Original Article

Concurrent Validity and Clinical Utility of the HCR-20^{V3} Compared With the HCR-20 in Forensic Mental Health Nursing: Similar Tools but Improved Method

AQ1 Stål Bjørkly, PhD[1,2], Gunnar Eidhammer[2,3], and Lars Erik Selmer[2,3]

ABSTRACT

The main scope of this small-scale investigation was to compare clinical application of the HCR-20^{V3} with its predecessor, the HCR-20. To explore concurrent validity, two experienced nurses assessed 20 forensic mental health service patients with the tools. Estimates of internal consistency for the HCR-20 and the HCR-20^{V3} were calculated by Cronbach's alpha for two levels of measurement: the H-, C-, and R-scales and the total sum scores. We found moderate (C-scale) to good (H- and R- scales and aggregate scores) estimates of internal consistency and significant differences for the two versions of the HCR. This finding indicates that the two versions reflect common underlying dimensions and that there still appear to be differences between V2 and V3 ratings for the same patients. A case from forensic mental health was used to illustrate similarities and differences in assessment results between the two HCR-20 versions. The case illustration depicts clinical use of the HCR-20^{V3} and application of two structured nursing interventions pertaining to the risk management part of the tool. According to our experience, Version 3 is superior to Version 2 concerning (a) item clarity, (b) the distinction between presence and relevance of risk factors, (c) the integration of risk formulation and risk scenario, and (d) the explicit demand to construct a risk management plan as part of the standard assessment procedure.

KEY WORDS:

clinical utility; concurrent validity; HCR-20^{V3}; violence risk

Assessment and management of a patient's[1] risk for violence toward others during a stay and after discharge from a psychiatric hospital is an important challenge for clinicians. Different strategies and tools have been developed to evaluate risk, embracing three main approaches: unstructured clinical judgment, structured professional judgment (SPJ), and actuarial assessment (e.g., Douglas, Hart, Webster, & Belfrage. 2013). Whereas the unstructured strategy does not presuppose a set of risk factors to be assessed, the actuarial approach has a list of risk factors and a scoring method and generates a probability estimate of risk contingent on the scores. The SPJ tradition is similar to the actuarial methods in providing a list of risk factors to be assessed. However, the SPJ strategy differs markedly by allowing the clinician to individualize risk judgment without predefined decision rules. This has a particular relevance to forensic mental health nursing because the emphasis on recovery may be a particular challenge because of the coercive and restrictive nature of forensic mental health services, especially in terms of patient involvement and social inclusion (Pouncey & Lukens, 2010). In particular, the strategy guides nurses to implement individualized

Author Affiliations: *[1]Faculty of Health and Social Sciences, Molde University College; [2]Centre for Research and Education in Forensic Psychiatry, Oslo University Hospital; and [3]Division of Mental Health and Drug Abuse, Vestre Viken Trust.*

The authors declare no conflict of interest.

Correspondence: *Stål Bjørkly, PhD, Faculty of Health and Social Sciences, Molde University College, P.O. Box 2110, NO-6402 Molde, Norway. Tel: +47 478 498 00; Fax: +47 71 21 40 50; E-mail: stal.bjorkly@himolde.no.*

Received March 19, 2014; accepted for publication September 18, 2014.

Copyright © 2014 International Association of Forensic Nurses

DOI: 10.1097/JFN.0000000000000047

[1]Although the term consumer is frequently used in contemporary practice, the term patient is used here as this is consistent with the language used in Norway.

interventions based on risk formulation and risk scenario planning to achieve risk reduction, recovery, user involvement, and personal growth.

The HCR-20 belongs to the SPJ category. It has been in clinical use in Norway since the first version was implemented in 1995 (Webster, Eaves, Douglas, & Wintrup, 1995). However, the big breakthrough came with the Norwegian translation of the second version (Rasmussen, Jakobsen, & Urheim, 2002; Webster, Douglas, Eaves, & Hart, 1997). The emphasis on interprofessional cooperation to enhance risk assessment and management has opened up for nurses to become more significant contributors in this field of clinical work. The psychometric properties of the HCR-20 have been documented to be very good, and its clinical relevance has been supported by its extensive use across the world (e.g., Fazel, Singh, Doll, & Grann, 2012; Singh, Grann, & Fazel, 2011). Still, based on many years of research and feedback from clinicians and researchers, the authors of the HCR-20 recently developed a revised and hopefully improved version, the HCR-20^{V3} (Douglas et al., 2013). The HCR-20^{V3} has a seven-step structure to guide risk assessment and the development of treatment and management strategies to mitigate risk: (a) to gather information; (b) to assess the presence of each of the 20 risk factors; (c) to decide the relevance of the risk factor for a given individual (see the case illustration for further elaboration on the difference between presence and relevance assessment); (d) to develop a risk formulation by establishing a theoretical and empirical understanding of the individual's violence risk; (e) to make individualized risk scenario planning; (f) to develop risk management plans and interventions; and (g) to document summary judgments in a clear, simple manner that facilitates appropriate action. Whereas Version 2 was scored by numbers (0, 1, and 2) to assess the risk level, the HCR-20^{V3} is coded by using letters: "y" (yes), "p" (partially/possibly), "n" (no), or "o" (omit). Compared with Version 2, seven items have been broadened, four items have been narrowed, and one item has been added in the HCR-20^{V3} (for more detailed information about the HCR-20^{V3}, see Douglas et al., 2013).

