search. Our search procedure was performed in consultation with a research librarian (VB).

The search yielded 492 potentially relevant papers from the three databases (see Fig. 1). 154 items were excluded from this stage due to replications and 338 research articles were included for initial assessment for inclusion. If the data required could not be obtained from the abstract, full article papers were assessed for eligibility for inclusion.

Fig. 1. Flow Diagram of the Identification and Selection of Articles in the Review.

3.2. Inclusion and exclusion criteria

Studies were considered to meet inclusion criteria if the abstract obtained was from an empirical study that made reference to the HCR-20 risk assessment tool. Question one of the study checklist tool (Section 3.3 Study Categorisation; Table 1) of the methodology states "was it a research study?". A "research study" was defined in terms of whether there had been testing of a hypothesis with empirical data. Fig.

1 shows the studies that were excluded from the review as they did not meet the criteria of being an empirical study that made reference to the HCR-20 (i.e., "not a research study", as outlined in Fig. 1).

Table 1. Study review checklist items.

1	Was this a research study?
2	Which version of HCR-20 was being used?
3	Was the HCR-20 the main focus of the study?
4	How was the HCR-20 performance assessed (quantitative/qualitative)?
5	Which HCR-20 steps were assessed?
6	Did the study use numerical summary scores?
7	Does the study assess predictive validity of the HCR-20?
8	Is the HCR-20 as an intervention being assessed?

Papers were also excluded if the full paper was not available for those studies where sufficient data could not be obtained from the abstract alone (Fig. 1; "full paper not available", $n = 34$). One study was excluded where this dissemination did not include adequate information about how the tool was used to assess for violence risk (Fig. 1; "Inadequate information in full paper", $n = 1$). Five papers were excluded as the study did not assess the HCR-20 risk assessment tool (Fig. 1; "HCR-20 not assessed", $n = 5$).

A total of 206 papers were included in the study to be further analysed.

3.3. Study categorisation

In order to identify the study design characteristics of relevance to the in-practice application of the HCR-20, a checklist of eight items (table 1), and an accompanying narrative description of each item was developed (by TN and AO). The first item ("was this a research study?") was defined in terms of whether there had been testing of a hypothesis with empirical data. The papers that were not considered to have described a research study were excluded and not analysed further with the checklist. The third checklist item ("was the HCR-20 the main focus of the study?") was based on a judgement about whether the research questions or aims made specific reference to the analysis of the HCR-20. With regard to the fifth item ("which HCR-20 steps were assessed?"), the checklist listed each HCR-20 step (i.e., item presence, relevance, formulation, scenarios, management, or final options for HCR-20 V3; or items or final risk judgement for HCR-20 V2). The sixth item on the study review checklist ("did the study use summary scores?") involved determining whether a score was generated by adding the individual item scores in some way (usually by separately adding the items in each of the three sections of the HCR-20 and/or all 20 items). The sixth item looked at the use of numerical individual scores, subscale scores, and total scores. Additionally, if the paper investigated the "final options" step of the HCR-20 V3, or "final risk judgement" step of the HCR-20 V2 using a non-numerical methodology (i.e., using clinical final risk judgements of high, moderate, and low), this was recorded.

3.4. Data collection

Two reviewers of the papers (AO and AC) both independently applied the checklist to the first 65 research papers and inter-rater reliability (IRR) was statistically

analysed using Cohen's kappa statistic. This measured inter-rater agreement for qualitative categorical items collected in the data. Each Kappa result showed very good agreement (each variable Kappa ≥0.93, *p*-values <0.001) demonstrating excellent IRR (Table 2).

Table 2. Inter-rater reliability of study review checklist results.

Checklist Variable	Kappa Value	P-value	Interrater Reliability Descriptor
1	0.98	< 0.001	Very good agreement
2	0.97	< 0.001	Very good agreement
3	1.00	< 0.001	Very good agreement
4	1.00	< 0.001	Very good agreement
5 (HCR-20 V3))	0.93	< 0.001	Very good agreement
5 (HCR-20 V2))	1.00	< 0.001	Very good agreement
6	1.00	< 0.001	Very good agreement
7	0.99	< 0.001	Very good agreement
8	1.00	< 0.001	Very good agreement

In this stage, the checklist was applied to each of the 206 papers by one or other of the two researchers (AO and AC). The authors examined full copies of each of the 206 papers and applied each question in the study categorisation checklist to the paper. A consensus discussion with TN was held for any disagreements between the

two reviewers, or for any uncertainty with the appraisal of the studies with the checklist.

4. Results

The flow diagram in Fig. 1 shows the number of papers initially identified by searching the three databases, the number after duplicates excluded, the number excluded after screening for eligibility, and the number of papers which were included in the literature review.

4.1. Studies included

Of the 338 studies assessed for eligibility, 132 were excluded. In total, 206 papers met inclusion criteria (see Fig. 1). Of these studies, 197 were published empirical studies, two of which were published audits, and seven papers were unpublished dissertations or theses.

4.2. Version of HCR-20

One hundred and eighty-five studies used the HCR-20 V2 alone while 15 studies used the HCR-20 V3 alone. Six studies used both the HCR-20 V2 and V3.

4.3. HCR-20 as the focus of the study

The HCR-20 was the main focus of the study in 95 of the 206 (46%) studies reviewed.

