Received: 19 January 2024 | Revised: 25 April 2024 | Accepted: 8 May 2024

DOI: 10.1002/ab.22150

BRIEF REPORT

AGGRESSIVE BEHAVIOR WILEY

Risk and protective factors in risk assessment: Predicting inpatient aggression in adult males detained in a forensic mental health setting

Jane L. Ireland[1] | Yara Levtova[1] | Christa-Maria Abi Semaan[2] | Lisa M. B. Steene[1] | Sören Henrich[1] | Lisa Gaylor[1] | Lina Driemel[2] | Sophie Volz[2] | Julia Homann[2] | Mareike Dickopf[2] | Leah Greenwood[1] | Simon Chu[1]

[1]School of Psychology and Humanities, University of Central Lancashire, Preston, UK

[2]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands

Correspondence
Jane L. Ireland, Darwin Building, University of Central Lancashire, Preston, PR1 2NE, UK.
Email: JLIreland1@uclan.ac.uk

Abstract

Structured clinical risk assessments represent a preferred means of assessing levels of aggression risk at different times and in different individuals. Increasing attention has been given to capturing protective factors, with sound risk assessment critical to high-secure forensic mental health care. The aim was to assess the predictive value of the HCR-20^{V3} for aggression risk and the long-term care pilot version of the SAPROF (the SAPROF-LC-pilot) in a high-secure forensic mental health inpatient population and to determine the incremental value of protective over risk factors. Participants were adult males detained in a high secure forensic mental health service, with a primary diagnosis of schizophrenia and/or personality disorder. The focus was on examining hospital based aggression (self- and other-directed) at two time points; up to 6 months (T1) and between 7 and 12 months (T2). The HCR-20^{V3} and SAPROF-LC-pilot demonstrated good predictive validity but with variability across subscales and aggression types/periods. Historical factors of the HCR-20^{V3} and External factors of the SAPROF-LC-pilot failed to predict, aside from a medium effect at T1 for verbal aggression and self-harm, for Historical factors. There was evidence for protective factors adding to prediction over risk factors alone, with the integration of protective and risk factors into a risk judgement particularly helpful in improving prediction accuracy. Protective factors contributed to risk estimates and particularly if integrated with risk factors. Combining risk and protective factors has clear predictive advantages, ensuring that protective factors are not supplementary but important to the aggression assessment process.

KEYWORDS

forensic patients, HCR-20, protective risk assessment, SAPROF, violence risk assessment

Aggression in inpatient forensic mental health settings is recognised as a common occurrence (Persson et al., 2017) that can extend to severe acts of aggression (Tuente et al., 2021). The accurate prediction and prevention of future institutional aggression is important in terms of effective management but remains a complex task (Abbiati et al., 2020; de Vogel et al., 2022). The implementation of empirically based, structured and clinically effective risk assessment tools are considered crucial (Fazel et al., 2022) to the safe management of aggression, with reliance to date on structured risk assessments for aggression, such as the HCR-20^{V3} (Historical Clinical Risk Management guide; Douglas et al., 2013). The HCR-20^{V3} has been the most widely utilised structured assessment for assessing risk, with this applied to those with a history of physical aggression (Douglas et al., 2014; Neil et al., 2020). The HCR-20^{V3} has demonstrated good predictive validity for future aggression, yielding moderate to large effect sizes (Brookstein et al., 2020; Persson et al., 2017).

However, contemporary views on risk management point to a need to capture risk factors alongside those that may ameliorate risk (Cappon et al., 2023; de Vogel et al., 2012; Robbé et al., 2012). Such approaches identify value in capturing protective factors (Ttofi et al., 2016). It is accepted that there are complicated conceptual and methodological issues with regard to protective factor research in comparison to risk factor research, which has led to differences in how the term "protective factor" has been operationalised (Ttofi et al., 2016). For example, some researchers define such factors as those predicting a low probability of offending (Loeber et al., 2008) whereas others consider them factors that interact with risk factor(s) to reduce or nullify risk effects (Rutter, 1987). This has led to some researchers describing these either as an "interactive protective factor" (or 'buffering protective factor') or a "risk based protective factor" (Ttofi et al., 2016). Regardless as to the specific definition, there is a consensus on the value of protective factors and the importance of capturing these in assessments of aggression.