According to Campbell and Fiske (1959), there are two main issues to take into account when revising a test, scale, or, as is the case with the HCR-20, a clinical guideline: (a) making it better and (b) avoiding constructing a completely different guideline. They stressed the importance of using both discriminant and concurrent validation techniques when assessing a new or revised test, scale, or guideline. In short, a revised version should be different but still more similar than different from the original. This means that there should be a significant association between the two versions in terms of joint internal consistency of ratings, or put in another way, the tools must have good concurrent validity. Concurrent validity applies to validation studies in which two measures are administered approximately at the same time. This article reports findings from a small-scale test of the concurrent validity of the HCR-20 and the HCR-20^{V3} and a case-based comparison of the clinical utility and relevance of the two versions for forensic mental health nursing.

Method

Setting

The ratings were carried out in a medium security forensic mental health unit at the Clinic of Addiction and Psychiatry, Vestre Viken Trust, in Norway. The unit has a catchment area of 250,000 inhabitants and covers urban and rural areas. The study is a part of the SAFE pilot project that was approved by the Norwegian Social Science Data Services and the Regional Committee for Medical Research Ethics. The approval granted exemption from asking for persons' consent to participate in the study.

Participants and Study Design

The investigation was conducted with an observational cross-sectional design. Two mental health nurses compared the second HCR-20 (V2) with the third version (HCR-20^{V3}, V3) in an assessment of 20 forensic mental health service consumers. The raters had extensive experience in forensic mental health care and further education in violence risk assessment and management and had used V2 for over 10 years. Consumer characteristics are presented in Table 1. T1

Procedure

The SPSS SamplePower 3 was used to estimate statistical power (SP). The analysis showed that a sample size of 20 would yield a power of 68.5% to test the null hypothesis that the population correlation is zero with alpha set at .05. A convenience sample of 20 patients was selected. The raters assessed 10 patients each. After completion of a V2 assessment, the same procedure was followed for a V3 assessment of the same person. The V2 was scored numerically according to the old procedure—2, 1, 0, or nk (*not known*)—whereas the V3 was coded with letters—"y" (*yes*), "p" (*probably/possibly*), "n" (*no*), or "o" (*omit*). Assessment data were gathered from multiple sources: patient files, observations, and consultations with colleagues who knew the patients well. All information was obtained, and every rating was completed independently. One of the patients in the sample was selected as a case illustration in this article. The case illustration depicts the clinical application of the HCR-20^{V3}, by using the seven steps as basis for implementing two structured nursing-driven interventions.

Data Analyses

To obtain data for statistical analysis, a common score scale of 0, 1, and 2 was chosen to transform V3 ratings into scores. This procedure was chosen after consulting with the first author of the HCR-20^{V3}. The statistical analyses were conducted in line with Campbell and Fiske's (1959)

TABLE 1. Survey of Diagnosis and Most Serious Violence (*N* =20 Patients)

Diagnosis	Most serious violence
Paranoid schizophrenia	Homicide
Paranoid schizophrenia	Homicide
Paranoid schizophrenia	Homicide
Paranoid schizophrenia	Severe violence
Paranoid schizophrenia	Severe violence
Paranoid personality disorder	Double homicide
Antisocial personality disorder	Violence and threats
Psychosis because of substance abuse	Arson and homicidal threats
Schizophrenia	Severe violence
Schizophrenia	Violence and threats
Psychosis because of substance abuse	Homicidal threats
Paranoid schizophrenia	Violence and homicidal threats
Antisocial personality disorder, psychosis	Violence and threats
Paranoid personality disorder	Violence
Psychosis because of substance abuse	Violence and homicidal threats
Paranoid schizophrenia	Violence
Schizophrenia	Knife stabbing
Schizophrenia	Rape
Paranoid schizophrenia	Arson
Schizophrenia	Violence in institution