4.4. Quantitative or qualitative evaluation of HCR-20

Of the 206 studies, a majority (n = 203 / 99%) used quantitative methodology. Of the 3 studies that used qualitative methodology, the HCR-20 was the focus of the study in two (Gough, Richardson, & Weeks, 2015; Travers & Kumar, 2011). Gough et al. (2015) examined the quality of, and service-user involvement in, the HCR-20 assessments completed on two wards. Travers and Kumar (2011) qualitatively compared the professional judgement of a clinician with violent risk assessment tools. The third paper using qualitative methodology, where the HCR-20 was not the main focus of the study, investigated the use and perceived utility of a variety of standardised risk measures across 29 medium secure units (Khiroya, Weaver, & Maden, 2009).

All of the studies evaluating the HCR-20 V3 were done so quantitatively. The three qualitative studies were performed on the HCR-20 V2.

4.5. Evaluation of HCR-20 steps

This section of the results shows *what* steps of the HCR-20 have been analysed in the studies included in the review. The following sections (4.6, 4.7, 4.8) describe *how* the steps of the HCR-20 have been studied (i.e., using scores for validity testing).

The use of the HCR-20 V2 includes an initial "item" coding judgement (in which the assessors allocate one of three codes (0, 1, or 2 which correspond to absent, possibly/less seriously present, or present, respectively), and a second step of integrating the items into one "final risk judgement" (involving the integration of the item-level information to reach a final decision regarding the risk of violence). This review explored what the HCR-20 V2 steps were investigated for all of the 191

studies that used the HCR-20 V2. A majority of the studies of the HCR-20 V2 used a variable based on "items" ($n = 187/191$; 98%). Forty of the HCR-20 V2 studies ($n = 40/191$; 21%) used a variable based on the "final risk judgement".

On further breakdown of steps tested in the HCR-20 V2 studies, one-hundred and fifty ($n = 150/191$; 79%) studies tested the variable "item" alone. Three ($n = 3/191$; 2%) investigated "final risk judgement" alone, and thirty-eight ($n = 38/191$; 20%) analysed both "items" and "final risk judgement" of the HCR-20 V2 in their study.

The structure of the HCR-20 was changed considerably with the introduction of the HCR-20 V3, which uses seven steps (gather information, presence of risk factors, relevance of risk factors, formulations, scenario planning, management, and final options). This review investigated what steps were studied for all 21 papers that tested the HCR-20 V3. Of those studies ($n = 21$), the variable most commonly used was based on "presence of risk factors" ($n = 18/21$; 86%) followed in descending order by "final options" ($n = 10/21$; 48%), "relevance of risk factors" ($n = 4/21$; 19%) and formulation (n = 1/21; 0.5%). No studies in this review used variables based on the "scenario planning" or "management" steps of the HCR-20 V3.

Further breakdown of the steps used in the analysis of HCR-20 V3 shows that ten (n = 10/21; 47%) studies only analysed the variable "presence of risk factors". Two studies (n = 2/21; 10%) only investigated the step "final options". Four (n = 4/21; 19%) studies tested both "presence of risk factors" and "final options" steps within their study. Four (n = 4/21; 19%) papers analysed three steps in their study, namely "item presence", "item relevance", and "final options". Finally, one (n = 1/21; 5%) study analysed "formulation" (Hopton et al., 2018). This study analysed the quality of the HCR-20 formulations using a quantitative scale (Hopton et al., 2018). No studies investigated the "scenario planning" or "management" steps of the HCR-20 V3.

4.6. Use of scores / summary scores

This section refers to the use of item-level scoring and of "summary scores", where a score (numerical value) was generated by either adding "coding items" with the HCR-20 V2, or, scoring via the additional step of the conversion of non-numerical data into scores with the HCR-20 V3. It may appear somewhat artificial to suspend scores in isolation from the variables with which they may be associated (such as violent incidents), but this will be explored in the next section on predictive validity. The focus in this section is on whether or not the summary scores have been used in analyses. This is important because it would be outside the clinical rules for the use of the HCR-20 V2 and V3, which advise that judgements for risk assessment and management should not be based on the numerical integration of item-level information. The use of item-level scoring and of "summary scores" removes the step of non-numerical item integration.

One hundred and seventy-six out of one-hundred and ninety-one ($n = 176/191$; 92%) studies assessed the HCR-20 V2 using a score. Nineteen out of the twenty-one ($n = 19/21$; 90%) studies involving the HCR-20 V3 used a score. This score was utilised in different ways within the research, involving specific item-level risk factor scores, grouped subscale summative scores (subscale categories of HCR-20; historical, clinical, risk management), or a total score.

Within the HCR-20 V2 "final risk judgements" and the HCR-20 V3 "final options", risk is categorised on a scale of high, moderate, and low. This review found that studies used two methodologies, either alone, or in combination, to assess this step of the HCR-20. One method involved the use of a "total score", whereby numerical item-level codes were added together. This "total score" was then used to form a final

judgement ("total score" categorised as high, moderate, low). The "total score" was also used in combination with a non-numerical summary risk judgement. The "final risk judgement" (the HCR-20 V2) or "final options" (the HCR-20 V3) step was also studied via the non-numerical categorisation of risk into high, moderate, or low. This method, if used in isolation, is completed as it is intended by the clinical use of the HCR-20 and does not use a "summary score".

