Violence Risk Assessment with the HCR-20^{V3} in Legal Contexts: A Critical Reflection

Vivienne de Vogel, Tamara De Beuf, Stephane Shepherd & Richard D. Schneider

Published online: 21 Jan 2022.

Submit your article to this journal ⬈

View related articles ⬈

View Crossmark data ⬈

Routledge
Taylor & Francis Group

SPECIAL ISSUE: PERSONALITY ASSESSMENT IN LEGAL CONTEXTS

OPEN ACCESS | Check for updates

Violence Risk Assessment with the HCR-20^{V3} in Legal Contexts: A Critical Reflection

Vivienne de Vogel[1,2,3] (iD), Tamara De Beuf[4] (iD), Stephane Shepherd[5], and Richard D. Schneider[6]

[1]De Forensische Zorgspecialisten, Utrecht, The Netherlands; [2]Maastricht University, Maastricht, The Netherlands; [3]University of Applied Sciences Utrecht, Utrecht, The Netherlands; [4]Ottho Gerhard Heldring Institution, Zetten, The Netherland; [5]Centre for Forensic Behavioural Science, Victorian Institute of Forensic Mental Health, Swinburne University of Technology, Hawthorn, Victoria, Australia; [6]Faculties of Medicine and Law, University of Toronto, Toronto, Ontario, Canada

ABSTRACT

The HCR-20^{V3} is a violence risk assessment tool that is widely used in forensic clinical practice for risk management planning. The predictive value of the tool, when used in court for legal decision-making, is not yet intensively been studied and questions about legal admissibility may arise. This article aims to provide legal and mental health practitioners with an overview of the strengths and weaknesses of the HCR-20^{V3} when applied in legal settings. The HCR-20^{V3} is described and discussed with respect to its psychometric properties for different groups and settings. Issues involving legal admissibility and potential biases when conducting violence risk assessments with the HCR-20^{V3} are outlined. To explore legal admissibility challenges with respect to the HCR-20^{V3}, we searched case law databases since 2013 from Australia, Canada, Ireland, the Netherlands, New Zealand, the UK, and the USA. In total, we found 546 cases referring to the HCR-20/HCR-20^{V3}. In these cases, the tool was rarely challenged (4.03%), and when challenged, it never resulted in a court decision that the risk assessment was inadmissible. Finally, we provide recommendations for legal practitioners for the cross-examination of risk assessments and recommendations for mental health professionals who conduct risk assessments and report to the court. We conclude with suggestions for future research with the HCR-20^{V3} to strengthen the evidence base for use of the instrument in legal contexts.

The importance of violence risk assessment in legal decision-making regarding appropriate sentences, judicial interim release, or matters regarding civil commitment is widely acknowledged. There are high stakes for public safety as violence risk assessment, including risk management strategies may prevent violent recidivism, but it is also crucial for justice-involved persons, as the results may seriously limit a person's liberty. Over the past three decades, empirical knowledge about violence risk assessment has grown tremendously (see for reviews Douglas & Otto, 2021; Singh et al., 2016) and violence risk assessment instruments are widely used in different settings and contexts, mostly in forensic mental health and correctional settings, but also in civil psychiatric, social service, and corporate settings, as well as for pretrial decision-making. Inclusion of a risk assessment tool is nowadays considered best practice in forensic mental health and criminal justice (Heilbrun et al., 2021). Moreover, in multiple contexts and different jurisdictions, laws or policies mandate the use of violence risk assessment tools (Cox et al., 2018; Viljoen & Vincent, 2020). Currently, more than 400 risk assessment instruments are being used in over 40 countries (Singh et al.,

2014). Initially, structured violence risk assessment instruments were mainly developed to ascertain a person's level of risk for violence or dangerousness and inform decision-making. More recently, the field of risk assessment has evolved from focusing largely on risk prediction to greater consideration of risk formulation and understanding of risk as well as risk management and risk reduction (see for a more elaborate discussion about generations of violence risk assessment instruments Heilbrun et al., 2021).

Notwithstanding the improvements in structured violence risk assessment over the past 30 years, there remain challenges with respect to the use of these tools in clinical practice as well as in legal contexts. As the use of risk assessment instruments has increased, so has the diversity of professionals providing violence risk expert evidence in court (Storey et al., 2013). As a result, the quality of violence risk assessments may fluctuate: not all evaluators are equally competent and not all instruments have a strong empirical basis or have been appropriately validated (e.g., Fazel, 2019; Hopton et al., 2018). Despite these issues, research to date shows that courts generally accept the findings of risk assessments without

---

evidentiary challenges being raised (Cox et al., 2018; Neal et al., 2019). Challenging the admissibility of expert evidence in court is imperative given the influence forensic (psychological) evidence can have on the legal decision-making process (Neal et al., 2019). Therefore, it is essential that violence risk assessment instruments tendered as evidence are relevant to the legal issue in dispute, to the particular justice-involved person, and that they are employed by trained and experienced evaluators who take measures to limit potential biases. More professional attention and research into the use and legal admissibility of risk assessment as expert evidence in court, is warranted (Neal et al., 2019; Slobogin, 2020).

In the present critical reflection, we will discuss the use of the *Historical, Clinical, Risk management-20 Version 3* (HCR-20^V3; Douglas et al., 2013), a widely used violence risk assessment tool constructed according to the structured professional judgment (SPJ) approach in legal contexts. We aim to provide legal practitioners with an overview of the strengths and weaknesses of the HCR-20^V3, particularly when applied in legal settings. First, the SPJ approach and the HCR-20^V3 will be described in more detail, followed by a discussion of its psychometric properties. Second, we present issues that may affect the legal admissibility of violence risk assessments with the HCR-20^V3, including potential bias and use in different justice-involved groups. Third, we explore the use and challenges of the HCR-20 Version 2/HCR-20^V3 in court since 2013 by examining case law databases from Australia, Canada, Ireland, the Netherlands, New Zealand, the UK, and the USA. The paper ends with suggestions for future research with the HCR-20^V3 and recommendations for legal practitioners relevant to cross-examination and critical evaluation of expert evidence involving the HCR-20^V3 as well as recommendations for forensic mental health professionals who conduct risk assessments and report to the court.

Violence risk assessments in legal contexts

Many legal decisions, such as sentencing or civil commitment, call for expert evaluation with a strong emphasis on risk prediction (Heilbrun et al., 2021). Therefore, it may not be surprising that most deliberations on the use of risk assessment instruments in legal settings have focused on actuarial risk assessment instruments which are aimed at predicting outcomes, such as violent recidivism (e.g., see a special issue of *Behavioral Sciences & the Law*, Slobogin 2020). Actuarial risk assessment instruments are initially designed to predict future offending based on group-statistics and are widely used in legal contexts. Risk factors in actuarial tools are coded based on fixed and explicit rules and subsequently summed up or weighted and combined according to an algorithm, resulting in a risk score. Many actuarial tools contain mainly historical risk factors that cannot be changed by intervention or forensic treatment, but several actuarial tools also consider dynamic risk factors and aim to provide risk management strategies, for instance, the Level of Service Instruments (see for a recent overview Wormith & Bonta, 2021) and the STABLE-2007 (Fernandez et al., 2014) for assessing risk for sexual recidivism.

Structured professional judgment

During the late 1990s, the SPJ approach was developed in Canada by a group of scholars with the aim to focus on the individual and on mitigating risks by providing guidelines for risk management, that is, the needed level of security and supervision, and treatment interventions. These scholars had expressed concerns about the actuarial approach with respect to the use of actuarial tools in clinical and correctional contexts, mainly relating to sample dependence, exclusion of potentially relevant risk factors, limited relevance to risk management, and under-emphasis on dynamic or changeable risk factors (see Hart & Cooke, 2013 for an elaborate discussion). Although SPJ tools also comprise empirically established risk factors, the methodology to reach a conclusion is fundamentally different from the actuarial approach. That is, the risk factors in SPJ instruments are interpreted, integrated, combined, and weighed by the evaluator to arrive at a final individualized risk judgment. Thus, the essence of the SPJ approach is more about the process of coding and about individualizing, understanding and explaining than about the summing of risk factors.

As Heilbrun et al. (2021) describe, the SPJ approach is valued by mental health professionals, mainly because they feel acknowledged in their forensic clinical expertise and they feel strengthened by the empirical basis of the SPJ instruments. That said, it remains vital to subject these instruments to rigorous evaluation of their reliability and validity for target groups of interest. It should be noted that in prospective research designs, the SPJ approach inherently hampers predictive validity analyses, because when these tools are used as intended in clinical practice, mental health professionals intervene in cases of high risk, for instance, by treatment programs aimed at reducing risk factors and enhancing protective factors, with the aim of mitigating the risk of recidivism.