recommendations: to test (a) concurrent validity and (b) discriminant validity. (a) Internal consistency estimates for the HCR-20 and the HCR-20^{V3} were calculated by Cronbach's alpha (Cronbach, 1951). We calculated alpha values for two levels of measurement: the three HCR-20 and HCR-20^{V3} parts (historical, 10 items; clinical, five items; and risk management, five items) and the total sum scores (20 items) for the two HCR-20 versions. This yielded a total number of 400 scores (20 patients × 2 HCR-20 versions × 10 items) for the H-part, 200 scores for the C and R parts, and 800 scores for the total sum. Because the presence and relevance of a risk factor are rated separately on the HCR-20^{V3}, Cronbach's alpha was calculated for both. Pearson's *r* was used to calculate correlation coefficients for the association between the two versions. (b) A one-sample Kolmogorov–Smirnov test was used to control for normal distribution of the scores in the H, C, and R parts and for the total sum scores. This was done to decide whether a parametric or nonparametric test should be used for analyzing sum score differences and correlations between V2 and V3. Because the Kolmogorov–Smirnov test showed that the scores had normal distribution (see Table 3), we conducted a paired sample *t* test to compare V2 and V3 sum scores for the H, C, and R parts and for the total sum scores. A conventional 5% significance level and 95% confidence interval were used. Cohen's *d* was applied to estimate the effect sizes of the differences between scores on the V2 and V3. The SPSS SamplePower 3 was used in post hoc SP analyses for the *t* test estimates. All other statistical analyses were conducted using the statistical program package SPSS Version 16.0 for Windows.

Results

Concurrent Validity, Correlations, and Differences Between V2 and V3 Sum Scores

The principal aim was to compare the concurrent validity of V2 and V3 scores. The results are presented in Table 2. T2 We found moderate (C items) to good (H and R items and aggregate scores) estimates of internal consistency for the two versions of the HCR. According to Nunnally and Bernstein (1994), for the early stages of research, alpha values of about .60 and .80 reflect moderate and good internal consistencies, respectively. Our finding indicates that the two versions reflect common underlying dimensions and that there still appear to be differences between the V2 and V3 ratings of the same patients. The sum scores for the H, C, R, and aggregate scores were normally distributed for both the V2 and V3 (see Table 3). T3

Total score correlations. There was a significant Pearson product–moment correlation between V2 and V3 concerning total scores for presence ($r(20) = .58$, $p = 0.007$) and relevance ($r(20) = .55$, $p = 0.007$). We conducted a paired sample *t* test to estimate differences between V2 and V3 sum scores.

Presence scores. There were significant score differences between the two versions for H items ($t(19) = -2.80$, $p = 0.012$, Cohen's $d = 0.37$, SP = 81%), C items ($t(19) = -4.04$, $p = 0.001$, Cohen's $d = 1.09$, SP = 98%), and total score ($t(19) = -3.11$, $p = 0.006$, Cohen's $d = 0.65$, SP = 84%), but not for R items ($t(19) = 0.22$, $p = 0.830$, Cohen's $d = 0.05$, SP = 6%). There was a marked mean score difference for two items in the R part. However, a 30% lower estimated risk in V3 on R3 (personal support) and a 37% increased risk rated on R4 (treatment or

TABLE 2. Internal Consistency (Cronbach's Alpha With 95% Confidence Intervals) of Ratings of the Same Patients (*N* = 20) With Versions 2 and 3 of the HCR-20 (Presence and Relevance Sum Scores)

	Presence	Relevance
H-part sum scores	.85 (0.73–0.94)	.85 (0.73–0.94)
C-part sum scores	.59 (0.26–0.81)	.58 (0.25–0.81)
R-part sum scores	.81 (0.66–0.91)	.79 (0.63–0.90)
Total sum scores	.84 (0.71–0.94)	.84 (0.71–0.93)

TABLE 3. Kolmogorov–Smirnov Test for Normal Distribution of Scores in Versions 2 and 3 of the HCR-20

HCR-20	Scores	*M*	SD	*p*2[1]
Version 2	Historical	12.10	3.92	0.46
	Clinical	4.90	1.41	0.60
	Risk management	5.80	1.85	0.46
	Total sum	22.80	4.63	0.46
Version 3: presence	Historical	13.45	3.28	0.53
	Clinical	6.35	1.23	0.67
	Risk management	5.70	2.00	0.14
Version 3: relevance	Total sum	25.50	3.61	0.36
	Historical	13.10	3.39	0.73
	Clinical	7.25	1.41	0.33
	Risk management	7.20	2.04	0.13
	Total sum	27.55	3.93	0.78

[1] $p > 0.05$; the distribution of scores is normal.

AQ2

supervision response) appeared to level the difference between sum scores for R items out.