In studies investigating the HCR-20 V3, nine out of the ten studies (n = 9/10; 90%) investigating the "final options" also involved the use of numerical item-level scores and a total score. This review found that these nine studies used a combination of numerical total scores, and a non-numerical summary judgement (high, moderate, low). One study by Cawood (2017) investigated the interrater reliability and predictive validity of the HCR-20 V3 in common workplace settings using only non-numerical summary judgements of high, moderate, and low. We found only one study (Cook et al., 2018) that incorporated the HCR-20 V3 recommended three types of final judgements (case prioritisation, risk for serious physical harm, and risk for imminent violence). This study analysed the concurrent validity of the Hamilton Anatomy of Risk Management – Forensic Version SPJ Tool (Cook et al., 2018). For those studies assessing the HCR-20 V2, our review found three studies (n = 3) that investigated the "final risk judgement" alone without employing a numerical score via the use of the "items" step. Keulen-de Vos et al. (2017) examined the relationship between "schema modes" in personality disorders and violence risk (assessed using the "final risk judgement" of high, moderate, and low in the HCR-20 V2). A paper by Seidel and Kilgus (2014) compared psychiatrist's evaluations of accident and emergency department patients made via telepsychiatry or face-to-face, whereby the HCR-20 V2 "final risk judgements" were analysed for agreements

of "dangerousness" (Seidel & Kilgus, 2014). The third study analysed the quality of the HCR-20 V3 formulations and the HCR-20 V2 "final risk judgements" using a quantitative scale (Hopton et al., 2018).

Thirty-eight ($n = 38/191$; 20%) studies assessing the HCR-20 V2 did so by analysing both "items" and "final risk judgements". Thirty studies ($n = 30/38$; 79%) testing both these steps of the HCR-20 V2 did so by using a non-numerical categorisation of risk (high, moderate, low), and also by using scores with the use of a "summary score" in the study. Four papers ($n = 4/38$; 11%) tested the "final risk judgement" using only a numerical score (with conversion of score into categories of high, moderate, low). (Archibald, Campbell, & Ambrose, 2014; Gray et al., 2004; Smith & White, 2007; Vogel & De Ruiter, 2004). The remaining four studies did not involve scores or a non-numerical categorisation of risk. Three of these studies were qualitative and are described in section 4.4 (Gough et al., 2015; Khiroya et al., 2009; Travers & Kumar, 2011). The fourth study is an audit of the quality and competency of the HCR-20 V2 assessments (Sen, Lindsey, Chatterjee, Rama-Iyer, & Picchioni, 2015).

4.7. Predictive validity

When the HCR-20 was the focus of the study, the majority ($n = 64$; 67%) of the studies assessed its predictive validity.

Ninety-five ($n = 95$; 50%) of the studies using the HCR-20 V2 assessed the predictive validity of the tool. All of these papers did so via the use of a score.

Twenty-four ($n = 24/95$; 25%) tested the predictive validity of the HCR-20 V2 using a combination of scores and non-numerical categorisation of the "final risk judgement".

Ten of the 21 studies using V3 (48%) looked at the predictive validity of the HCR-20. Of those studies, that investigated predictive validity, nine studies did so using

scores. One study analysed predictive validity using summary ratings in the HCR-20 V3 "final options" without the use of a numerical score (Cawood, 2017). Five of the ten studies (n = 5/10; 50%) did so by using numerical scores and summary ratings in the HCR-20 V3 "final options".

4.8. HCR-20 as an intervention

This section refers to whether the HCR-20 has been tested as an intervention; a decision-making tool aimed at reducing violence. Only one study assessed the HCR-20 as an intervention tool (Belfrage, Fransson, & Strand, 2004). Empirical testing of the HCR-20 as an intervention mirrors its clinical applicability as a decision-making tool in the reduction of violence. The researchers compared the number of violent incidents enacted by residents (n = 47) on a wing of a maximum-security prison before and after the in-vivo introduction of the HCR-20 (i.e., training of staff in risk assessment and the HCR-20, and assessment of the patients using the HCR-20, followed by team generated risk management strategies). Following the introduction of the HCR-20 there was a statistically significant reduction in the mean number of violent incidents (14/year versus 5/year).

5. Discussion

This review confirms that the HCR-20 has been subject to extensive empirical testing. Most studies examining the performance of the HCR-20 have created numerical variables in the form of summated scores transposed from the HCR-20 codes. The predominant methodology has involved the assessment of validity, particularly predictive validity. This has been done by examining the ability of summary scores (total or the summary of the historical, clinical, or risk items) to

predict violent outcomes. Positive findings from studies of the HCR-20 using predictive validity methodology provide evidence in favour of the summary scores in predicting violent outcomes. In addition to the use of summary scores, studies have used the non-numerical clinically generated categorisation of high, moderate, and low final risk judgements. The use of this methodology has become more common over time and therefore is seen in more in studies of the HCR-20 V3 in comparison to the HCR-20 V2. However, in the majority of studies, this non-numerical method is usually accompanied by the use of a summary score.

The intended use of the HCR-20, as described in the manual, was that it is not used in an algorithmic fashion. This review shows that as the literature base expanded, along with the re-design and emergence of the HCR-20 V3, the use of non-numerical clinically based risk judgements has become more prevalent. The increase in the evaluation of the HCR-20 using non-numerical final risk judgements is important because the methodological approach correlates with how the tool is intended to be used in clinical practice. As an SPJ tool, empirical studies relying only on "scores" from "item" codes (whether individually or combined) would not be a clinically valid test of the overall HCR-20 scheme. Thus, studies using non-numerical risk judgements have more direct applicability to the use of the HCR-20 in clinical practice. Despite the increase in studies evaluating the HCR-20 in this manner (through the use of non-numerical final risk judgements), our review shows that the use of item-level scoring has remained prominent throughout the literature. Additionally, "scores" have been retained as a component in its overall evaluation even when non-numerical final risk judgements are used. Evaluating the HCR-20 via the use of item-level scoring contravenes the recommended clinical approach of this SPJ instrument.