There has certainly been a pivot in the last 15 years to specifically include protective factors in assessments of risk, with the Structured Assessment of Protective Factors (SAPROF: de Vogel et al., 2012) the first guide developed specifically for aggression that focuses solely on protective factors. The SAPROF adopts a strength-based approach to identifying factors that brings someone away from aggression. Research has demonstrated good predictive validity for the SAPROF in relation to violent[1] recidivism (Robbé et al., 2012), institutional misconduct and self-harm, after short- and long-term follow-ups (Abidin et al., 2013), including for inpatient aggression (Burghart et al., 2023: reporting moderate to good predictive validity). Whilst acknowledging the predictive validity of the SAPROF and increased interest in this area, there has been increased focus on specifically assessing predictive and incremental validity (Burghart et al., 2023). This has noted the value of including protective factor tools when used alongside a risk focused tool with regard to predicting risk, whilst also noting shortcomings in the research base in terms of failing to report calibration indices (Burghart et al., 2023).

There are also differences noted in relation to how the subscales of these assessment guides predict risk, particularly regarding the value of dynamic factors over static factors (Burghart et al., 2023; Neil et al., 2020), although research is neglected in higher secure populations. Furthermore, although including risk and protective factors are presented as a more holistic assessment of risk (Burghart et al., 2023; Robbé et al., 2012), again, forensic patients detained in conditions of raised security have been under-researched. This is surprising considering this population is considered at elevated risk for aggression. To address this gap in the literature, we aimed to assess the predictive validity of the HCR-20^{V3} and examine the predictive and incremental validity of the SAPROF-LC-pilot in the prediction of inpatient aggression in a high secure forensic mental health hospital. We predicted that the HCR-20^{V3} and Final Risk Judgement would demonstrate good predictive validity for predicting future inpatient aggression (Green et al., 2016; Smith et al., 2020), that the HCR-20^{V3} subscales would perform differently in terms of predictive value (Hogan & Olver, 2016; Neil et al., 2020), and that the SAPROF-LC-pilot Final Protection Judgement would demonstrate good predictive validity for predicting the absence of future inpatient aggression (de Vries Robbé et al., 2017), with differences observed between SAPROF-LC-pilot subscales.

1 | METHOD

One hundred and six adult male participants, detained in a high secure forensic mental health service in England, housing patients with severe mental illness and/or personality disorder, volunteered to take part after being approached. Fifteen were subsequently excluded due to incomplete data, leaving a total of 91 participants at baseline. The average age at baseline was 38.45 years (s.d. = 10.23; range 22–63 years). All were detained due to aggression. One-third had committed a lethal index offence (31.9%), with the remainder involved in a non-lethal index offence (Supporting Information S1: Table S1 includes participant demographics). The hospital predominately houses a white adult male population.

For each participant, a HCR-20^{V3} (Douglas et al., 2013) and SAPROF-LC-pilot (de Vries Robbé et al., 2017) was rated retrospectively from clinical records.[2] The HCR-20^{V3} comprises 20 risk items (each rated as either present, possibly present, or not evidenced) divided into 10 historical, five dynamic, and five that are projective for future risk management. This allowed for subscale scores and a total HCR-20^{V3} score to be calculated. The original SAPROF comprises 17 items (each rated as clearly present, present to some extent, or not present) divided into five internal (static/dynamic), seven motivational (dynamic), and five external (dynamic) items. In this study, suggested alterations to two of the original factors into subfactors and one additional Long-care specific factor were included, altogether making

TABLE 1 Descriptive statistics for the HCR-20^{V3} and SAPROF-LC-pilot.