Relevance scores. There were significant score differences between the two versions for H items ($t(19) = -2.08$, $p = 0.052$, Cohen's $d = 0.27$, SP = 49%), C items ($t(19) = -5.53$, $p \leq 0.001$, Cohen's $d = 1.67$, SP = 99%), R items ($t(19) = -2.75$, $p = 0.013$, Cohen's $d = 0.88$, SP = 73%), and total score ($t(19) = -5.16$, $p = 0.001$, Cohen's $d = 1.10$, SP = 99%).

Case Illustration

John is a man in his mid-30s with a diagnosis of paranoid schizophrenia. He was raised in several foster homes. His father was mostly absent and neglected John, and his mother experienced a major mental disorder and was frequently hospitalized. When he was 10 years old, his parents divorced. He had early problems in school because of impaired concentration and a very early start of fighting and bullying. During his adolescence, he was involved in several types of crime and antisocial activities. His first serious violent episode occurred at the age of 16 years when he stabbed a friend with a knife. The following year, he inflicted bodily harm on strangers and family members. When he was 17 years old, he entered a right-wing extremist group and also served as money collector for a group of organized criminals. John's criminal record covers a variety of criminal behaviors, and most of his peers also engaged in criminal or antisocial activities. He has practically no work experience, and since the age of 19 years, he has mostly been incarcerated or in a psychiatric hospital.

According to John, he experienced symptoms of mental illness for as long as he could remember: "I've been restless all my life, and I'm convinced that I have ADHD. I do not trust anyone, and I must admit that I become paranoid when I take 'speed.'" He reported that he sometimes heard voices in his head while on drugs and that this started when he was a teenager. During his adolescence, he was in and out of short-term psychiatric treatment because of acute symptoms after heavy use of alcohol and drugs. In his early 20s, he killed a man while under the influence of amphetamines and alcohol. The victim was a "friend." According to John, the motive for killing his friend was that the victim had harassed him over a long time. John was first sentenced to prison, but after a brief stay, he was transferred to a forensic mental health unit for mandated treatment.

Using this case illustration, we report on HCR-20 evaluations of John with V2 and V3. The aims of the case illustration are (a) to compare the evaluations concerning risk factors (items) and judgment procedures and (b) to especially focus on the extent to which V2 and V3 inform and instruct individualized risk management strategies. John's scores are presented in Table 3, and the presentation parallels the seven-step procedure of the HCR-20[V3].

Step 1: Gathering data. See the "Procedure" section above for an overview of the sources of information the evaluators used.

Steps 2 and 3: Risk factor (item) evaluation. Only minor differences, manifested as higher scores on H7 and H9, appear when comparing John's V2 scores with V3 presence scores. However, substantial differences emerge when we focus on relevance evaluations. Three historical factors were assessed to be of moderate relevance, indicating that the impact of these past risk factors was only partially or probably relevant to current risk. A high relevance rating indicates that the factor is relevant to the development and maintenance of risk management strategies. This means that a person with a history of violence under the influence of drugs may have a low coding on presence and still have a high relevance code because he, at the time of assessment, is in an efficient drug relapse prevention program. The high relevance code reflects that it is essential to maintain this risk management approach to mitigate future risk for violence. Hypothetically, if a person has a drug problem but only is violent when sober, drug abuse will be coded to be present, but not as a relevant risk factor for violence. This may contribute to a more balanced risk evaluation and reduce rates of false positives. Apart from C4, all clinical and risk management items were rated to be of high relevance to the development of risk management strategies. This information invites a many-faceted, dynamic understanding of John's life challenges and risk for violence. In our opinion, assessing relevance is a substantial development in the V3. It may help the raters to comprehend the person's background and history more dynamically in the sense that a positive identification of a historical risk factor does

AQ3 **TABLE 4. John's Scores[1] on Versions 2 and 3 of the HCR-20**

	V2	V3	
		Presence	Relevance
H1:[2] violence	2	2	2
H2: antisocial behavior	2	2	2
H3: relationships	2	2	1
H4: employment	2	2	1
H5: substance use	2	2	2
H6: major mental disorder	2	2	2
H7: personality disorder	0	1	2
H8: traumatic experiences	2	2	1
H9: violent attitudes	1	2	2
H10: treatment or supervision response	2	2	2
C1:[3] insight	1	1	2
C2: violent ideation or intent	1	1	2
C3: symptoms of major mental disorder	2	2	2
C4: instability	1	1	1
C5: treatment or supervision response	1	1	2
R1:[1] professional services	0	0	2
R2: living situation	1	1	2
R3: personal support	1	1	2
R4: treatment or supervision response	0	0	2
R5: stress or coping	1	1	2
Sum score	26	28	36

AQ2 [1]Scores for the HCR-20[V3]; only used for presentation clarity (0 = *not present/not relevant*, 1 = *possibly present/moderate relevance*, 2 = *present/high relevance*).
[1]Historical items.
[2]Clinical items.
[3]Risk management items.

not necessarily mean that it is a valid risk factor in the here and now.