HCR-20^V3

The most widely used SPJ tool for adults is the *Historical Clinical Risk management-20 Version 2* (Webster et al., 1997) for the assessment of risk for future violence. In an international survey of mental health professionals from 44 countries on six continents, the HCR-20 was found to be the most commonly used tool in violence risk assessment practices (Singh et al., 2014). After reviewing the scientific and professional literature, consulting with international experts and beta-testing of the draft version, a third edition of the HCR-20 was published in 2013 (HCR-20^V3; Douglas et al., 2013). The purpose of the HCR-20^V3 is to provide structure to the risk assessment process and to inform clinical practice with respect to case prioritization. It may also assist in monitoring a person's risk over time given the periodical reevaluation of the dynamic factors. The authors recommend formal reassessment of risk at least every 6 to 12 months, or whenever there is an important change in the status of the case (e.g., unsupervised leave). Therefore, the HCR-20^V3 is *not* designed to predict for the longer term. Instead, its purpose is to establish the presence and relevance of risk factors within a specific context and timeframe enabling the development of adequate risk management

strategies and treatment plans aimed at mitigating those context-specific risks.

The HCR-20[V3] contains 10 static risk factors (Historical scale) and 10 dynamic risk factors (Clinical and Risk management scale) and is intended for use with men and women age 18 and above to "evaluate risk for violence when there is a legal or clinical need to do so" (Douglas et al., 2013, p. 35). This implies that the HCR-20[V3] can be used for decision-making about violence risk within correctional, civil psychiatric, and forensic mental health settings, whether institutional or community-based. The user guide informs about common applications of HCR-20[V3] including: "release decision-making (from correctional, psychiatric, or forensic facilities); admission decision-making (upon entry to correctional, psychiatric, or forensic facilities); monitoring of risk while a person is incarcerated or institutionalized; and monitoring of risk while a person is under a term of community supervision by correctional, forensic, or psychiatric authorities" (Douglas et al., 2013, p. 35). In the HCR-20[V3], violence is defined as "actual, attempted, or threatened infliction of bodily harm of another person" (Douglas et al., 2013, p. 36). In most jurisdictions this would include psychological harm. This definition does not require a prior conviction for a violent offense or an official mental health diagnosis. The HCR-20[V3] manual guides evaluators through seven steps: 1) gathering case information; 2) coding the presence of 20 risk factors and their sub-items; 3) judging the relevance of these risk factors with respect to the development of future risk management strategies; 4) risk formulation (developing an individualized theory of violence); 5) planning risk scenarios; 6) recommending risk management strategies; and 7) documenting conclusory opinions (summary risk ratings in terms of low, moderate or high).

The new steps in the HCR-20[V3], of relevance rating, risk formulation, and scenario planning, are considered useful by practitioners in clinical contexts to help them understand and explain violence and to develop individualized strategies to prevent violence (see Hopton et al., 2018; Logan, 2014). For instance, in the step of risk formulation, evaluators should integrate separate risk factors into a conceptual meaningful framework that explains a person's violence. However, research on these steps is still limited, as well research on the translation of risk assessment findings into risk management strategies (see also Viljoen & Vincent, 2020). Moreover, the vast majority of risk management instruments, like the HCR-20, are validated primarily for the extent to which risk judgments and item/total scores are associated with future violent offending. While this provides evidence of predictive ability, it provides little guidance in determining which risk factors are causally related to violence and what effect a change in these factors would have on the probability of future violence, which is presumably the purpose of a risk management instrument (Spivak & Shepherd, 2020).

Summary of research results with the HCR-20[V3]

The HCR-20 Version 2 has been the subject of studies in multiple countries and settings, showing in general good psychometric properties and clinical value (for an overview see Douglas et al., 2017). However, research into the psychometric properties and clinical value of the HCR-20[V3] is still in its infancy. Recently, Douglas and Shaffer (2021) summarized the research conducted with the HCR-20[V3] thus far.

Reliability

Douglas and Shaffer (2021) found 17 published evaluations of interrater reliability of codings and summary risk ratings, including 465 participants from nine different countries across multiple settings (forensic, civil psychiatric, correctional). Overall, the interrater reliability is reported to be acceptable, with the majority of ICC values in the .70 s to .80 s. Furthermore, concurrent validity was found to be good. A strong association was found between HCR-20 Version 2 and 3 (e.g., de Vogel et al., 2014; Strub et al., 2014), as well as a significant association with other established risk assessment instruments such as the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995; see Persson et al., 2017).

Predictive validity

In their review, Douglas and Shaffer (2021) summarize 15 studies conducted in eight countries in several, mostly forensic populations that evaluated the predictive validity of the HCR-20[V3] for post-discharge community violence and inpatient violence and conclude that the tool resembles its predecessor, the HCR-20 Version 2. They conclude that, overall, there is evidence that the HCR-20[V3] items predict violence, with some indication that the dynamic items might be more strongly related to institutional or short-term violence. For example, in a retrospective study with a sample of 99 Canadian forensic psychiatric inpatients, Hogan and Olver (2016) found significant predictive accuracy for the HCR-20[V3] total score for inpatient aggression (AUC = .76). The HCR-20[V3] dynamic scores demonstrated incremental predictive validity for inpatient aggression to varying degrees after controlling for static risk factors. In a more recent follow-up study, Hogan and Olver (2019) studied the predictive value of several tools for community violence in 82 patients discharged from a maximum security forensic psychiatric hospital and found that dynamic change scores computed from the HCR-20[V3] relevance ratings demonstrated incremental predictive validity for community violence, controlling for baseline scores. Furthermore, several studies have shown that there is evidence that the dynamic risk factors can measure change during interventions or treatment. To illustrate, Penney et al. (2016) conducted a study with a sample of 87 forensic psychiatric patients transitioning to the community and showed that the HCR-20[V3] dynamic risk factors exhibited significant change across time and this change was related to clinically-relevant outcomes, such as violence or rehospitalization.

More recent studies not included in the Douglas and Shaffer review reached similar conclusions. For example, in a retrospective study in 100 Australian forensic psychiatric patients, Brookstein et al. (2020) found that the HCR-20[V3] significantly predicted violent recidivism after discharge

(AUCs= .70 to .77). Most of the above-described studies have been conducted in forensic or correctional samples assessing predictive accuracy of the HCR-20[V3] for inpatient violence or post-discharge violence and not in pretrial samples. Smith et al. (2014) explored the use of the HCR-20[V3] in a pretrial sample and noted several obstacles for coding the tool, for instance, limited file information and uncertainty regarding participants' legal status. The authors did not examine predictive validity, but found some noteworthy differences in mean item scores compared to other samples. Still, they concluded that their results indicated that the HCR-20[V3] seems to be adaptable for use in a pretrial context, but more research is needed.

Incremental validity of summary risk ratings

An important question is if the summary risk ratings (step 7) add incrementally to the mere coding of the items (step 2). This is important, as this step is one of the most distinguishing aspects of SPJ tools like the HCR-20[V3]. Most studies into the HCR-20[V3] to date have only examined the predictive validity of the numerical codings. A few studies have studied numeric totals of scales as well as summary risk ratings of the HCR-20[V3] finding good predictive validity for both (see Hogan & Olver, 2016, 2019; Persson et al., 2017). A smaller number of studies tested incremental validity and found that the summary risk ratings added incrementally to numerical scores (Neil et al., 2020; Strub et al., 2014). For example, in a Canadian sample of 56 offenders and 50 civil psychiatric patients, Strub et al. (2014) found that summary risk ratings added incremental validity to both presence (step 2) and relevance (step 3) ratings and for short and longer-term post-discharge violence prediction. These findings are in line with studies into other SPJ tools finding incremental validity for the summary risk ratings/ final risk judgments (see Heilbrun et al., 2021).

The good predictive validity of summary risk ratings (in HCR-20 Version 2 named final risk judgments) has also been found in a substantial number of studies and meta-analyses with the HCR-20 Version 2 (Douglas et al., 2014; Singh et al., 2011; see for a summary of study results Douglas et al., 2017). Guy (2008) applied meta-analytic techniques to examine the predictive validity of the SPJ model using 113 disseminations about a number of SPJ tools including the HCR-20 and concluded that results supported the utility of the SPJ model (i.e., when summary risk ratings were used) and indicated no distinct superiority for either the actuarial or SPJ model. In a descriptive review of research, Douglas et al. (2014) reported that in 30 out of 34 published studies that have investigated HCR-20 summary risk ratings have found support for them and in 15 of the 17 studies that examined both numerical codings and summary risk ratings, incremental validity was found for the summary risk ratings. To illustrate, in a Dutch sample of 127 male forensic psychiatric patients, Cox regression analyses showed that the HCR-20 final risk judgment produced a significant improvement to the model's fit ($\chi2$ change (1, 127) = 6.8, p < .01) (de Vogel & de Ruiter, 2006).