Risk and protective factors	Mean (SD)	Median	Range
HCR-20^{V3}			
Historical scale	17.73 (2.16)	18	10–20
Clinical scale	5.14 (2.60)	5	0–10
Risk management scale	4.54 (2.27)	5	0–13
HCR-20^{V3} total	27.41 (4.93)	28	18–38
SAPROF-LC-pilot			
Internal scale	4.80 (2.72)	5	0–11
Motivational scale	8.43 (2.54)	9	1–11
External scale	9.54 (2.62)	9	5–14
SAPROF-LC-pilot total	22.77 (6.46)	23	7–36

21 factors: six internal, eight motivational, and seven external (see Table 1). This allowed for subscale scores and a total SAPROF-LC-pilot score to be calculated. The SAPROF-LC-pilot further instructs assessors to make an Integrative Final Risk Judgement. This overall risk estimate is a judgement made by the evaluator based on the SAPROF-LC-pilot (Final Protection Judgement) and the HCR-20^{V3} (Final Risk Judgement). These ratings were captured as "low," "moderate," or "high" in terms of aggression risk. Each assessment was double coded for interrater reliability analyses. Comprehensive incident reports of aggression were collated from routinely gathered hospital information for T1 (up to 6 months) and T2 (7–12 months) periods following the point of assessment. Raters were blind to the outcome data at the time of assessment. Incident data was captured using the Overt Aggression Scale (Yudofsky et al., 1986). This comprises four categories; verbal aggression, physical aggression against objects, physical aggression against self, and physical aggression against others. All statistical analyses were conducted using SPSS Statistics v24 and JASP 0.14.3

1.1 | Ethical approval

Procedures were approved by South Central - Berkshire NHS Research Ethics Committee—approval: 20/SC/0168. All adult participants provided written informed consent to participate in this study.

2 | APPROACH TO ANALYSIS

To establish interrater reliability, all cases in the study were double-coded. For numerical items such as total scores and subscales, the Intraclass Correlation Coefficient (ICC_2) was utilised, while weighted kappa was employed for ordinal variables (i.e., overall risk judgements). To establish convergent validity, correlation analyses between assessments of protective factors and risk factors were conducted, with the alpha level adjusted using the Bonferroni correction to account for multiple comparisons. The validity of the prediction model was assessed by analysing predictive discrimination and calibration. Although the SAPROF's discriminatory ability has been extensively studied, its calibration still requires further evaluation (see Burghart et al., 2023). Discrimination, or the model's capacity to differentiate between individuals who desist from aggression and those who do not, was evaluated using Receiver Operating Characteristics (ROC) curve analysis. Area Under the Curve (AUC) values were interpreted based on established thresholds: 0.64 for a medium effect and 0.71 for a large effect (Rice & Harris, 2005).

For the calibration assessment the Brier score was used, ranging from 0 to 1, with lower scores indicating better model calibration. The model's Brier score was compared with a score of zero and the mean to evaluate calibration, where a score lower than these comparative values indicates good calibration (Ferro, 2007, 2014). Finally, to assess the incremental predictive validity of protective factors beyond the risk assessment, hierarchical logistic regressions were applied for each Overt Aggression Scale outcome at T1 and T2 (T1 equals up to 6 months and T2 between 7 and 12 months post risk assessment completion[3]). This analysis was bootstrapped with 5000 samples and Bias-Corrected bootstrap confidence intervals were calculated to ensure the robustness and stability of the predictive model.

3 | RESULTS

At T1 25.3% of patients had recorded incidents of physical aggression, 34.1% had incidents of verbal aggression, and 16.5% had incidents of self-harm. At T2 the corresponding proportions were 26.5% physical aggression, 39.8% verbal aggression and 9.2% self-harm. At both time points, 44% had engaged in "any aggression."