Steps 4, 5, and 6: Risk formulation, risk scenario, and management strategies. The relevance scores contribute to further individualize the clinical assessment and, most importantly, to guide evaluators in developing and implementing individualized risk management plans to mitigate risk factors and triggers. Naturally, this is not a brand new invention emanating from the HCR-20 tool alone. During the last couple of decades, this approach has emerged as a mainstream component in numerous risk assessment tools within the SPJ tradition. Of course, users were advised to address these issues in V2 as well. However, by introducing Steps 4–6 as part of a standard procedure, V3 guides and guards that this actually will be done. We believe that, if the structured procedure of the HCR-20[V3] is established in risk assessment practice, there will be an increase in demands from colleagues, courts, and other kinds of experts to include recommendations for individualized risk management strategies in risk assessment reports.

We selected the following out of John's most relevant risk factors to illustrate the rationale and use of risk formulation: H1, history of problems with violence; H5, history of problems with substance use; C1, recent problems with insight; C2, recent problems with violent ideation or intent; C3, recent problems with symptoms of major mental disorder; R2, future problems with living situation; and R4, future problems with treatment or supervision response. We want to emphasize that the clinical relevance of historical risk factors rests on evidence of their current impact on risk for violence. If a person has successfully gone through trauma therapy, the relevance of early childhood abuse (H8: traumatic experiences) as a historical risk factor is reduced. Formulation 1 refers to reactive violence in response to minor or idiosyncratic perceptions of provocations. Precipitating events are insufficient external boundaries: changes in living environment (reduced predictability) and drug abuse. It is a psychosis-driven, cognitive regulation impairment characterized and reinforced by lack of insight, persecutory delusions, and auditory command hallucinations. Emotional distress in the form of severe anxiety attacks functions as a catalyst of violent behavior, and it is further strengthened by drug and alcohol intoxication. Violence appears to be reinforced by temporary relief from psychotic symptoms and concomitant emotional distress.

Formulation 2 is of a proactive violent nature but has two characteristics in common with the first formulation. Substance abuse plays a similar role as in the first formulation, but even if persecutory symptoms are also present, they are not so dominant in John's proactive violence. The main precipitant is free and unsupervised interaction with criminal peers. Their influence on John is enhanced by his low self-esteem and impaired self-control. This dynamic of violence has played a paramount role in several incidents of violent crime. The main reinforcements appear to be social peer acceptance and monetary gain.

Whereas risk formulations help evaluators to apprehend and conceptualize individual causal factors of violence, risk scenarios (Step 5) are intended to guide them into risk management planning. In John's case, risk management strategies had to address two different but interrelated risk formulations. The first one was characterized by exacerbation of psychotic symptoms and concomitant anxiety; the other one was associated with his long-lasting criminal career. Two risk scenarios emerge from these formulations. In the first psychosis-driven scenario, John has no professional follow-up, he starts using drugs, he develops severe

persecutory symptoms with concomitant impaired insight and anxiety attacks, and some kind of relationship change occurs. The second scenario starts like the first one with supervision failure; this allows for contact from criminal peers whom he actually fears, he gets easy access to drugs, and he is paying for the drugs by committing violent crimes.

To illustrate Step 6 in our HCR-20^{V3} evaluation of John, we will very briefly sketch some main points in a hypothetical and yet conceivable risk management plan. Two related approaches to develop individualized risk management strategies are suggested as being feasible in this process: the Progression Ladder (ProLad; Bjørkly, 2004) and the Early Recognition Method (ERM; Fluttert et al., 2008). ProLad and ERM are based on principles from cognitive behavior therapy (CBT), and both are structured approaches designed for gaining insight into coping strategies that may enhance persons' abilities to identify and manage their precursors of aggression and stressful situations. The main aim is to enhance personal self-control and growth. ProLad is a multifocused intervention strategy with both violence-specific and violence-related treatment targets. Individual warning signs and risk factors for violence are examples of the violence-specific category. ProLads that target risk factors are based on the assessment of individual stress–vulnerability interactions such as exposure to physical contact from strangers, limit setting, certain persons, drug abuse, and so forth (Bjørkly, 1994). To identify risk situations and warning signs in detail and to minimize those by means of systematic enhancement of individual self-management are violence-specific ProLad and ERM features. Training in social skills and basic living skills are examples of violence-related treatment targets. The treatment is provided in natural situations within different types of milieu treatment contexts (e.g., hospital and community-based settings). In John's case, the ERM may be used to develop his coping skills in tense and violence-escalating situations (Formulation 1 concerning H1, H5, C1, C2, and C3), and the ProLad may be implemented to strengthen his social skills by developing a structured, stepwise plan for his reintegration into community living (Formulation 2 concerning H1, H5, C1–C3, R2, and R4).