Critique of the current HCR-20[V3] Research Base. A critical note about the current body of knowledge regarding the HCR-20[V3] is that several of the reported studies in the above-described review have been conducted with the draft version[1] and by HCR-20[V3] authors or translators, making them vulnerable to potential authorship bias (see also Judges et al., 2016). Authors or translators who know the tool very well may have better coding skills and show more enthusiasm and fidelity in the use of it, although no clear evidence has been found for authorship bias so far (Singh et al., 2013). Still, more research from independent research groups and in different samples, settings and countries is highly needed and is currently emerging (e.g., Brookstein et al., 2020; Hogan & Olver, 2016, 2019; Penney et al., 2016). To date, most of the studies with the HCR-20[V3] – like with the HCR-20 Version 2 - have been conducted in Western countries. This trend is gradually changing with research from Non-Western countries, particularly from East Asia (for an overview of international risk assessment practices, see: Singh et al., 2016). There are currently several published studies from Non-Western countries, like China (Zhou et al., 2016), Ghana (Adjorlolo & Chan, 2019), and Latin American countries (Folino, 2018) finding mixed results with the HCR-20. For example, Zhou et al. (2016) reviewed risk assessment studies in China and concluded that there is limited evidence to support the use of violence risk assessment instruments (including the HCR-20 Version 2) in general and forensic psychiatric patients in China. Adjorlolo and Chan (2019) found significant predictive validity for the HCR-20[V3] for general recidivism in their preliminary study in prisoners in Ghana. They did not examine violent recidivism however. Based on their findings, the authors state that using the HCR-20[V3] in African countries could contribute significantly to justice delivery, offender rehabilitation, and ultimately, to public safety. Obviously, more research on the psychometric properties of the HCR-20[V3] in Non-Western populations is needed.

Furthermore, it is important to recognize that several of the studies in the Douglas and Shaffer (2021) review are considered to be lab studies (i.e., conducted by researchers and without real-life implications) and that these results cannot always be generalized to routine clinical practice (for a discussion about the importance of field studies, see: Edens & Boccaccini, 2017). Moreover, most of these studies have used the tool in an actuarial way and have not sufficiently examined (incremental) validity of the summary risk ratings. A final critique is that the HCR-20[V3] has not yet been sufficiently validated in specific groups of justice-involved people, including ethnic minorities, women, or persons with mental disabilities or specific offenses (see section below on *Value of the HCR-20[V3] for Different Justice-Involved Groups*).

---

[1]There were no substantial differences in items between the Draft and final version, except for item *H6 Major mental disorder* and *C3 Symptoms of major mental disorder*, as the DSM-V was just released and incorporated in the definitions of the final version.

Use of the HCR-20[V3] in legal contexts

With respect to the use of the HCR-20[V3] in legal contexts, it is stated that "the user should have full understanding and knowledge of the legal criteria relevant to risk that evaluees will be subject to, as well as to broader legal context and legal procedures" (Douglas & Shaffer, 2021, p. 261). These legal criteria may differ across jurisdictions. Generally, the HCR-20[V3] seems to be most suited to assist decision-making for institutional violence and for legal decisions such as conditional release and prolongation of involuntary treatment or civil commitment. This is corroborated by studies with the HCR-20 Version 2. For example, Vitacco et al. (2018) found that violence risk assessments with the HCR-20 Version 2 did not predict long-term outcomes for insanity acquittees in the community, but may be valuable in identifying relevant treatment targets aimed at reducing violence within this specialized population.

Legal admissibility challenges

Admissibility of expert evidence refers to whether the evidence can be accepted by the court to assist legal decision-making (Neal et al., 2019). The criteria on which the admissibility[2] of evidence is based, differ considerably across jurisdictions. Legal standards for admissibility vary from no standards to legislation incorporating scientific considerations. For example, in some jurisdictions (e.g., England, Wales, Australia, and some states in the USA) expert evidence is admitted when it is judged to be sufficiently established and generally accepted as a body of knowledge, while other jurisdictions require (scientific) evidence of the reliability of the used method (e.g., most states in USA, Canada; Edmond et al., 2014).[3] Admissibility criteria based on scientific reliability and validity are arguably the most rigorous. One of the most referenced sets of admissibility criteria are the American Daubert criteria[4]: 1) whether the method used by the expert has been subjected to empirical testing; 2) whether it has been subjected to peer review and publication; 3) whether the error rate for the method is known (or potentially known); 4) the existence and maintenance of standards controlling its operation; and 5) whether procedure or test used by the expert is generally accepted within the relevant scientific community. It is important to provide relevant information about the "fit" between the method used and the specific case facts. This requirement suggests that courts should be particularly attentive to whether there are data from the field about the reliability and validity of the tool and whether the tool was designed to address both the issue at hand (e.g., parental

fitness, risk of reoffending) and the population in question (Neal et al., 2019). Whether or not expert evidence will be permitted will be decided by the presiding judge. Typically, the threshold to be met is whether the expert (opinion) evidence is necessary (not just *helpful*) to answer the question(s) before the court. As to whether particular types of data/evidence may be relied upon or received by the court will be subject to the sorts of considerations articulated in Daubert.

Below, we discuss two issues that may arise as potential challenges to the legal admissibility of violence risk assessments with the HCR-20[V3]: 1) potential biases in risk assessment with the HCR-20[V3]; 2) validity in different justice-involved groups, including cross-cultural validity.

Potential bias in HCR-20[V3] assessments

When violence risk assessment is included as expert evidence, it is expected to be unbiased. The issue of bias in forensic science and forensic psychology has recently gained more attention as the field came to realize that forensic evaluators –no different than other humans– are susceptible to bias (Neal & Grisso, 2014; Zapf & Dror, 2017). In forensic risk assessment, evaluators learned to rely on instruments to limit the impact of bias. Recent studies, however, suggest that these instruments are not a cure-all and biases may still impact the risk assessment (Neal & Brodsky, 2016). There is, for example, proof of adversarial allegiance in risk assessment, evidenced by experts assigning higher risk scores when they thought they were working for the prosecution compared to experts who believed they were working for the defense (Murrie et al., 2013).

It should be noted that most studies about bias in risk assessment concern actuarial instruments, few studies have examined the role of bias in HCR-20 assessments specifically. In their review of psychological assessments in legal contexts, Neal et al. (2019) stated that SPJ tools like the HCR-20[V3] are challenging to evaluate with respect to legal admissibility because they are designed with the individual in mind rather than for group-based predictions. One type of bias that has been demonstrated with the HCR-20 is attribution bias, that is, evaluators rated the HCR-20 differently depending on whether they attributed behaviors of the evaluee to internal (i.e., individual characteristics) versus external (e.g., due to circumstances) factors. In experiments with 40 undergraduate students, Murray et al. (2014) found that evaluators gave higher scores on the Historical and Clinical scales when internal attribution was triggered compared to external attribution. Attitudes toward offenders have also been found to influence HCR-20 ratings. Recently, Kamorowski, de Ruiter, et al. (2021) demonstrated that risk evaluators (i.e., individuals trained to complete any version of the HCR-20) with a more negative attitude toward offenders gave higher ratings on the HCR-20[V3] Clinical and Risk Management scales and gave a higher estimate on the summary risk ratings. The opposite was true for evaluators with more positive attitudes toward offenders. These studies, conducted in Western countries, demonstrate that, although

---

[2]Admissibility requires that the evidence be logically probative of an issue in dispute and that its probative value is greater than its prejudicial impact.

[3]For a more comprehensive discussion of legal admissibility, we refer to Schneider (2016) with respect to expert evidence in general, and Slobogin (2021) for violence risk assessment in particular.

[4]In the United States federal law, the Daubert standard is a rule of evidence regarding the admissibility of expert witness testimony (Daubert v. Merrell Dow Pharmaceuticals, 1993). See for more elaborate discussion of the Daubert criteria Glancy and Saini (2009) and Neal et al. (2019).

structured risk assessment with instruments like the HCR-20 is meant to minimize error and bias, bias may still play a role. It should be noted that biases do not automatically result in inaccurate judgments (de Vogel & de Ruiter, 2006). Still, it is important for evaluators and decision-makers to be aware of the possibility that bias may affect the reliability and validity of the assessment and potentially have serious consequences for evaluees (Neal & Grisso, 2014).