Interrater reliability of the SAPROF-LC-pilot total scores were excellent (ICC_2 = 0.91) and the weighted Kappa for the Final Protection Judgement was 0.66, indicating good agreement. The internal, motivational, and external subscales had reliabilities ranging from good to excellent (ICC_2 = 0.84, 0.73, and 0.88 respectively), with the HCR-20^{V3} subscales (H, C, and R) and total score presenting with excellent reliability (ICC_2 = 0.91, 0.92, 0.81, and 0.93 respectively). The Final Risk Judgement had a weighted Kappa of 0.64, indicating a good level of agreement. The Integrative Final Risk Judgement had excellent reliability (Fleiss, 1986), with a weighted Kappa of 0.94. Means across the HCR-20^{V3} and SAPROF-LC-pilot are indicated in Table 1. Regarding the Final Protection Judgement, 25.3% were scored as having a low level of protective factors, 53.8% a moderate level, and 20.9% a high level. In terms of Final Risk Judgement, 29.7%

were rated as a low level of risk, 37.4% moderate, and 33% high. Finally, regarding the Integrative Final Risk Judgement, 42.9% were classified as low risk, 36.3% as moderate and 20.9% as high risk.

3.1 | Convergent validity

The HCR-20^{V3} and SAPROF-LC-pilot totals demonstrated a strong and negative association (r = −0.67, p < 0.000), indicating that as risk factors increased, protective factors decreased. The SAPROF-LC-pilot total score was negatively associated with any type of aggression at T1 (r = −0.41, p < 0.000) and T2 (r = −0.43, p < 0.000), with the HCR-20^{V3} total score having a positive and significant relationship with any type of aggression and T1 (r = 0.47, p < 0.000) and T2 (r = 0.45, p < 0.000). Table 2 presents the correlation between subscales.

3.2 | Predictive accuracy

ROC analyses examining the discrimination properties of the HCR-20^{V3} and SAPROF-LC-pilot for the Overt Aggression Scale outcomes are presented in Table 3. In relation to the HCR-20, the Clinical subscale, Risk Management subscale and HCR-20^{V3} total predicted "any aggression" and "physical aggression to objects," to a large effect at T1, with Risk Management and Total HCR-20^{V3} similarly predicting physical aggression to others, verbal aggression and (for total HCR-20^{V3}) self-harm at T1. These findings were broadly replicated at T2, although the Clinical subscale was predicting more consistently. The Final Risk Judgement performed consistently well for all outcomes at both time points. The SAPROF-LC-pilot total and subscales also predicted risk in the expected direction, although there was more variability; the External subscale was consistently not predicting, the Internal subscale was consistently predicting (although less so for physical aggression towards others at T2), with the Motivational subscale, Total SAPROF-LC-pilot and Final Protective Judgement appearing to predict any aggression and longer term predictions in an improved manner. Consistently, however, the Integrative Final Risk Judgement was predicting all forms of aggression at both time points, with a consistently large effect size.

3.3 | Calibration

To assess the estimation of an exact probability of an aggressive/violent incident, the Brier score was calculated using the total protection and risk scores as well as the Integrative Final Risk Judgement. For T1 and any aggressive incidents, the Brier score (*Br*: 0.1476) was lower than that using mean predicted probability (*Br*: 0.4396) or using a predicted probability of zero (*Br*: 0.2725). The Brier score for T2 was slightly higher for any aggression (*Br*:

TABLE 2 Correlations between HCR-20^{V3}, SAPROF-LC-pilot and structured professional judgement estimates.