A brief illustration of the ProLad approach in John's risk management plan. John is in need of enhanced social skills to achieve personal growth and protection against relapse into further contact with criminal peers. His starting point is at a forensic mental health ward, so it is evident that his integration into the community is a long-term project. His first ProLad aim is being alone outside the ward. A ladder with 12 steps is constructed. The first step is being alone 1 minute outside the entrance door of the ward. To move on to the next step, he must meet the criteria for this step in 14 days. The next steps are 2, 3, and 4 minutes before moving on to walk a fixed route that takes 5 minutes; after that, free movement for 5 minutes; and so forth. At the final step of this ladder, John must cope with 1 hour unaccompanied by nursing staff within the hospital area. If he fails at one step, he will have to stay on that step until he succeeds for 14 days. If he absconds, uses drugs, or is violent, he will have to start from the beginning. The next ladder is even more challenging for John. The aim is to manage to go to the public indoor swimming pool by bus, to stay alone there for 1 hour, and finally, to go back to the ward by bus. The first step is to do the whole behavior chain together with a nurse, except for the walk from the bus stop to the ward. The next step is to go by bus from the public pool and to walk to the ward alone and so forth. The criterion for a new step is to succeed five subsequent times on each step. A number of other ProLads may follow to address a variety of aims like sheltered work, shopping alone in a nearby city center, joining a leisure group for persons living in the community, and so forth. The final ladder is used to make John's move to a 24-hours-a-day staffed, sheltered collective. John's problems with trusting people who he does not know well may be mitigated by a stepwise reduction in overlap between personnel from the collective and forensic mental health ward. In Step 1, John stays 1 hour together with the new personnel, then 2 hours, and up until the forensic mental health nurses withdraw from the house with a telephone back-up option in case John shows warning signs of coping failure.

A brief illustration of the ERM approach in John's risk management plan. Persecutory delusions and concomitant anxiety attacks are major obstacles to successful coping in John's life. During the work with the ERM protocol, John reports that he feels helpless in these situations and that he always seems to "submit to an uncontrollable force." He describes a sequence of escalating warning signs that start with a feeling of "inner unrest." If John does not cope with this, it develops into hearing condescending voices, and finally, he becomes convinced that he is in imminent danger of being killed. John says that he depends on the nurses to initiate the first ERM intervention step. To manage this, the nurses must know observable warning signs and an intervention strategy that John believes in and feels comfortable with. John says that, when he feels inner unrest, he cannot sit still and starts drifting around in the ward. They agree to use this specific warning sign as a trigger for the nurses to invite John to participate in a chosen physical activity, such as, for example, going for a walk or playing table tennis. Provided that John makes progress, the next two interventions are geared to help John take over the responsibility for initiating interventions. First, this may be done by nurses prompting the option of some kind of intervention by saying, for instance, that he appears restless and asking if there is anything they can do for him. In the last phase, John is responsible, himself, for initiating the intervention. Continuous cooperation on further development

of the intervention plan and daily feedback to John are important facets of the ERM.

Notably, John's drug abuse is one of the most crucial challenges to his future progress, and therefore, efficient relapse prevention is pivotal for his success. This may be addressed by regular blood and hair sample tests implemented as part of a ProLad and by an ERM intervention, similar to the one described concerning persecutory delusions. In this context, John learns to identify early warning signs of drug craving and, subsequent to this, to ask nursing personnel for help. John's coping with the risk management strategies described above serves three important risk assessment functions. First, it relates the risk assessment to a certain context as opposed to an open, unspecified living situation; second, it provides important information about possible changes in dynamic risk factors (e.g., C1, C2, and C3); and, third, John's responses to the plan help inform future risk assessment and what kind of future living context will be best for him.

Discussion

The main scope of this investigation was to obtain a preliminary picture of the clinical utility and quality of the HCR-20^{V3}. We did this by (a) evaluation of the association between V2 and V3 assessments in a sample of 20 forensic mental health service consumers (group level) and (b) a more detailed comparison of the assessments of one person (individual level). Our findings must be interpreted within the limitations set by the small sample size and the convenience-based selection of persons and evaluators.