A summary score derived from the individual HCR-20 in itself does not tell the assessor sufficient information about risk in that case. However, if contextualised within an estimate of central tendency (i.e., means and medians) derived from groups of individuals with similar characteristics to the patient, an individual score can be used as an index of likelihood. If comparable group-based data which aligns scores with likelihood is available, then the clinician could use the score to assign a case with a numerical probability that there will be a violent outcome within a defined period.

The problems inherent in using individual probability estimates in forensic clinical decision-making have been widely reported (Clark & Natarajan, 2011). In brief, these problems include the inability (or limited ability) of a numerical probability estimate to (i) reflect the nature of the risk, (ii) highlight the circumstances in which the risk may change, (iii) identify the areas that if addressed could reduce risk, and (iv) accurately reflect the risk likelihood for an individual outside the groups for which there is normative data. These are areas that clinicians need to consider when making decisions in practice.

While the SPJ approach arose as a reaction against the reported limitations of actuarial risk assessment tools (Hockenhull et al., 2012), the findings of this review indicate that the evidence base cited in favour of a commonly used SPJ approach, the HCR-20, is based on a methodology that would primarily support its use as an actuarial tool. Furthermore, whereas the HCR-20 V3 manual advises that "evaluators should not attempt to quantify and combine algorithmically judgements of risk made using the HCR-20 V3" (Douglas et al., 2013), this review shows that studies evaluating the HCR-20 have relied very heavily on such a numerical approach.

The use of clinically generated final risk judgements as described by the HCR-20 manual does partially address this concern. However, there remains a theoretical issue about whether the predictive validity methodologies used in the HCR-20 studies provide any support for the use of the HCR-20 as a SPJ tool. In general terms, positive results from studies of the HCR-20 using this methodology can be taken as support for inclusion of the items that form part of the HCR-20. It would be difficult to justify including a factor in assessing and managing risk if there was no demonstrable association between the factor and the risk outcome. Even this limited support for the chosen items must be qualified when seen in light of the study methodologies. The results of the studies considered in this review do not allow a conclusion to be drawn about a specific association between each factor and a violent outcome. Rather, there is an association between certain combinations of the factors (i.e., in their totality or grouped according to their subcategorization under the headings historical, clinical, and risk management) and the violent outcome. Risk assessment and management emphasises the prevention and reduction of violence risk. The predictive validity design does not allow us to truly assess the SPJ process, where the goal is to prevent violence and manage risk, rather than predict it.

The limitations of research on violence risk tools are not unique to the HCR-20, with very few studies assessing violence risk instruments as an intervention that may reduce risk. Notable exceptions include the study by Abderhalden et al. (2008) that performed a randomised controlled trial (RCT) using the Brøset Violence Checklist, a short term 6-item actuarial tool, and van de Sande et al. (2011) who performed an extension of that by Abderhalden et al. (2008), using the Brøset Violence Checklist and the dangerousness scale. Viljoen, Cochrane, & Jonnson (2018) performed a comprehensive systematic review and narrative synthesis to evaluate whether there

was evidence that risk assessment tools help manage the risk of violence. The results of this review are consistent with the findings from our narrative review that (i) research on the clinical utility of tools is scarce in comparison to research on the predictive validity of the tools, (ii) studies evaluating risk assessment tools did not do so by directly comparing professionals' risk management practices with and without a tool, and (iii) it is not clear whether these tools were being used in research studies as they were *"properly"* intended to be used in clinical practice (Viljoen, Cochrane, & Jonnson, 2018). Our review answers this question, demonstrating that the research on the HCR-20 has predominantly used an approach that is not recommended for its use in clinical practice. The authors conclude that "greater attention to how tools are applied to guide real-world decisions" are needed, such as evaluating pathways between assessment and management of risk and developing strategies that result in risk assessment translating into better risk management outcomes (Viljoen, Cochrane, & Jonnson, 2018).

To achieve an improved balance between risk assessment, management, and reduction, Penney (2021) has proposed a re-conceptualisation of how the validity of violence risk assessment tools is measured. Drawing on key lessons from analogous concepts in risk assessment and management within the aviation industry, Penney (2021) argued that a shift in focus will improve our understanding of how risk assessment tools are used to guide real-world decision making, and in turn, how these decision-making tools can actually reduce the level and extent of violence. A stronger emphasis on the steps of risk formulation and management in risk assessment tools, rather than prediction, may help the validity of risk factors to be re-conceptualised in terms of their ability to forge effective risk management plans, that can measurably reduce risk and rates of violent outcomes (Penney, 2021).

We think that approaches that may prove valuable in testing the effectiveness and applicability of risk assessment and management tools are those used in the evaluation of complex interventions and clinical guidelines. The design of the HCR-20 V3 underwent several revisions with the creators setting out clear guiding principles and goals for the process. This has been done in part to the HCR-20 (through an iterative process of revisions underpinned by guiding principles and goals aimed at identifying and incorporating existing evidence and theory into the instrument, and improving acceptability, delivery, and feasibility of the tool for real-world clinical settings) (Craig et al., 2008; Douglas et al., 2014; Moore et al., 2015). The guidance on the evaluation of complex interventions warrants further attention when investigating the effectiveness of SPJ tools as an intervention.

Of the approaches for evaluating effectiveness of interventions, the RCT is the gold standard (Moore et al., 2015). This remains the case for complex interventions (Hawe, Shiell, & Riley, 2004). Given that the HCR-20 is a tool that enables the clinician to reach a risk judgement to guide management to reduce harm, it can be conceptualised as an intervention. The difficulties in carrying out a RCT with SPJ tools are well-known and have been documented in the literature (Douglas & Kropp, 2002). To date, there has been one grade 1 RCT evidence for the effectiveness of a SPJ tool, the Short-Term Assessment of Risk and Treatability (START), in reducing violence, resulting in a negative outcome (Troquete et al., 2013a, Troquete et al., 2013b). From our review, only one study assessed the HCR-20 as an intervention with a naturalistic design that utilised the HCR-20 as is intended for use in clinical practice (Belfrage et al., 2004). This research demonstrated a statistically significant reduction in the mean number of violent incidents within the highly controlled environment of a high-security prison (Belfrage et al., 2004).