Value of the HCR-20^{V3} for different justice-involved groups

In addition to the question of whether violence risk assessment with the HCR-20^{V3} can be accepted as expert evidence in general, there is the question of whether the validity findings can be generalized to the specific case at hand (i.e., the issue of fit; Slobogin, 2021). There are recent examples of court rulings in which a risk assessment expert testimony was considered not relevant, because the respective risk assessment instrument was not validated for the particular subgroup to which the defendant arguably belonged (Slobogin, 2021). For example, the Canadian Supreme Court ruled that when using risk assessment instruments to assess Indigenous offenders, one should ensure that there is proof of the instrument's accuracy for this population (Ewert v. Canada, 2018). This is in line with the *Specialty Guidelines for Forensic Psychologists* from the American Psychological Association (APA): "Forensic practitioners use assessment instruments whose validity and reliability have been established for use with members of the population assessed. When such validity and reliability have not been established, forensic practitioners consider and describe the strengths and limitations of their findings" (APA, 2013, p. 15). Most courts, however, are less strict in their approach to this issue of fit (i.e., validation in local populations) and awareness of this validation issue suffices (Slobogin, 2021). Relevant to the issue of fit, we present an overview of the current knowledge regarding the performance of the HCR-20^{V3} in several justice-involved minority groups (i.e., cultural minorities, justice-involved women, people with intellectual disabilities, specific offender types).

Cultural minority groups

It is important that violence risk assessment approaches are applicable to different cultural populations to avoid discriminatory medico-legal practice. Instruments that are less effective for particular cultural groups could disadvantage them and potentially further entrench them in the criminal justice system. There is now an extensive body of literature spanning several decades and a growing research base considering cross-cultural risk assessment.

Findings from this literature indicate that commonly employed forensic risk instruments (both SPJ and actuarial tools) are generally able to predict recidivism for nonwhite populations at similar levels of accuracy to White populations (Lowder et al., 2019; Olver et al., 2014; Shepherd et al., 2014; Wilson & Gutierrez, 2014). It should be noted,

however, that for the HCR-20 in particular, there are almost no validation studies employing the HCR-20 with key minority populations (e.g., African-Americans, Hispanic, and First Nations communities), with few exceptions (see Fujii et al., 2005; Snowden et al., 2010). This is unusual given the prominence of the instrument - future studies with the HCR-20^{V3} should endeavor to include minority cohorts.

While these methodological improvements are encouraged, it is important to remember that using risk assessment instruments protects against the impressionistic nature of unstructured clinical judgments, which are more prone to bias and negative/stereotypical heuristics. Such biases could disproportionately impact justice-involved cultural minorities. SPJ instruments, which allow for clinical discretion are already susceptible to such biases. For assessors employing SPJ instruments, caution is advised when injecting cultural information/perceptions into the discretionary spaces into the spaces in an SPJ assessment. There are no agreed-upon scientifically grounded cultural-specific risk or protective factors related to offending in the forensic literature that one can reliably draw upon (Shepherd & Spivak, 2021).

Justice-involved women

Mixed research results have been found with respect to the predictive validity of the HCR-20 Version 2 for women (Garcia-Mansilla et al., 2009; O'Shea et al., 2013) and little is known about the predictive validity of the HCR-20^{V3} for justice-involved women. A small-scale study with the HCR-20^{V3} draft version in a group of 100 male and 24 female insanity acquittees showed that the relationship between scale scores and violence was stronger for men than for women. However, gender was not a significant moderator in logistic regression analyses predicting the likelihood of violence (Green et al., 2016). In a Dutch multi-center study with a sample of 78 discharged female forensic psychiatric patients, it was found that various risk assessment tools, including both Versions 2 and 3 of the HCR-20 and a gender-specific additional tool, the Female Additional Manual (FAM; de Vogel et al., 2014) had moderate predictive validity for general recidivism, but low predictive validity for violent recidivism (de Vogel et al., 2019). The gender-specific tool performed better than the HCR-20, but did not provide incremental predictive value to the HCR-20^{V3}. It should be noted that these studies with justice-involved women comprised small samples and are likely underpowered, thus no firm conclusions can currently be drawn.

People with intellectual disability

The application of the HCR-20^{V3} to justice-involved people with an intellectual disability is less extensively researched. Hounsome et al. (2018) reviewed the literature on the predictive validity of risk assessment tools for violence in adults with an intellectual disability. Their review included four studies with the HCR-20 Version 2, finding significant predictive validity for general and violent reconviction in this

population. O'Shea et al. (2015) demonstrated that, after controlling for a range of potential covariates, the HCR-20 Version 2 was a significant predictor of inpatient aggression among patients with intellectual disability and performed equally well compared to mentally-disordered individuals without intellectual disability. To our knowledge, there are no published studies yet into the use of the HCR-20^{V3} in justice-involved populations with intellectual disability.

People who committed sexual offenses

Different risk factors are found to be predictive of different types of offenses. A substantial amount of the literature on risk assessment is concerned with research into risk factors for sexual offending. There are multiple reliable and validated instruments for the assessment of sexual violence risk, such as the Static-99R (Helmus et al., 2009), STABLE-2007, and the Sexual Violence Risk-20 Versions (SVR-20; Boer et al., 1997, 2017). While acknowledging the importance of using specialized instruments for the assessment of sexual recidivism, it should be noted that the literature shows that people who have committed sexual offenses are at greater risk of nonsexual than sexual recidivism (Cartwright et al., 2018). This underscores the relevance of assessing risk for general violence in sexual offender populations. Cartwright et al. (2018) examined several risk assessment instruments for predicting institutional aggression among patients detained for sexual offenses or civilly-committed pursuant to sexually violent predator legislation and found support for the predictive validity for general violence of HCR-20 Version 2, which was stronger than for the Static-99R. The authors advise to use tools like the HCR-20 in addition to instruments like the Static-99R to assess risk of general violence.

As far as we know, there are no studies yet with the HCR-20^{V3} in sexual offender populations. A relevant question in this respect is if and how the HCR-20^{V3} should be used in combination with risk assessment tools for sexual violence, considering the fact that sexual violence is included in the HCR-20^{V3} definition of violence. The HCR-20^{V3} authors strongly encourage evaluators to "use a risk assessment measure developed specifically for sexual violence in addition to the HCR-20^{V3} when a history of sexual violence is present" (Douglas et al., 2013, p. 69). However, they do not provide suggestions on how to integrate the results of multiple tools in a risk assessment report. An option for evaluators is to make a clear distinction between types of violence by applying the HCR-20^{V3} solely for the assessment of general violence risk and using a specialized tool for the judgment of risk for sexual violence. This choice should then be highlighted in the risk assessment report.

To conclude from the above described literature into the use of the HCR-20^{V3} for different justice-involved groups, there is not much information yet about the relative value of the tool across specific target groups. Further research is required to validate the HCR-20^{V3} for use with cultural minority groups, justice-involved women, individuals with intellectual disability, and those who have committed sexual offenses. Mental health professionals, and also legal decision-makers should be cognizant of these caveats and should be cautious when interpreting HCR-20^{V3} results for these groups. Other subgroups that warrant further examining are life-term inmates (Cox et al., 2018) and specific types of violence, such as terrorism.

Case law review

It is highly relevant to know how risk assessment tools are being used in real court cases and how often the results of the risk assessment are challenged with respect to admissibility. Cox et al. (2018) searched for the use of HCR-20 Version 2 and the Violence Risk Appraisal Guide (VRAG; Harris & Rice, 1997) in American case law through the legal database LexisNexis and found 134 cases for the years 2010 to 2016 that included the HCR-20 ($n = 107$), VRAG ($n = 16$) or both of the tools ($n = 11$). The tools were typically introduced by the prosecution to inform opinions regarding violence risk. The authors found that the HCR-20 Version 2 was challenged in 11 (9.3%) of the 118 cases, usually by defense attorneys who were concerned about the validity of the tool when used with inmates serving life-imprisonment. Challenges were rarely successful and, if successful, they never directly involved the tool itself.

To our knowledge, there are no case law reviews published with the HCR-20 in other countries nor with the HCR-20^{V3}. Therefore, we conducted a case law review to examine how often the HCR-20^{V3} has been challenged in court, for what reason, and whether these challenges were successful, that is, if the court upheld the challenge and the risk assessment report was considered inadmissible.

Case law review HCR-20^{V3}

We searched English and Dutch language case law databases and selected cases from the year 2013 when the HCR-20^{V3} was officially published. The following databases were searched: WestLawNext for Australia and the USA; WestLawUK for the UK; Canadian Legal Information Institute (CanLII) for Canada (https://www.canlii.org/en/); De Rechtspraak for the Netherlands (www.rechtspraak.nl); New Zealand Legal Information Institute (NZLII) for New Zealand (http://www.nzlii.org/); and British and Irish Legal Information Institute (BAILII) for Ireland (https://www.bailii.org/). The databases were accessed via the online library of Maastricht University or were freely available. There was one exception; we searched in the USA WestLaw database from January 2017 in addition to the study of Cox et al. (2018) who performed their search until December 2016. Using *HCR-20* as the search term, we found 546 cases: Australia ($n = 175$), Canada ($n = 161$), Ireland ($n = 2$) the Netherlands ($n = 127$), New Zealand ($n = 21$), UK ($n = 21$) and USA ($n = 39$). The coding of the cases was performed by the first and second author, in line with previous studies (Cox et al., 2018; Edens et al., 2015). There was no interrater reliability check, but the raters had frequent meetings during the coding process and discussed all challenges.