	Clinical	Risk management	HCR-20^{V3} total	FRJ	Internal	Motivational	External	SAPROF total	FPJ	IFRJ
Historical	−0.06 (0.545)	0.11 (0.320)	0.45 (0.000)	0.08 (0.434)	−0.10 (0.327)	0.06 (0.546)	0.00 (0.993)	−0.02 (0.865)	0.04 (0.728)	0.11 (0.322)
Clinical	0.63*(0.000)		0.79(0.000)	0.65*(0.000)	−0.72(0.000)	−0.60*(0.000)	−0.36*(0.000)	−0.69*(0.000)	−0.64*(0.000)	−0.64(0.000)
Risk management			0.84*(0.000)	0.65(0.000)	−0.57*(0.000)	−0.54(0.000)	−0.49*(0.000)	−0.65*(0.000)	−0.59*(0.000)	0.58*(0.000)
HCR-20^{V3} total				0.68*(0.000)	−0.68*(0.000)	−0.54*(0.000)	−0.42*(0.000)	−0.67(0.000)	−0.59*(0.000)	−0.65*(0.000)
FRJ					−0.64(0.000)	−0.40*(0.000)	−0.24 (0.023)	−0.52*(0.000)	−0.55*(0.000)	0.81*(0.000)
Internal						0.59*(0.000)	0.40*(0.000)	0.82*(0.000)	−0.71*(0.000)	−0.68*(0.000)
Motivational							0.54(0.000)	0.86*(0.000)	0.75*(0.000)	−0.53*(0.000)
External								0.78(0.000)	0.56*(0.000)	−0.30 (0.002)
SAPROF-LC-pilot total									0.82* (0.000)	−0.62*(0.000)
FPJ										−0.66*(0.000)

Note: N = 91.

Abbreviations: FPJ, Final Protection Judgement; FRJ, Final Risk Judgement; HCR-20^{V3}, Historical Clinical Risk management-20 Version 3; IFRJ, Integrative Final Risk Judgement; SAPROF-LC-pilot, Structured Assessment of Protective Factors for violence risk Long-term Care pilot version.

*r_s is significant at the Bonferroni corrected p-value of 0.0008 (0.05/66).

TABLE 3 Predictive validity of the HCR-20^{V3}, SAPROF-LC-pilot, and final estimates for inpatient aggression.

	Six months follow-up										Seven to 12-months follow up									
	Any aggression		Physical aggression others		Physical aggression objects		Verbal aggression		Self-harm		Any Aggression		Physical aggression others		Physical aggression objects		Verbal aggression		Self-harm	
	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI
HCR-20^{V3}																				
Historical subscale	0.60	0.49–0.72	0.54	0.37–0.71	0.61	0.46–0.75	0.63*	0.50–0.75	0.69*	0.53–0.84	0.61	0.50–0.73	0.61	0.46–0.76	0.55	0.39–0.71	0.57	0.45–0.69	0.56	0.38–0.74
Clinical subscale	0.71***	0.61–0.82	0.64	0.50–0.78	0.73***	0.61–0.86	0.65*	0.53–0.78	0.64*	0.51–0.78	0.76***	0.65–0.86	0.77***	0.63–0.90	0.81***	0.71–0.91	0.73***	0.62–0.84	0.71**	0.58–0.85
Risk management subscale	0.75***	0.65–0.85	0.74***	0.62–0.86	0.73***	0.63–0.84	0.70**	0.59–0.81	0.64*	0.51–0.77	0.73***	0.63–0.83	0.76***	0.64–0.88	0.75***	0.64–0.87	0.70***	0.60–0.81	0.70**	0.57–0.83
HCR-20^{V3} total	0.76***	0.67– 0.86	0.71**	0.58–0.84	0.77***	0.67–0.87	0.69**	0.58–0.81	0.70**	0.58–0.83	0.76***	0.67–0.86	0.78***	0.65–0.92	0.77***	0.65–0.90	0.72***	0.61–0.83	0.73**	0.57–0.88
Final risk judgement	0.83***	0.75–0.92	0.79***	0.68–0.90	0.85***	0.77–0.94	0.77***	0.67–0.87	0.79***	0.68–0.90	0.84***	0.75–0.92	0.75***	0.63–0.86	0.89***	0.82–0.96	0.82***	0.73–0.91	0.83***	0.72–0.94
SAPROF-LC-pilot																				
Internal subscale	0.81***	0.71–0.90	0.75***	0.63–0.87	0.78***	0.66–0.90	0.75***	0.64–0.86	0.76***	0.65–0.88	0.77***	0.67–0.87	0.68*	0.54–0.83	0.82***	0.70–0.94	0.75***	0.64–0.85	0.79***	0.64–0.95
Motivational subscale	0.72***	0.61–0.82	0.60	0.45–0.75	0.66*	0.52–0.79	0.65*	0.53–0.77	0.58	0.44–0.73	0.73***	0.62–0.83	0.76*	0.51–0.80	0.76***	0.65–0.88	0.70***	0.60–0.81	0.71*	0.54–0.87
External subscale	0.55	0.43–0.67	0.47	0.32–0.62	0.55	0.42–0.69	0.52	0.40–0.64	0.50	0.35–0.65	0.61	0.49–0.72	0.51	0.33–0.769	0.63	0.48–0.79	0.59	0.47–0.71	0.62	0.45–0.80
SAPROF-LC-pilot total	0.73***	0.63–0.83	0.62	0.47–0.76	0.69**	0.56–0.82	0.67**	0.55–0.78	0.64	0.50–0.78	0.74***	0.64–0.84	0.62	0.45–0.78	0.77***	0.63–0.90	0.71***	0.61–0.82	0.74**	0.58–0.89
Final protection judgement	0.71***	0.60–0.82	0.70**	0.56–0.83	0.71**	0.59–0.83	0.66**	0.55–0.78	0.66*	0.53–0.80	0.71***	0.61–0.82	0.67*	0.54–0.81	0.76***	0.64–0.88	0.69**	0.58–0.80	0.72**	0.56–0.88
Integrative final risk judgement	0.83***	0.74–0.92	0.81***	0.72– 90	0.86***	0.79–0.94	0.78***	0.67–0.88	0.79***	0.66–0.92	0.86***	0.77–0.94	0.72***	0.60–0.84	0.89***	0.82–0.96	0.77***	0.66–0.87	0.88***	0.79–0.97