Concurrent Validity

Our findings indicate that the concurrent validity of the second and third versions of the HCR-20 is significant and solid. The correlations between the total scores of the two versions were .55 (presence scores) and .58 (relevance scores). This is actually a good result when taking into consideration that seven items were broadened, four were narrowed, and a new item was included in V3. The internal consistency between the versions was very good for the H items, even if the sum score of H items increased by 11% when V3 was used. It is noteworthy that almost 60% of this increase pertained to the H7 item, which has been expanded from being limited to psychopathy in V2 to cover problems with personality disorder in the recent version. There were only four other H items that yielded a substantial change in scores. The alpha value of the C items had markedly lower values than those found for the H and R parts. This was reflected in the finding that the highest and most general increase in scores was found for the C items (mean increase of 30% and a range from 11% to 59%). The instability item (C4) and recent problems with treatment or supervision response (C5) accounted for almost 70% of the increased variance. One may hypothesize that this is because of a more accurate description of indicators in the V3. The internal consistency value for the R items was very good. However, even if there was only a minor average change (2% decrease in scores) from V2 to V3, there were two important changes at a single-item level. Parallel to the results for C4, there was a marked increase in score values for R4 (37%) that elicits information concerning future problems with treatment or supervision response. This may be because of improved accuracy of explanation and/or widened content of this item in V3. However, this score increase was almost balanced out by a reduction of score values of 30% for R3, future problems with personal support. This result is not very easy to interpret, but it may invite further elaboration on a possible relationship between the assessed impact of private and professional networks in risk management plans. Results from the eight tests of differences between V2 and V3 sum scores must be interpreted with caution. Although six of the tests showed significant differences, only the results for C items and total scores had an SP of 80% or higher and a large effect size value.

In terms of psychometrics, the HCR-20 and HCR-20^{V3} are indexes and not scales (Streiner, 2003). A scale for the measurement of, for instance, anxiety is developed to monitor different indicators of anxiety based on an understanding of anxiety as a single underlying dimension. However, the HCR-20 is an index because, rather than assessing a single underlying dimension, the items are used to evaluate different aspects of risk, such as individual and contextual risk factors, and the feasibility and efficiency of future risk management strategies. A very high correlation among the items would have pointed to deficiencies in the index and that, rather than taping a broad set of indicators, the actual items may be too narrow. Our results indicate that this is not the case with the HCR-20 and the HCR-20^{V3}.

Clinical Utility

In our opinion, the V3 manual is significantly improved compared with the current HCR-20 version. Each risk factor is more elaborated on, detailed, and more thoroughly explained. "Indicators" and "coding notes" provide good guides for the rater. The seven steps (gathering information, the presence of risk factors, relevance of the risk factors, risk formulation, "scenario planning," risk management strategies, and summary) provide a very robust platform and guide for violence risk management.

Coding

In our opinion, the change from numbers to *yes*, *partially/possibly*, and *no* as rating decisions makes the V3 more clinically relevant and dynamic. Notably, it mitigates the risk for importing the actuarial approach into the use of the V3. Violence risk assessment rarely generates a clear-cut, definite answer. We also believe that the new structure in

V3 may substantially reduce the risk for empirically unfounded conclusions.

Relevance

The introduction of the new risk assessment category (relevance) in the HCR-20 to emphasize individual risk factors of special importance is another asset of the V3. It informs the development of a tailor-made and recognizable risk management plan to the best for the consumer and his or her social and professional network.

Risk Formulation, Risk Scenario, and Risk Management

Overall, these features in the V3 contribute to a more systematic and detailed violence risk assessment, with reinforced guidance to integrate development and evaluation of risk management strategies. The requirement in the seven-step procedure to introduce and discuss future risk directly in relation to a risk management plan is a significant step forward emphasizing one of the key facets of the SPJ tradition. Although dynamic factors have been found to enhance the quality of violence risk assessment, the most important implication for clinical practice is that these factors inform the risk management of the assessed individuals (Mills, 2005). In the HCR-20^{V3} manual, the concepts of risk formulation and risk scenario are exemplified by referring to a nursing-driven intervention: "Fluttert et al. (2008) described an 'Early Recognition Method' in which persons and nurses act together to identify the person's signature threats or warning signs, and then develop risk management strategies to mitigate risk early in the process, before violence occurs. Notably, this model also draws on theory to formulate or conceptualize such warning signs. This method has shown some potential to reduce subsequent violence" (Douglas et al., 2013, p. 56). In our case illustration, the identification of the most relevant risk factors forms a starting point for developing nursing-driven structured risk management interventions. In our opinion, there is a mutual relationship between the HCR-20^{V3} approach and developing and implementing structured nursing interventions.