Table 1. Challenges case law review HCR-20/HCR-20^{V3} since 2013, N = 546.

	Number of cases	Percentage
Instrument		
HCR-20	107	19.60%
HCR-20^{V3}	153	28.02%
Not specified	286	52.38%
Challenged		
Referred, but not challenged	524	95.97%
Challenged	22	4.03%
By defense attorney	20	3.82%
By applicant	2	0.35%
Challenges tool		% of challenged cases
Not sufficiently validated	3	13.63%
Not validated for sex offenders	1	4.55%
Not validated for Indigenous offenders	7	31.81%
Not validated for people with intellectual disabilities	1	4.55%
HCR-20^{V3} is not developed to predict but to evaluate treatment and case prioritization	1	4.55%
Challenges assessor/process		
Risk assessment is outdated/not based on recent information	3	13.63%
HCR-20 is not used according to the manual (no coding sheet used, assessor renamed items)	1	4.55%
Lack of reliable data to code the tool	1	4.55%
Defendant does not have a history of violence	2	9.09%
Too much weight given to instrument and neglect of other data	2	9.09%
Outcome challenge		
Court did not uphold challenge, i.e., risk assessment was still considered admissible	22	100%

We emphasize that this case law review should be considered preliminary: it does not provide a complete and thorough picture, as we were dependent on English/Dutch language databases that were freely available or available via Maastricht University library. These databases do not contain all legal cases (see for instance https://www.bailii.org/bailii/summary-cases.html). Furthermore, there are important differences between jurisdictions, for instance, in legal definitions, making it difficult to compare results. Therefore, we will not compare results per country, but present an aggregated overview.

Results of case law review HCR-20^{V3}

Table 1 presents the results of the case law review. Although we searched databases after the publication date of the HCR-20^{V3}, most of the cases still referred to HCR-20 Version 2. In 153 (28.0%) of the cases the HCR-20^{V3} was used, in the other cases HCR-20 Version 2 was used or the version was not defined. Most of the cases that referenced the HCR-20/HCR-20^{V3} were criminal cases ($n = 397$, 72.7%), including fitness to stand trial, or civil law ($n = 81$, 14.8%); dangerousness determination, civil rights during admission to forensic psychiatric hospital). A few cases dealt with migration law ($n = 31$, 5.7%), administrative law ($n = 32$, 5.9%), or family law ($n = 5$, 0.9%). The HCR-20/HCR-20^{V3} was usually introduced by the prosecution to inform opinions regarding future violence risk.

The HCR-20 was rarely challenged, on average in 4.03% of the cases ($n = 22$; range countries 0% − 7.4%). All but two of these cases concerned HCR-20 Version 2. All challenges were presented by defense attorneys. The challenges can be divided into challenges to the tool itself ($n = 13$, 59.1%) and to questionable practices of the assessor or the

assessment process ($n = 9$, 40.9%). Challenges were never successful; that is, the court was never persuaded by such a challenge that the risk assessment was so unreliable as to be inadmissible. Finally, we made some notable observations during the reviewing of the cases. In several cases, we noticed that the assessor used the instrument in an actuarial way by reporting the total score and applying numerical communication, usually based on an empirical study (e.g., "similar scoring individuals on the HCR-20 test reoffended at a rate of 93% over 7.5 years in the community"), instead of communicating the level of risk in descriptive words (categories: low, moderate, high). This is notable as there are no official norms for HCR-20 coding and this goes against the user guidelines outlined in the manual. However, this numerical communication was not challenged by the defense attorneys in these cases. Virtually all of these cases concerned the Version 2 of the HCR-20, and as HCR-20^{V3} sets clear guidelines for risk communication in descriptive terms, this may have changed in recent years.

Understanding of risk assessment reports in court: Risk communication

Risk communication forms an essential part of the risk assessment process and is the link between assessment results and the subsequent decision-making about risk management strategies, but has only recently received empirical attention (Heilbrun et al., 2016; Hilton et al., 2015). The literature to date shows that the way of communicating may affect decision-making. For example, research has shown that decision-makers can be influenced how numeric risk estimates are framed or whether the estimate is formulated as a frequency or a probability). Frequency reporting (1 in 10) leads to more conservative decisions as compared to

Table 2. Questions for legal decision-makers to consider for cross-examination about HCR-20[V3] results.

Evaluator	
1.	Is the evaluator educated and trained in using the HCR-20[V3]?
2.	Does the evaluator have expertise performing violence risk assessments with the HCR-20[V3] or was the assessment conducted under supervision of a trained and experienced evaluator?
Evaluee	
3.	Does the evaluee belong to a specific justice-involved group (e.g., relating to gender, ethnicity, type of offense, intellectual disabilities) for whom the HCR-20[V3] has been proven reliable and valid?
4.	In the case of sexual offenses: is the HCR-20[V3] being used for assessment of sexual and/or non-sexual violence risk? Are specialized instruments used for assessment of sexual risk in addition to the HCR-20[V3]?
Risk assessment	
5.	Did the evaluator have access to reliable and recent information about the evaluee?
6.	Did the evaluator conduct the risk assessment by strictly following the guidelines of the HCR-20[V3] user guide?
7.	Were precautions taken to mitigate possible biases in coding the HCR-20[V3]?
8.	Is the likely dispositional outcome taken into account when reaching a risk conclusion (i.e., is the probable context taken into consideration)?
Risk communication	
9.	Is the aim and context of the assessment clearly explained, as well as the valid time frame?
10.	Are the most important risk factors explained and discussed in relation to each other?
11.	Are the conclusory opinions well-substantiated (i.e., not merely mentioned as low, moderate or high risk, but also risk scenarios, nature, type, frequency of risks)?
12.	Are recommendations provided for risk management strategies (i.e., needed level of security and supervision, appropriate interventions or treatment)?

percentages (10%), similarly, negative framing (20% chance of recidivism) leads to more conservative decisions than positive framing (80% chance of success; Scurich, 2018). Overall, numerical risk communication is often misunderstood by receivers, also referred to as *risk illiteracy* (Heilbrun et al., 2016).

Both mental health and legal practitioners show a preference for categorical risk communication; that is, in descriptive terms of low, moderate, high, in line with the SPJ approach (Singh et al., 2014). Categorical risk communication, however, has been criticized that it is not sufficiently specified, lacks an empirical base, and is not always well-understood or appreciated in a consist way by judges (Hilton et al., 2015; Scurich, 2018). According to a Canadian case law review (Storey et al., 2013), judges seemed to prefer the following risk communication strategies: identify and describe the instrument, how it was employed, how conclusions were reached, and the probability and severity of risks, including a worst-case scenario of re-offense. In another study, Storey et al. (2015) found that most pre-sentence reports to the court were not consistent with the SPJ guidelines, that is, most evaluators failed to document or concretely discuss risk factors, scenarios and management strategies. In summary, risk communication in categorical terms is preferred by both mental health professionals and legal practitioners, however, clarification of the meaning of risk categories is needed, as well as more empirical research on how the results are understood by decision-makers or by evaluees.

Suggestions for future research with the HCR-20[V3]

It becomes apparent from our review that more research is needed to enhance the legal admissibility of violence risk assessment with the HCR-20[V3]. Both retrospective and prospective studies should focus on interrater reliability, concurrent and predictive validity for violence (both inpatient and official recidivism), incremental validity of summary

risk ratings over numerical codings, and clinical applicability for the assessment and management of violence risk. More attention should be paid to the value of the HCR-20[V3] in different countries, for different legal contexts, and in different subgroups of justice-involved persons, for instance, related to ethnic minority, offender type, gender, specific mental disorders or intellectual disabilities. More specifically, research into the relevance rating, risk formulation, risk scenarios, and the translation from risk assessment into risk management strategies is urgently needed (see also Viljoen & Vincent, 2020). Further research into the changeability of the HCR-20[V3] ratings and if these changes are actually related to changes in recidivism risk is also desirable. Furthermore, more attention is warranted to adequate risk communication with the HCR-20[V3] and how the results are understood and used by decision-makers. Specifically for the HCR-20[V3], Cox et al. (2018) recommend studies exploring how the updated HCR-20[V3] might change the expert's communication of risk to legal decision-makers. For example, the new step in the risk assessment process of risk formulation and relevance rating in addition to presence rating may facilitate communication to the court, but may also increase subjectivity, opening the door for potentially biased risk judgments and possible subsequent legal challenges.