*P < .05; **P < .01; ***P < .001

0.1740), compared to T1, indicating a reduced predictive accuracy at the second follow-up. However, this was still lower than the mean predicted probability (*Br*: 0.4176) or the predicted probability of zero (*Br*: 0.2995).

3.4 | Incremental predictive validity of HCR-20^{V3} and SAPROF-LC-pilot

The following analyses considered if protective factors (SAPROF), including the Integrative Final Risk Judgement, improve prediction of aggression outcomes beyond the contribution of risk factors (HCR-20^{V3}) alone. Hierarchical Logistic Regressions were employed, with the aggression outcomes representing any aggression, verbal aggression, aggression towards objects and aggression towards others.

3.4.1 | Any aggression at T1 and T2

The initial model (Block 1) with only the HCR-20^{V3} total score significantly predicted any aggression at T1 (Wald = 16.38, $p < 0.001$) and classified 69.1% of cases correctly. The model remained significant with the addition of SAPROF-LC-pilot total score in Block 2. However, neither the improvement to the model (change $\chi^2 = 2.18$, $p = 0.144$) nor the predictor was significant, and resulted in a slight decrease in classification accuracy to 68.1%; therefore, adding protective factors did not improve the prediction of any aggression at T1, beyond the risk factors. Notably, the addition of the Integrative Final Risk Judgement in Block 3 resulted in the model with the best fit ($\chi^2 = 41.10$, $p < 0.001$, $R^2 = 0.36–0.49$); this was a significant improvement from the risk and protective factors alone (change $\chi^2 = 16.96$, $p < 0.001$). This final model correctly classified 78% of cases and accounted for up to 49% of the variance in any aggression at T1 ($R^2 = 0.36–0.49$), although it is important to note that although both Integrative Final Risk Judgements (low to moderate; moderate to high) were significant predictors, the associated standard error for moderate to high was large (SE = 9.16), suggesting caution in interpreting this for moderate to high ratings.

When exploring the prediction of any aggression at T2 results differed to a degree. The addition of protective factors improved the change and the model, and they were a significant predictor when added to risk factors, which also remained significant (Block 2). However, when the Integrative Final Risk Judgements were added (Block 3) the difference between moderate and low risk (as determined by the Integrative Final Risk Judgement) was the sole significant predictor of any aggression at T2, with the difference between moderate and high not. As before, the final model provided the best fit (−2 Log Likelihood = 89.16, $\chi^2 = 34.51$, $p < 0.001$, $R^2 = 0.32–0.43$) and correctly classified 75.8% of cases.