Where conducting an RCT is difficult or unfeasible researchers may undertake studies of the performance of risk assessment tools in real-word settings (Fazel & Wolf, 2018). Studies have examined the field validity of the HCR-20, that is, its relevance to actual assessment of clinical risk, by testing IRR between clinicians and researchers, alongside the predictive validity of clinical ratings (Jeandarme, Pouls, De Laender, Oei, & Bogaerts, 2017). The findings of these studies have called into question the utility of the HCR-20 in clinical settings demonstrating that the HCR-20 may be of limited benefit as a predictive tool in real-world forensic practice (Jeandarme et al., 2017; Neal, Miller, & Shealy, 2015; Vojt, Thomson, & Marshall, 2013).

Guidance for the evaluation of an intervention argues for a systematic approach, drawing on clear descriptions of the risk intervention tools theory and to identify key questions for process evaluation (Craig et al., 2008). Process evaluation would capture whether the intervention is inherently faulty (failure of the interventions theory or concept), and those that are not delivered as intended (implementation failure) (Oakley, Strange, Bonell, Allen, & Stephenson, 2006; Rychetnik, Frommer, Hawe, & Shiell, 2002). The authors believe it would be worthwhile to apply these principles to the evaluation of the effectiveness of the HCR-20.

6. Conclusion

In conclusion, this review found that most studies tested the HCR-20 as if it were an actuarial measure of violence. One study used a methodology for testing the effectiveness of the HCR-20 as an intervention and did so on participants who remained incarcerated during the study. The findings of this review suggests that there is no research to date that shows that the HCR-20 leads to more effective risk

management. The ultimate aim in the validation of a violence risk assessment is to demonstrate that it reduces or prevents violence as a result of using the tool. This would be by demonstrating efficacy in the way in which the tool was intended for use in clinical practice and by judging its utility on its its ability to contribute to harm reduction.

The authors are not advocating a return to unstructured clinical risk assessments, only that practitioners, researchers, and policymakers are more aware of the empirical foundation of SPJ tools, and to guide future research to approaches that test the effectiveness of existing and future tools in clinical practice. The development of risk assessment instruments, the identification of risk factors, and the statistical predictive value of these factors within risk assessment tools clearly have informative value, even if the methodology used to reach this are somewhat fundamentally limited. The HCR-20 can inform a prevention-based model, highlighting changes in risk factors that are important for constructing and revising risk management plans. However, we must practice epistemic humility and strive to conduct research with robust study designs that incorporate an accurate representation of the complexity of risk assessment and management in clinical practice. We recommend that further empirical support for the HCR-20 should be derived from studies that do not convert judgements into scores and that employ study designs that would test the effectiveness of interventions or the evaluation of decision-making tools. The opening words of the HCR-20 Version 1 (Webster, Eaves, Douglas, & Wintrup, 1995) and the HCR-20 V2 (Webster et al., 1997a, Webster et al., 1997b) were that the "challenge in what remains is to integrate the almost separate words of research on the prediction of violence and the

clinical practice of assessment" (Douglas et al., 2014). This review shows that despite progress, this challenge still remains.

References

Abderhalden et al., 2008

C. Abderhalden, I. Needham, T. Dassen, R. Halfens, H.-J. Haug, J.E. Fischer

Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial

British Journal of Psychiatry, 193 (1) (2008), pp. 44-50, 10.1192/bjp.bp.107.045534

CrossRefView Record in ScopusGoogle Scholar

Archibald et al., 2014

S.-J. Archibald, C. Campbell, D. Ambrose

Prediction of treatment outcomes for personality disordered offenders

Journal of Forensic Practice, 16 (4) (2014), pp. 281-294, 10.1108/JFP-05-2013-0037

View Record in ScopusGoogle Scholar

Beazley et al., 2017

P. Beazley, A. Carter, S. Stewart, J. Renton

Practicalities of HCR-20 implementation within secure psychiatric services

Journal of Psychiatric Intensive Care, 13 (2017), pp. 47-56

CrossRefView Record in ScopusGoogle Scholar

Belfrage et al., 2004

H. Belfrage, G. Fransson, S. Strand

Management of violent behaviour in the correctional system using qualified risk assessments

Legal and Criminological Psychology, 9 (1) (2004), pp. 11-22, 10.1348/135532504322776825

View Record in ScopusGoogle Scholar

Buchanan, 2008

A. Buchanan

Risk of violence by psychiatric patients: beyond the "actuarial versus clinical" assessment debate

Psychiatric Services, 59 (2) (2008), pp. 184-190, 10.1176/ps.2008.59.2.184

View Record in ScopusGoogle Scholar

Cawood, 2017

J.S. Cawood

The interrater reliability and predictive validity of the HCR-20V3 in common workplace environments

Journal of Threat Assessment and Management, 4 (1) (2017), pp. 1-11, 10.1037/tam0000071

CrossRefView Record in ScopusGoogle Scholar

Cheng et al., 2019

J. Cheng, A.M. Haag, M.E. Olver

Predictors of historical clinical risk management-20 version 3 (HCR-20:V3) summary risk ratings

Psychiatry, Psychology, and Law, 26 (4) (2019), pp. 682-692, 10.1080/13218719.2019.1618753