Recommendations for legal scholars interpreting risk assessments with the HCR-20[V3]

Based on the literature described in this critical reflection, the case law review and our own experiences, we formulate recommendations for the critical appraisal and cross-examination of risk assessment reports that use the HCR-20. For legal practitioners to be able to effectively evaluate and challenge violence risk assessment evidence, they need to be familiar with the core principles of the SPJ approach and risk communication. These core elements involve the systematic collection, reviewing, combining, weighing, and integration of relevant information on empirically based risk

factors by a qualified evaluator for a specific context and timeframe. Note that the results of the HCR-20^{V3} are valid for a limited time period rather than a prediction for the long-term. Due to the dynamic nature of the items in the clinical and risk management scales, the instrument is better suited for short-term violence risk prediction. The authors recommend a formal re-assessment of risk at least every 6 to 12 months, or whenever there is an important change in the status of the case. Legal practitioners should be aware of the differences between actuarial and SPJ tools when assessing the relevance of the risk assessment report to the specific issue before the court. More specifically, they need to be cognizant that the HCR-20^{V3} has been primarily developed for decision making in clinical practice and release planning. Training in risk assessment should therefore not be limited to forensic mental health professionals: judges and attorneys can also benefit from risk assessment training tailored to their role. Vincent et al. (2012) recommend training with background information on the risk assessment instrument (i.e., rationale and research, benefits and limitations) and how the instrument relates to other tools used by forensic experts. Furthermore, legal professionals should be informed about how the risk information can guide risk management and how they can contribute to good quality violence risk assessment in court. Similarly, Boccaccini et al. (2013) underscore the importance for judges and jurors to understand the relevance of risk assessment results for legal decisions.

When it comes to critically appraising the written reports presented as evidence in court, legal practitioners should evaluate whether these reports clearly state the goals, context, and timeframe of the risk assessment in unambiguous language, use terminology consistently, and present a coherent risk formulation (e.g., how are risk factors related? Are there also protective factors?). Finally, it is important that risk assessment reports specify the risk (i.e., nature, frequency, severity, timeframe, potential victim) and provide concrete recommendations for risk management (e.g., needed level of security and desired interventions or treatment). In Table 2, we suggest questions for legal scholars to consider during the cross-examination of HCR-20^{V3} assessments. These can be used in relation to the Daubert criteria, or admissibility criteria in other jurisdictions.

Recommendations for mental health professionals

Forensic mental health professionals who report to the court are responsible for conducting risk assessments as objectively as possible. Awareness of the existence and possible presence of biases among expert witnesses is a precondition to the minimization of its impact, however, in and of itself, awareness and introspection are unsuccessful debiasing strategies. As an alternative to introspection, Neal and Brodsky (2016) suggest searching for behavioral indicators of one's biases. For example, risk evaluators can investigate whether they systematically assign high scores, or whether they tend to minimize historical risk factors. Such patterns in decision-making may become more apparent when forensic

experts consult with colleagues, which has been suggested as a promising debiasing strategy (Croskerry et al., 2013). Other bias mitigation strategies include making use of reliable base rate information and slowing down. The latter technique involves taking time to consciously think about the available information and add a (brief) time-out during the risk assessment (Croskerry et al., 2013). Lastly, and perhaps one of the most effective debiasing strategies, is limiting exposure to irrelevant contextual information (i.e., exposure control; Neal & Brodsky, 2016; Kamorowski, Ask, et al., 2021). Indeed, the HCR-20^{V3} user guide describes that evaluators should focus on non-redundant, reliable, and immediately relevant information. However, it may be challenging in routine clinical practice to establish what is relevant and irrelevant information in structured risk assessments. Thus far, there have been few empirical studies on the effectiveness of bias mitigation strategies in violence risk assessment practice. However, when and if effective, it will likely improve the reliability and validity of the risk assessments (Neal & Grisso, 2014).

In addition to limiting bias during risk assessments, reliability can be enhanced by implementing and maintaining good practices. There are several recommendations to provide for these good practices (see also de Vogel et al., 2014; Douglas et al., 2013; Logan, 2014): 1) adherence to the HCR-20^{V3} user guide during each assessment (even when highly experienced); 2) frequent consultation with colleagues; 3) attending periodical (refresher) training; and 4) keeping up with new research and developments regarding violence risk assessment. An important aspect to consider in the risk assessment process is risk communication. Forensic mental health professionals are advised to think carefully in advance about the purpose of the risk assessment and consider who will be the recipient of the risk communication (e.g., the court, the assessee, mental health professional colleagues). Furthermore, it is important to write in clear, unambiguous language, to specify the behavioral outcome, including nature, frequency, term, potential victims and clearly explain the risk in context of the individual (for more information about the process of risk formulation, see, for example Logan, 2014). We advise evaluators to always be transparent about the caveats and clearly describe the strengths and limitations of their findings and specify the context and time frame (see also the *Specialty Guidelines for Forensic Psychologists*, APA, 2013). Especially, when the risk assessment is used for pretrial evaluations, it may be difficult to code the Risk management items as the future context is not clear. If there is not sufficient reliable information to code the items of the HCR-20^{V3}, forensic practitioners should refrain from coding the tool and presenting conclusions about the violence risk. There are no clear rules how many items can be omitted, but generally, more than four omitted items seems to be problematic (Douglas et al., 2013). The HCR-20^{V3} manual states: When factors are omitted, evaluators should document this fact and qualify their opinions accordingly, acknowledging whether and how their opinions might have changed if complete information were available. They should also make recommendations

concerning how missing information could be obtained, given additional time or resources (Douglas et al., 2013, p. 44).

Conclusion

Based on the empirical and theoretical research described in this paper, we conclude that the HCR-20^V3 is most suited for legal decision-making with respect to conditional release and prolongation of involuntary treatment or civil commitment, and less suitable for (capital) sentencing and long-term prediction. With respect to the use of the HCR-20^V3 for pretrial evaluation, we believe that if there is enough reliable information available to code the items and if the (potential) future context for the individual is clearly described, the tool can be used, but with caution. We emphasize again that evaluators should always be transparent about the caveats of the risk assessment and clearly describe the strengths and limitations of their findings (see also the guidelines for forensic psychologists, APA, 2013).

Overall, it is important for legal practitioners to be cognizant that the HCR-20^V3 was developed to aid mental health professionals to gain insight into the risk level of the evaluee and develop appropriate risk management strategies. Assessments with the HCR-20^V3 should be seen as dynamic and strongly dependent on the (future) context and therefore, better suited for short-term violence risk assessment and conditional release decisions, more so than for long-term prediction. Additionally, the proficiency of evaluators in conducting the HCR-20^V3 assessments is crucial: do they demonstrate appropriate experience, education, and best-practice knowledge of risk assessment and risk communication? Moreover, evaluators should acknowledge potential biases and document which bias mitigation strategies were implemented. Overall, our expert opinion on the legal admissibility of the HCR-20^V3 is that the tool can be valuable as expert evidence, but only when used properly by trained and experienced raters and within the right legal or clinical context.

Disclosure statement

Vivienne de Vogel is one of the translators of the Dutch version of the HCR-20^V3 but does not receive personal financial benefits related to this.

ORCID

Vivienne de Vogel http://orcid.org/0000-0001-7671-1675
Tamara De Beuf http://orcid.org/0000-0001-5273-8523

Data availability statement

The case law review data is available upon requested from the first author.

References

Adjorlolo, S., & Chan, H. C. O. (2019). Risk assessment of criminal offenders in Ghana: An investigation of the discriminant validity of the HCR-20^V3. *International Journal of Law and Psychiatry*, 66, 101458. https://doi.org/10.1016/j.ijlp.2019.101458

American Psychological Association (APA). (2013). Speciality guidelines for forensic psychology. *The American Psychologist*, 68(1), 7–19. https://doi.org/10.1037/a0029889

Andrews, D. A., & Bonta, J. (1995). *The level of service-inventory-revised*. Multi-Health Systems.

Boccaccini, M. T., Turner, D. B., Murrie, D. C., Henderson, C. E., & Chevalier, C. (2013). Do scores from risk measures matter to jurors? *Psychology, Public Policy, and Law*, 19(2), 259–269. https://doi.org/10.1037/a0031354

Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk-20: Professional guidelines for assessing risk of sexual violence*. The Mental Health, Law, and Policy Institute, Simon Fraser University.

Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (2017). *Manual for version 2 of the Sexual Violence Risk-20: Professional guidelines for assessing and managing risk of sexual violence*. Protec International Risk and Safety Services Inc.