3.4.2 | Verbal aggression at T1 and T2

For verbal aggression at T1 the model with only the HCR-20^{V3} significantly predicted aggression (Wald = 9.59, $p < 0.001$). The addition of SAPROF-LC-pilot in Block 2 did not significantly improve the model ($\chi^2 = 0.43$, $p = 0.557$). Again, the addition of the Integrative Final Risk Judgement resulted in the model with the best fit (−2 Log Likelihood = 93.07, $\chi^2 = 23.68$, $p < 0.001$, $R^2 = 0.23–0.32$) classifying 75.8% of the cases correctly. The difference between moderate and low risk (Integrative Final Risk Judgement) was the sole predictor ($b = −1.88$, SE = 1.83, Wald = 7.01, $p = 0.006$).

Similar results were obtained for verbal aggression at T2; the model with HCR-20^{V3} only significantly predicted aggression (Wald = 11.35, $p < 0.001$). The addition of SAPROF-LC-pilot in Block 2 did not significantly improve the model. Again, the difference between moderate and low risk (Integrative Final Risk Judgement) was the sole significant predictor in the final model (B = −1.66, $p = 0.011$). The final model classified 68.1% of the cases correctly.

3.4.3 | Aggression against objects at T1 and T2

For aggression against objects at T1 risk factors alone significantly predicted aggression (Wald = 11.13, $p < 0.001$). The addition of SAPROF-LC-pilot did not significantly improve the model ($\chi^2[1] = 0.63$, $p = 0.469$). Both the change from moderate to low risk and moderate to high risk (Integrative Final Risk Judgement) were significant predictors in the final model and improved the prediction of aggression against objects at T1 (change $\chi^2 = 19.73$, $p < 0.001$) and resulted in the model with the best fit (−2 Log Likelihood = 58.48, $\chi^2 = 34.77$, $p < 0.001$, $R^2 = 0.32– 0.50$), classifying 80.2% of the cases correctly. At T2 the model with only risk factors again significantly predicted aggression (Wald = 10.64, $p < 0.001$). However, the addition of SAPROF-LC-pilot in Block 2 significantly improved the model ($\chi^2 = 4.83$, $p = 0.031$), with protective factors the sole significant predictor in the model. When the Integrative Final Risk Judgement was added (Block 3), both the change from moderate to low risk and moderate to high risk were significant predictors. A large confidence interval for the change from moderate to high (−1.29, 21.91, 95% CI) again highlighted that this result needs to be interpreted with caution. The final model resulted in the model with the best fit (−2 Log Likelihood = 51.25, $\chi^2 = 36.40$, $p < 0.001$, $R^2 = 0.33–0.53$), with 86.8% of cases correctly classified.

3.4.4 | Aggression against others at T1 and T2

At T1 the model with only the HCR-20^{V3} significantly predicted aggression (Wald = 6.58, $p = 0.001$). The addition of SAPROF-LC-pilot in Block 2 did not significantly improve the model ($\chi^2 = 0.00$, $p = 0.960$). The Integrative Final Risk Judgement demonstrated significant associations, specifically for predicting low versus

moderate risk. This produced the best fitting model, with 83.5% of cases correctly classified. At T2 the model with the HCR-20^{V3} only, significantly predicted aggression (Wald = 9.86, p = 0.004). The addition of SAPROF-LC-pilot in Block 2 did not improve the model. In the last block, change from moderate to low risk (Integrative Final Risk Judgement) was the sole predictor but showed a very broad confidence interval (ranging from −20.20 to −0.76) and was associated with a high standard error (St. E = 8.72), indicating uncertainty.