CrossRefView Record in ScopusGoogle Scholar

Clark and Natarajan, 2011

T. Clark, M. Natarajan

Risk of Violence Assessment

CRC Press (2011)

Google Scholar

Cook et al., 2018

A.N. Cook, H.M. Moulden, M. Mamak, S. Lalani, K. Messina, G. Chaimowitz

Validating the hamilton anatomy of risk management–forensic version and the aggressive incidents scale

Assessment, 25 (4) (2018), pp. 432-445, 10.1177/1073191116653828

CrossRefView Record in ScopusGoogle Scholar

Craig et al., 2008

P. Craig, P. Dieppe, S. Macintyre, S. Michie, I. Nazareth, M. Petticrew

Developing and evaluating complex interventions: the new medical research council guidance

BMJ: British Medical Journal, 337 (2008), p. 1655, 10.1136/bmj.a1655

Google Scholar

Douglas et al., 2013

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). HCR-20v3

Assessing risk for violence. Mental Health, Law and Policy Institute.

Google Scholar

Douglas et al., 2014

K.S. Douglas, S.D. Hart, C.D. Webster, H. Belfrage, L.S. Guy, C.M. Wilson

Historical-clinical-risk management-20, version 3 (HCR-20V3): development and overview

International Journal of Forensic Mental Health, 13 (2) (2014), pp. 93-108, 10.1080/14999013.2014.906519

CrossRefView Record in ScopusGoogle Scholar

Douglas and Kropp, 2002

K.S. Douglas, P.R. Kropp

A prevention-based paradigm for violence risk assessment: clinical and research applications

Criminal Justice and Behavior, 29 (5) (2002), pp. 617-658, 10.1177/009385402236735

View Record in ScopusGoogle Scholar

Douglas et al., 1999

K.S. Douglas, J.R. Ogloff, T.L. Nicholls, I. Grant

Assessing risk for violence among psychiatric patients: the HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening Version

Journal of Consulting and Clinical Psychology, 67 (6) (1999), pp. 917-930, 10.1037//0022-006x.67.6.917

View Record in ScopusGoogle Scholar

M. Doyle, M. Dolan

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 (6) (2002), pp. 649-657, 10.1046/j.1365-2850.2002.00535.x

View Record in ScopusGoogle Scholar

Fazel and Wolf, 2018S. Fazel, A. Wolf

Selecting a risk assessment tool to use in practice:a 10-point guide

Evidence-Based Mental Health, 21 (2) (2018), pp. 41-43, 10.1136/eb-2017-102861

CrossRefView Record in ScopusGoogle Scholar

Ginty, 2013

A.T. Ginty

Construct validity

M.D. Gellman, J.R. Turner (Eds.), *Encyclopedia of Behavioral Medicine* (pp. 487–487), Springer New York (2013), 10.1007/978-1-4419-1005-9_861

Google Scholar

Gough et al., 2015

K. Gough, C. Richardson, H. Weeks

An audit of service-user involvement and quality of HCR-20 version 2 risk assessments on rehabilitation and low secure wards

Journal of Psychiatric Intensive Care, 11 (S1) (2015), 10.1017/S1742646415000084 e2, Article e2

Google Scholar

Gray et al., 2004

N.S. Gray, R.J. Snowden, S. MacCulloch, H. Phillips, J. Taylor, M.J. MacCulloch

Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: a comparative study of HCR-20, PCL:SV, and OGRS

Journal of Consulting and Clinical Psychology, 72 (3) (2004), pp. 523-530, 10.1037/0022-006X.72.3.523

View Record in ScopusGoogle Scholar

Harris and Hanson, 2010

A.J.R. Harris, R.K. Hanson

Clinical, actuarial and dynamic risk assessment of sexual offenders: why do things keep changing?

Journal of Sexual Aggression, 16 (3) (2010), pp. 296-310, 10.1080/13552600.2010.494772

CrossRefView Record in ScopusGoogle Scholar

Hawe et al., 2004

P. Hawe, A. Shiell, T. Riley

Complex interventions: how "out of control" can a randomised controlled trial be?

BMJ: Britiwh Medical Journal, 328 (7455) (2004), pp. 1561-

1563, 10.1136/bmj.328.7455.1561

View Record in ScopusGoogle Scholar

Heilbrun et al., 2010

K. Heilbrun, K. Yasuhara, S. Shah

R.K. Otto, K.S. Douglas (Eds.), Violence risk assessment tools (2010)

Google Scholar

Hockenhull et al., 2012

J.C. Hockenhull, R. Whittington, M. Leitner, W. Barr, J. McGuire, M.G. Cherry, ..., R.

Dickson

A systematic review of prevention and intervention strategies for populations

at high risk of engaging in violent behaviour: update 2002-8

Health Technology Assessment (Winchester, England), 16 (3) (2012),

p. 1, 10.3310/hta16030

Google Scholar

Hopton et al., 2018

J. Hopton, A. Cree, S. Thompson, R. Jones, R. Jones

An evaluation of the quality of HCR-20 risk formulations: a comparison

between HCR-20 version 2 and HCR-20 version 3

International Journal of Forensic Mental Health, 17 (2) (2018), pp. 195-

201, 10.1080/14999013.2018.1460424

View Record in ScopusGoogle Scholar

Jeandarme et al., 2017

I. Jeandarme, C. Pouls, J. De Laender, T.I. Oei, S. Bogaerts

Field validity of the HCR-20 in forensic medium security units in Flanders

Psychology, Crime & Law, 23 (4) (2017), pp. 305-

322, 10.1080/1068316X.2016.1258467

View Record in ScopusGoogle Scholar

Judges et al., 2016

R. Judges, V. Egan, G. Broad

A critique of the historical clinical risk–20, version 3, risk assessment

instrument

Journal of Forensic Psychology Practice, 16 (4) (2016), pp. 304-

320, 10.1080/15228932.2016.1196102

CrossRefView Record in ScopusGoogle Scholar

Karras, 1997

D.J. Karras

Statistical methodology: II. Reliability and validity assessment in study design,