Brookstein, D. M., Daffern, M., Ogloff, J. R., Campbell, R. E., & Chu, C. M. (2020). Predictive validity of the HCR-20^V3 in a sample of Australian forensic psychiatric patients. *Psychiatry, Psychology and Law*, 1–18. https://doi.org/10.1080/13218719.2020.1775152

Cartwright, J. K., Desmarais, S. L., Hazel, J., Griffith, T., & Azizian, A. (2018). Predictive validity of HCR-20, START, and static-99R assessments in predicting institutional aggression among sexual offenders. *Law and Human Behavior*, 42(1), 13–25. https://doi.org/10.1037/lhb0000263

Cox, J., Fairfax-Columbo, J., DeMatteo, D., Vitacco, M. J., Kopkin, M. R., Parrott, C. T., & Bownes, E. (2018). An update and expansion on the role of the Violence Risk Appraisal Guide and Historical Clinical Risk Management-20 in United States case law. *Behavioral Sciences & the Law*, 36(5), 517–531. https://doi.org/10.1002/bsl.2376

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 2: Impediments to and strategies for change. *BMJ Quality & Safety*, 22(Suppl 2), ii65–ii72. https://doi.org/http://dx.doi.org/10.1136/bmjqs-2012-001713

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

de Vogel, V., Bruggeman, M., & Lancel, M. (2019). Gender-sensitive violence risk assessment. Predictive validity of six tools in female forensic psychiatric patients. *Criminal Justice and Behavior*, 46(4), 528–549. https://doi.org/10.1177/0093854818824135

de Vogel, V., & de Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law*, 12(3), 321–336. https://doi.org/10.1080/10683160600569029

de Vogel, V., de Vries Robbé, M., van Kalmthout, W., & Place, C. (2014). *Female Additional Manual (FAM): Additional guidelines to the HCR-20V3 for assessing risk for violence in women*. Van der Hoeven Kliniek.

de Vogel, V., van den Broek, E., & de Vries Robbé, M. (2014). The value of the HCR-20^V3 in Dutch forensic clinical practice. *International Journal of Forensic Mental Health*, 13(2), 109–121. https://doi.org/10.1080/14999013.2014.906518

Douglas, K. S., Hart, S. D., Groscup, J. L., & Litwack, T. R. (2014). Assessing violence risk. In I. Weiner & R. K. Otto (Eds.), *The handbook of forensic psychology* (4th ed.). John Wiley & Sons.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk of violence – User guide*. Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S. & Otto, R. K. (Eds.). (2021). *Handbook of violence risk assessment* (2nd ed.). Routledge.

Douglas, K. S., & Shaffer, C. S. (2021). The science of and practice with the HCR-20 V3 (Historical-Clinical-Risk Management – 20, Version 3). In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (2nd ed., pp. 253–293). Routledge.

Douglas, K. S., Shaffer, C., Blanchard, A. J. E., Guy, L. S., Reeves, K., Weir, J. (2017). *HCR-20 violence risk assessment scheme: Overview and annotated bibliography.* https://kdouglas.files.wordpress.com/2017/06/hcr-20-annotated-bibliography-version-13.pdf.

Edens, J. F., & Boccaccini, M. T. (2017). Taking forensic mental health assessment "out of the lab" and into "the real world": Introduction to the special issue on the field utility of forensic assessment instruments and procedures. *Psychological Assessment*, 29(6), 599–610. https://doi.org/10.1037/pas0000475

Edens, J. F., Cox, J., Smith, S. T., DeMatteo, D., & Sörman, K. (2015). How reliable are Psychopathy Checklist-Revised scores in Canadian criminal trials? A case law review. *Psychological Assessment*, 27(2), 447–456. https://doi.org/10.1037/pas0000048

Edmond, G., Cole, S., Cunliffe., & Roberts, A. (2014). Admissibility compared: The reception of incriminating expert evidence (i.e., forensic science) in four adversarial jurisdictions. *University of Denver Criminal Law Review*, 3(1), 31–109.

Ewert v. Canada, 2 S. C. R. 165 (2018). (Can.).

Fazel, S. (2019). The scientific validity of current approaches to violence and criminal risk assessment. In J. W. de Keijser, J. V. Roberts, & J. Ryberg (Eds.), *Predictive sentencing: Normative and empirical perspectives* (pp. 197–212). Bloomsbury Publishing.

Fernandez, Y. M., Harris, A. J. R., Hanson, R. K., & Sparks, J. (2014). *STABLE-2007 coding manual – revised 2014* [Unpublished report]. Public Safety Canada.

Folino, M. E. (2018). *Valoración del riesgo de violencia en pacientes forenses con el HCR-20 V3: análisis de sus propiedades psicométricas* [Doctoral dissertation]. Universidad Nacional de La Plata.

Fujii, D. E. M., Tokioka, A. B., Lichton, A. I., & Hishinuma, E. (2005). Ethnic differences in prediction of violence risk with the HCR-20 among psychiatric inpatients. *Psychiatric Services (Washington, D.C.)*, 56(6), 711–716. https://doi.org/10.1176/appi.ps.56.6.711

Garcia-Mansilla, A., Rosenfeld, B., & Nicholls, T. L. (2009). Risk assessment: Are current methods applicable to women? *International Journal of Forensic Mental Health*, 8(1), 50–61. https://doi.org/10.1080/14999010903014747

Glancy, G. D., & Saini, M. (2009). The confluence of evidence-based practice and Daubert within the fields of forensic psychiatry and the law. *The Journal of the American Academy of Psychiatry and the Law*, 37(4), 438–441.

Green, D., Schneider, M., Griswold, H., Belfi, B., Herrera, M., & DeBlasi, A. (2016). A comparison of the HCR-20^{V3} among male and female insanity acquittees: A retrospective file study. *International Journal of Forensic Mental Health*, 15(1), 48–64. https://doi.org/10.1080/14999013.2015.1134726

Guy, L. (2008). *Performance indicators of the structured professional judgment approach for assessing risk for violence to others: A meta-analytic survey* [Unpublished doctoral dissertation]. Simon Fraser University.

Harris, G. T., & Rice, M. E. (1997). Risk appraisal and management of violent behavior. *Psychiatric Services*, 48, 1166–1176. https://doi.org/10.1176/ps.48.9.1168

Hart, S. D., & Cooke, D. J. (2013). Another look at the (im-) precision of individual risk estimates made using actuarial risk assessment instruments. *Behavioral Sciences & the Law*, 31(1), 81–102. https://doi.org/10.1002/bsl.2049

Heilbrun, K., Newsham, R., & Pietruszka, V. (2016). Risk communication: An international update. In S. Fazel, J. P. Singh, & S. Björkly (Eds.), *International perspectives on risk assessment* (pp. 26–39). Oxford University Press.

Heilbrun, K., Yasuhara, K., Shah, S., & Locklair, B. (2021). Approaches to violence risk assessment. Overview, critical analysis, and future directions. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (2nd ed., pp. 3–27). Routledge.

Helmus, L., Babchishin, K. M., Hanson, R. K., & Thornton, D. (2009). *Static-99R: Revised age weights.* https://www.static99.org

Hilton, N. Z., Scurich, N., & Helmus, L. M. (2015). Communicating the risk of violent and offending behavior: Review and introduction to this special issue. *Behavioral Sciences & the Law*, 33(1), 1–18. https://doi.org/10.1002/bsl.2160

Hogan, N. R., & Olver, M. E. (2016). Assessing risk for aggression in forensic psychiatric inpatients: An examination of five measures. *Law and Human Behavior*, 40(3), 233–243. https://doi.org/10.1037/lhb0000179

Hogan, N. R., & Olver, M. E. (2019). Static and dynamic assessment of violence risk among discharged forensic patients. *Criminal Justice and Behavior*, 46(7), 923–938. https://doi.org/10.1177/0093854819846526

Hopton, J., Cree, A., Thompson, S., Jones, R., & Jones, R. (2018). 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), 195–201. https://doi.org/10.1080/14999013.2018.1460424

Hounsome, J., Whittington, R., Brown, A., Greenhill, B., & McGuire, J. (2018). The structured assessment of violence risk in adults with intellectual disability: A systematic review. *Journal of Applied Research in Intellectual Disabilities: JARID*, 31(1), e1–e17. https://doi.org/10.1111/jar.12295

Judges, R., Egan, V., & Broad, G. (2016). A critique of the Historical Clinical Risk–20, Version 3, risk assessment instrument. *Journal of Forensic Psychology Practice*, 16(4), 304–320. https://doi.org/10.1080/15228932.2016.1196102

Kamorowski, J., de Ruiter, C., Schreuder, M., Jelicic, M., & Ask, K. (2021). *Negative pretrial publicity, evaluator attitudes toward offenders, and risk assessment with the HCR-20V3* [Manuscript submitted for publication]. Department of Forensic Psychology, Maastricht University.