Limitations

We have already discussed possible shortcomings in the design of our study, such as convenience sampling, the selection of raters, and SP issues. There is a risk that the association between V2 and V3 may have been inflated because we did not use independent raters. However, the variability of scores and associations we found at item level and for the H, C, and R parts indicates no evidence of systematic rater bias in the results. It is also important to keep in mind that alpha coefficient values tend to increase with increasing numbers of items and numbers of points on the score scale. The use of a short, 3-point scoring scale (0, 1, and 2) and, particularly, the very brief five-item C and R parts is vulnerable to this effect. However, it is important to keep in mind that numbers are replaced by words in the HCR-20^{V3}. Because it is not an actuarial instrument with an automatic end-point conclusion concerning risk, the pivotal part of the SPJ approach is to use information obtained by the HCR-20^{V3} and other sources to develop risk scenarios and risk formulations. Therefore, the ultimate test of internal consistency of the HCR-20^{V3} within the SPJ tradition would be to monitor the association between scenarios and formulations for independent raters. Finally, our research design is limited by the fact that it did not include an investigation of interrater reliability. This is a natural and requisite focus for future research on the HCR-20.

Concluding Remarks

We have conducted small-scale research on the internal consistency and clinical utility of the HCR-20^{V3} compared with the HCR-20. Our overall conclusion is that the new version is a substantial contribution to the improvement of risk assessment within the SPJ tradition. The fact that the scores on the C items yielded lower internal consistency when comparing the two versions may indicate that the most substantial item difference in the new version pertains to clinical items. This may be an asset in the sense that the new version facilitates and demands more accurate clinical evidence of risk. On the other hand, there is always a risk for overestimating risk if a clinician misinterprets the use of indicators provided for each item. Even if no single part of the seven-step approach has its origin in the HCR-20^{V3}, the seven-step model has fortified a structure to secure that tailor-made risk management plans may be developed as part of the risk assessment process. According to our long experience with the HCR-20 and our brief use of the HCR-20^{V3}, the latter appears to be a major step forward in improving risk assessment of violence. In our opinion, the design of the new HCR-20 version invites nurses to play an even more direct and essential role in the assessment and management of violence risk.

References

Bjørkly, S. (1994). The Scale for the Prediction of Aggression and Dangerousness in Psychotic Patients (PAD)—A prospective pilot study. *Criminal Justice and Behavior, 21*, 341–356. doi:10.1177/009385489402100300

Bjørkly, S. (2004). Risk management in transition between forensic institutions and the community: A literature review and an introduction to a milieu treatment approach. *International Journal of Forensic Mental Health, 3*(1), 67–75. doi:10.1080/14999013.2004.10471197

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20^{V3} Assessing Risk for Violence*. Vancouver, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *British Medical Journal*, *345*, e4692. doi:10.1136/bmj.e4692

AQ4 Fluttert, F. A. J., Meijel, B. V., Nijman, H., Bjørkly, S., & Grypdonck, M. (2010). Preventing aggressive incidents and seclusions in forensic care by means of the "Early Recognition Method." *Journal of Clinical Nursing*, *19*, 1529–1537. doi:10.1111/j.1365-2702.2009.02986

Fluttert, F. A. J., Meijel, B. V., Webster, C., Nijman, H., Bartels, A., & Grypdonck, M. (2008). Risk management by early recognition of warning signs in forensic psychiatric patients. *Archives of Psychiatric Nursing*, *22*(4), 208–216. Doi:10.1016/j.apnu.2007.06.012

Mills, J. F. (2005). Advances in the assessment and prediction of interpersonal violence. *Journal of Interpersonal Violence*, *20*, 236–241.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Pouncey, C., & Lukens, J. (2010). Madness versus badness: The ethical tension between the recovery movement and forensic psychiatry. *Theoretical, Medical & Bioethics*, *3*, 3–105.

Rasmussen, K., Jakobsen, D., & Urheim, R. (2002). *HCR-20: Vurdering av voldsrisiko (HCR-20: assessment of violence risk)*. Trondheim, Norway: Psykologisk Institutt, NTNU.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and meta regression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*, 499–513. doi:10.1016/j.cpr.2010.11.009

Streiner, D. L. (2003). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment*, *80*(3), 217–222.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20 assessing risk for violence: Version 2*. Vancouver, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme. The assessment of dangerousness and risk*. Vancouver, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

AQ1 = Please provide academic degrees for the last two authors.

AQ2 = Please check whether footnotes were appropriately captured and structured in all tables.

AQ3 = Please provide citation of Table 4 in the text.

AQ4 = Please provide citation of this reference (Fluttert, Meijel, Nijman, Bjorkly, & Grypdonck, 2010) in the text.

END OF AUTHOR QUERIES