Part B

Academic Emergency Medicine, 4 (2) (1997), pp. 144-147, 10.1111/j.1553-

2712.1997.tb03723.x

View Record in ScopusGoogle Scholar

Keulen-de Vos et al., 2017

M. Keulen-de Vos, D.P. Bernstein, L.A. Clark, V. de

Vogel, S. Bogaerts, M. Slaats, A. Arntz

Validation of the schema mode concept in personality disordered offenders

Legal and Criminological Psychology, 22 (2) (2017), pp. 420-441, 10.1111/lcrp.12109

CrossRefView Record in ScopusGoogle Scholar

Khiroya et al., 2009

R. Khiroya, T. Weaver, T. Maden

Use and perceived utility of structured violence risk assessments in English medium secure forensic units

Psychiatric Bulletin, 33 (4) (2009), pp. 129-132, 10.1192/pb.bp.108.019810

CrossRefView Record in ScopusGoogle Scholar

Koller et al., 2017

I. Koller, M.R. Levenson, J. Glück

What do you think you are measuring? a mixed-methods procedure for assessing the content validity of test items and theory-based scaling

Frontiers in Psychology, 8 (2017), p. 126, 10.3389/fpsyg.2017.00126

View Record in ScopusGoogle Scholar

Lewis et al., 2013

K. Lewis, M.E. Olver, S.C. Wong

The violence risk scale: predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits

Assessment, 20 (2) (2013), pp. 150-164, 10.1177/1073191112441242

CrossRefView Record in ScopusGoogle Scholar

Meehl, 1954

P.E. Meehl

Clinical versus statistical prediction: A theoretical analysis and a review of the evidence

(1954)

Google Scholar

Monahan and Skeem, 2014

J. Monahan, J.L. Skeem

The evolution of violence risk assessment

CNS Spectrums, 19 (5) (2014), pp. 419-424, 10.1017/S1092852914000145

CrossRefGoogle Scholar

Moore et al., 2015

G.F. Moore, S. Audrey, M. Barker, L. Bond, C. Bonell, W. Hardeman, ..., J. Baird

Process evaluation of complex interventions: Medical Research Council guidance

BMJ: British Medical Journal, 350 (2015), p. h1258, 10.1136/bmj.h1258

CrossRefView Record in ScopusGoogle Scholar

Neal et al., 2015

T.M. Neal, S.L. Miller, R.C. Shealy

A field study of a comprehensive violence risk assessment battery

Criminal Justice and Behavior, 42 (9) (2015), pp. 952-968

CrossRefView Record in ScopusGoogle Scholar

Oakley et al., 2006

A. Oakley, V. Strange, C. Bonell, E. Allen, J. Stephenson

Process evaluation in randomised controlled trials of complex interventions

BMJ: British Medical Journal, 332 (7538) (2006), pp. 413-416, 10.1136/bmj.332.7538.413

CrossRefView Record in ScopusGoogle Scholar

Penney, 2021

S.R. Penney

Innovations in violence risk assessment: What aviation can teach us about assessing and managing risk for rare and serious outcomes

International Journal of Law and Psychiatry, 77 (2021), p. 101710

Advance online publication

https://doi.org/10.1016/j.ijlp.2021.101710

ArticleDownload PDFView Record in ScopusGoogle Scholar

Quinsey et al., 1998

V.L. Quinsey, G.T. Harris, M.E. Rice, C.A. Cormier

Mentally disordered offenders

Violent offenders: Appraising and managing risk (1998)

Google Scholar

Ramesh et al., 2018

T. Ramesh, A. Igoumenou, M. Vazquez Montes, S. Fazel

Use of risk assessment instruments to predict violence in forensic psychiatric hospitals: a systematic review and meta-analysis

European Psychiatry, 52 (2018), pp. 47-53, 10.1016/j.eurpsy.2018.02.007

ArticleDownload PDFCrossRefView Record in ScopusGoogle Scholar

Rice et al., 2013

M.E. Rice, G.T. Harris, C. Lang

Validation of and revision to the VRAG and SORAG: the violence risk appraisal guide—revised (VRAG-R)

Psychological Assessment, 25 (3) (2013), pp. 951-965, 10.1037/a0032878

CrossRefView Record in ScopusGoogle Scholar

Rychetnik et al., 2002

L. Rychetnik, M. Frommer, P. Hawe, A. Shiell

Criteria for evaluating evidence on public health interventions

Journal of Epidemiology & Community Health, 56 (2) (2002), pp. 119-127

View Record in ScopusGoogle Scholar

Salkind, 2010

N. Salkind

Research Design. In *Encyclopaedia of Research Design*

SAGE Publications, Inc (2010), 10.4135/9781412961288

Google Scholar

van de Sande et al., 2011

R. van de

Sande, H.L.I. Nijman, E.O. Noorthoorn, A.I. Wierdsma, E. Hellendoorn, C. van der

Staak, C.L. Mulder

Aggression and seclusion on acute psychiatric wards: effect of short-term risk

assessment

British Journal of Psychiatry, 199 (6) (2011), pp. 473-

478, 10.1192/bjp.bp.111.095141

CrossRefView Record in ScopusGoogle Scholar

Scurich, 2016

N. Scurich

An introduction to the assessment of violence risk

International perspectives on violence risk assessment, Oxford University

Press (2016), pp. 3-15, 10.1093/acprof:oso/9780199386291.003.0001

CrossRefView Record in ScopusGoogle Scholar

Seidel and Kilgus, 2014

R.W. Seidel, M.D. Kilgus

Agreement between telepsychiatry assessment and face-to-face assessment for Emergency Department psychiatry patients