Kamorowski, J., Ask, K., Schreuder, M., Jelicic, M., & de Ruiter, C. (2021). "He seems odd": The effects of risk-irrelevant information and actuarial risk estimates on mock jurors' perceptions of sexual recidivism risk. *Psychology, Crime & Law*, 1–30. https://doi.org/10.1080/1068316X.2021.1909016

Logan, C. (2014). The HCR-20 Version 3: A case study in risk formulation. *International Journal of Forensic Mental Health*, 13(2), 172–180. https://doi.org/10.1080/14999013.2014.906516

Lowder, E. M., Morrison, M. M., Kroner, D. G., & Desmarais, S. L. (2019). Racial bias and LSI-R assessments in probation sentencing and outcomes. *Criminal Justice and Behavior*, 46(2), 210–233. https://doi.org/10.1177/0093854818789977

Murray, J., Charles, K. E., Cooke, D. J., & Thomson, M. E. (2014). Investigating the influence of causal attributions on both the worksheet and checklist versions of the HCR-20. *International Journal of Forensic Mental Health*, 13(1), 8–17. https://doi.org/10.1080/14999013.2014.890978

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science*, 24(10), 1889–1897. https://doi.org/10.1177/0956797613481812

Neal, T. M. S., & Brodsky, S. (2016). Forensic psychologists' perceptions of bias and potential correction strategies in forensic mental health evaluations. *Psychology, Public Policy, and Law*, 22(1), 58–76. https://doi.org/10.1037/law0000077

Neal, T. M. S., & Grisso, T. (2014). The cognitive underpinnings of bias in forensic mental health evaluations. *Psychology, Public Policy, and Law*, 20(2), 200–211. https://doi.org/10.1037/a0035824

Neal, T. M. S., Slobogin, C., Saks, M. J., Faigman, D. L., & Geisinger, K. F. (2019). Psychological assessments in legal contexts: Are courts keeping "junk science" out of the courtroom? *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, 20(3), 135–164. https://doi.org/10.1177/1529100619888860

Neil, C., O'Rourke, S., Ferreira, N., & Flynn, L. (2020). Protective factors in violence risk assessment: Predictive validity of the SAPROF and HCR-20^{V3}. *International Journal of Forensic Mental Health*, 19(1), 84–102. https://doi.org/10.1080/14999013.2019.1643811

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, 26(1), 156–176. https://doi.org/10.1037/a0035080

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*(2), 255–270. https://doi.org/10.1016/j.avb.2012.11.016

O'Shea, L. E., Picchioni, M. M., McCarthy, J., Mason, F. L., & Dickens, G. L. (2015). Predictive validity of the HCR-20 for inpatient aggression: The effect of intellectual disability on accuracy. *Journal of Intellectual Disability Research*, *59*(11), 1042–1054. https://doi.org/10.1111/jir.12184

Penney, S. R., Marshall, L. A., & Simpson, A. I. (2016). The assessment of dynamic risk among forensic psychiatric patients transitioning to the community. *Law and Human Behavior*, *40*(4), 374–386. https://doi.org/10.1037/lhb0000183

Persson, M., Belfrage, H., Fredriksson, B., & Kristiansson, M. (2017). Violence during imprisonment, forensic psychiatric care, and probation: Correlations and predictive validity of the risk assessment instruments COVR, LSI-R, HCR-20^{V3}, and SAPROF. *International Journal of Forensic Mental Health*, *16*(2), 117–129. https://doi.org/10.1080/14999013.2016.1266420

Schneider, R. D. (2016). Expert evidence: Judge as gatekeeper. In C. Pakosh (Ed.), *The Lawyer's Guide to the Forensic Sciences* (pp. 669–718). Irwin Law Inc.

Scurich, N. (2018). The case against categorical risk estimates. *Behavioral Sciences & the Law*, *36*(5), 554–564. https://doi.org/10.1002/bsl.2382

Shepherd, S. M., Adams, Y., McEntyre, E., & Walker, R. (2014). Violence risk assessment in Australian Aboriginal offender populations: A review of the literature. *Psychology, Public Policy and Law*, *20*(3), 281–293. https://doi.org/10.1037/law0000017

Shepherd, S. M., & Spivak, B. (2021). Finding colour in conformity part II-Reflections on structured professional judgement and cross-cultural risk assessment . *International Journal of Offender Therapy and Comparative Criminology*, *65*(1), 92–99. https://doi.org/10.1177/0306624X20928025

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. https://doi.org/10.1023/A:1005575113507

Singh, J. P., Björkly, S., & Fazel, S. (Eds.). (2016). *International perspectives on risk assessment*. Oxford University Press.

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., Doyle, M., Folino, J. O., Godoy-Cervera, V., Grann, M., Ho, R. M. Y., Large, M. M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., de Ruiter, C., Seewald, K., & Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health*, *13*(3), 193–206. https://doi.org/10.1080/14999013.2014.922141

Singh, J. P., Grann, M., & Fazel, S. (2013). Authorship bias in violence risk assessment? A systematic review and meta-analysis. *Plos One*, *8*(9), e72484. https://doi.org/10.1371/journal.pone.0072484

Slobogin, C. (2020). Introduction to the special issue on implementing post-conviction risk assessment. *Behavioral Sciences & the Law*, *38*(3), 187–192. https://doi.org/10.1002/bsl.2466

Slobogin, C. (2021). Constitutional and evidentiary issues concerning risk assessment. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (2nd ed., pp. 70–90). Routledge.

Smith, S. T., Kelley, S. E., Rulseh, A., Sörman, K., & Edens, J. F. (2014). Adapting the HCR-20^{V3} for pre-trial settings. *International Journal of Forensic Mental Health*, *13*(2), 160–171. https://doi.org/10.1080/14999013.2014.906520

Snowden, R. J., Gray, N. S., & Taylor, J. (2010). Risk assessment for future violence in individuals from an ethnic minority group. *International Journal of Forensic Mental Health*, *9*(2), 118–123. https://doi.org/10.1080/14999013.2010.501845

Spivak, B. L., & Shepherd, S. M. (2020). Machine learning and forensic risk assessment: New frontiers. *Journal of Forensic Psychiatry and Psychology*, *31*(4), 571–581. https://doi.org/10.1080/14789949.2020.1779783

Storey, J. E., Campbell, V. J., & Hart, S. D. (2013). Expert evidence about violence risk assessment: A study of Canadian legal decisions. *International Journal of Forensic Mental Health*, *12*(4), 287–296. https://doi.org/10.1080/14999013.2013.867383

Storey, J. E., Watt, K. A., & Hart, S. D. (2015). An examination of violence risk communication in practice using a structured professional judgment framework. *Behavioral Sciences & the Law*, *33*(1), 39–55. https://doi.org/10.1002/bsl.2156

Strub, D. S., Douglas, K. S., & Nicholls, T. L. (2014). The validity of Version 3 of the HCR-20 violence risk assessment scheme amongst offenders and civil psychiatric patients. *International Journal of Forensic Mental Health*, *13*(2), 148–159. https://doi.org/10.1080/14999013.2014.911785

Viljoen, J. L., & Vincent, G. M. (2020). Risk assessments for violence and reoffending: Implementation and impact on risk management. *Clinical Psychology: Science and Practice*, e12378. https://doi.org/10.1111/cpsp.12378

Vincent, G. M., Guy, L. S., & Grisso, T. (2012). *Risk assessment in juvenile justice: A guidebook for implementation*. John D. & Catherine T. MacArthur Foundation. http://modelsforchange.net/publications/346

Vitacco, M. J., Balduzzi, E., Rideout, K., Banfe, S., & Britton, J. (2018). Reconsidering risk assessment with insanity acquittees. *Law and Human Behavior*, *42*(5), 403–412. https://doi.org/10.1037/lhb0000298

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing the risk of violence. Version 2*. Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Wilson, H. A., & Gutierrez, L. (2014). Does one size fit all? A meta-analysis examining the predictive ability of the Level of Service Inventory (LSI) with Aboriginal offenders. *Criminal Justice and Behavior*, *41*(2), 196–219. https://doi.org/10.1177/0093854813500958

Wormith, J. S., & Bonta, J. (2021). Risk/need assessment for adults and older adolescents. The Level of Service (LS) instruments. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (2nd ed., pp. 159–190). Routledge.

Zapf, P. A., & Dror, I. E. (2017). Understanding and mitigating bias in forensic evaluation: Lessons from forensic science. *International Journal of Forensic Mental Health*, *16*(3), 227–238. https://doi.org/10.1080/14999013.2017.1317302

Zhou, J., Witt, K., Xiang, Y., Zhu, X., Wang, X., & Fazel, S. (2016). Violence risk assessment in psychiatric patients in China: A systematic review. *The Australian and New Zealand Journal of Psychiatry*, *50*(1), 33–45. https://doi.org/10.1177/0004867415585580