Journal of Telemedicine and Telecare, 20 (2) (2014), pp. 59-62, 10.1177/1357633x13519902

View Record in ScopusGoogle Scholar

Sen et al., 2015

P. Sen, S. Lindsey, N. Chatterjee, R. Rama-Iyer, M. Picchioni

An audit of the quality of HCR-20 violence risk assessments in a low secure service

Journal of Psychiatric Intensive Care, 11 (S1) (2015), 10.1017/S1742646415000096 e3, Article e3

Google Scholar

Singh, 2013

J.P. Singh

Predictive validity performance indicators in violence risk assessment: a methodological primer

Behavioral Sciences & the Law, 31 (1) (2013), pp. 8-22, 10.1002/bsl.2052

CrossRefView Record in ScopusGoogle Scholar

Singh et al., 2011

J.P. Singh, M. Grann, S. Fazel

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) (2011), pp. 499-513, 10.1016/j.cpr.2010.11.009

ArticleDownload PDFView Record in ScopusGoogle Scholar

Singh et al., 2013

J.P. Singh, M. Grann, S. Fazel

Authorship bias in violence risk assessment? a systematic review and meta-analysis

PLoS One, 8 (9) (2013), Article e72484, 10.1371/journal.pone.0072484

CrossRefView Record in ScopusGoogle Scholar

Smith and White, 2007

H. Smith, T. White

Feasibility of a structured risk assessment tool in general adult psychiatry admissions

Psychiatric Bulletin, 31 (11) (2007), pp. 418-420, 10.1192/pb.bp.106.012518

CrossRefView Record in ScopusGoogle Scholar

Steadman et al., 2000

H.J. Steadman, E. Silver, J. Monahan, P.S. Appelbaum, P.C. Robbins, E.P. Mulvey, ..., S. Banks

A classification tree approach to the development of actuarial violence risk assessment tools

Law and Human Behavior, 24 (1) (2000), pp. 83-100, 10.1023/a:1005478820425

View Record in ScopusGoogle Scholar

Travers and Kumar, 2011

R. Travers, A. Kumar

P02-192 - Structured clinical decision making: Is it different ?

European Psychiatry, 26 (2011), p. 788, 10.1016/S0924-9338(11)72493-2

ArticleDownload PDFCrossRefView Record in ScopusGoogle Scholar

Troquete et al., 2013a

N. Troquete, R.H.S. van den Brink, H. Beintema, T. Mulder, T.W.D.P. van Os, R.A. Schoevers, D. Wiersma

Authors' reply

British Journal of Psychiatry, 202 (6) (2013), pp. 468-469, 10.1192/bjp.202.6.468a

CrossRefView Record in ScopusGoogle Scholar

Troquete et al., 2013b

N.A.C. Troquete, R.H.S. van den Brink, H. Beintema, T. Mulder, T.W.D.P. van Os, R.A. Schoevers, D. Wiersma

Risk assessment and shared care planning in out-patient forensic psychiatry: cluster randomised controlled trial

British Journal of Psychiatry, 202 (5) (2013), pp. 365-371, 10.1192/bjp.bp.112.113043

CrossRefView Record in ScopusGoogle Scholar

Tully, 2017

J. Tully

HCR-20 shows poor field validity in clinical forensic psychiatry settings

Evidence Based Mental Health, 20 (3) (2017), p. 95, 10.1136/eb-2017-102745

CrossRefView Record in ScopusGoogle Scholar

Viljoen, Cochrane, & Jonnson, 2018

J.L. Viljoen, D.M. Cochrane, M.R. Jonnson

Do risk assessment tools help manage and reduce risk of violence and reoffending? A systematic review.

Law and Human Behavior, 42 (3) (2018), pp. 181-214, 10.1037/lhb0000280

CrossRefView Record in ScopusGoogle Scholar

Vogel and De Ruiter, 2004

V. Vogel, C. De Ruiter

Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients

The Journal of Forensic Psychiatry, 15 (2004), pp. 145-164, 10.1080/14788940410001655916

Google Scholar

Vojt et al., 2013

G. Vojt, L.D.G. Thomson, L.A. Marshall

The predictive validity of the HCR-20 following clinical implementation: does it work in practice?

The Journal of Forensic Psychiatry & Psychology, 24 (3) (2013), pp. 371-385, 10.1080/14789949.2013.800894

CrossRefView Record in ScopusGoogle Scholar

Webster et al., 1997a

C. Webster, K. Douglas, D. Eaves, S. Hart

HCR-20 assessing risk for violence – version 2 manual

Mental Health, Law and Policy Institute (1997)

Google Scholar

Webster et al., 1997b

C.D. Webster, K.S. Douglas, D. Eaves, S.D. Hart

HCR-20: Assessing risk for violence, version 2

Mental Health, Law, and Policy Institute, Simon Fraser University (1997)

Google Scholar

Webster et al., 1995

C.D. Webster, D. Eaves, K.S. Douglas, A. Wintrup

The HCR-20 scheme: The assessment of dangerousness and risk

Simon Fraser University and Forensic Psychiatric Services Commission of British

Columbia (1995)

Google Scholar