4 | DISCUSSION

The HCR-20^{V3} and the SAPROF-LC-pilot demonstrated convergent validity; as risk factors increased (HCR-20^{V3}) protective factors decreased (SAPROF-LC-pilot). In addition, protective factors negatively associated with aggression, with the opposite relationship found for risk factors. When attending to overall predictive accuracy and incremental predictive validity a more mixed picture emerged. Findings demonstrated that the historical factors of the HCR-20^{V3} were not presenting as notable predictors of risk across aggression types or time points, a finding consistent with Neil et al. (2020) and also echoing the findings of Burghart et al. (2023) in relation to the value of dynamic factors. However, remaining HCR-20^{V3} subscales and the total were predicting, but this held particularly for the Risk Management subscale and HCR-20^{V3} total, with the Clinical (dynamic) items becoming increasingly of value only as the time period extended. These findings supported the prediction that there would be a difference in the predictive value of the HCR-20^{V3} subscales (Hogan & Olver, 2016; Neil et al., 2020), demonstrating that the HCR-20^{V3} dynamic subscales (i.e., Clinical, Risk Management) were better at predicting future inpatient aggression than static (Historical) factors. Indeed, the current study highlighted limited variability in the Historical scale risk factors in this sample. This is comparable to previous research (Green et al., 2016; Hogan & Olver, 2016; Neil et al., 2020) and perhaps provides a basis for the lack of predictive prowess for historical factors since they were uniform among this population. Rather, it appears the dynamic elements of the HCR-20^{V3} were particularly well suited to the assessment of inpatient aggression with this population. These results are promising as the dynamic scales, by definition, are most amenable to change in risk level and form the target of management strategies and intervention (Douglas et al., 2013; Fazel et al., 2022).

Notable differences were also found with the SAPROF-LC-pilot subscales, which supported the prediction. Here we noted that the Internal subscale was the strongest and most consistent predictor of the absence of (any) inpatient aggression. The Motivational subscale also robustly predicted all types of aggression at the second time point but its predictive utility was limited to verbal and any aggression at time point one. The External subscale was not proving valuable. Importance was, nevertheless, demonstrated in the SAPROF-LC-pilot in its entirety; having more protective factors and a higher rating on the Final Protection Judgement reduced the risk of aggression. The results support prior (limited) research (de Vries Robbé et al., 2017), highlighting the importance of capturing protective factors in detained forensic patients. Indeed, the SAPROF-LC-pilot Internal subscale presented as the best predictor of the absence of any future inpatient aggression, at both time points. This subscale captures personal characteristics and capabilities and is particularly consistent with strength-based models of forensic assessment that place emphasis on developing individual skills and abilities as opposed to a single focus on capturing risk factors (Robbé et al., 2012; Ttofi et al., 2016). Paying particular attention to these may prove especially salient for long-stay settings, where motivation for change and hope for the future becomes increasingly depleted among patients.

The findings also pointed to the importance of not making general judgements of "protective" and "risk" factors uniformly predicting aggression. Rather, what is supported is a more diverse picture that needs to account for aggression type, time period under review and how risk and protective factors are considered, which arguably reflects the diversity in how these factors have been defined and applied (e.g. Rutter, 1987; Ttofi et al., 2016). Connected to the latter, applying a judgement that combined protective and risk factors (i.e., the Integrative Final Risk Judgement) to predict aggression was clearly demonstrating advantages in relation to prediction. For example, when predicting "any aggression" there were evidenced advantages in applying a judgement that combined protective and risk factors to predict at both time points. The incremental predictive validity findings were particularly illuminating in this regard, demonstrating how protective factors were adding to improvement in prediction beyond risk factors alone. This appeared more uniformly the case when combined with risk factors into a single judgement of risk (i.e., the Integrative Final Risk Judgement). This improved model fit and appeared consistent most with judgements of low to moderate risk.

Ultimately, the findings in their entirety present some additional foundation in identifying the value of protective factors for high secure forensic mental health populations and how these are not merely supplementary to standard aggression risk assessments (e.g., HCR-20^{V3}) but are actually improving predictions for within hospital aggression. This may seem obvious but this finding is not well considered in practice, with a focus instead on conducting risk as opposed to risk and protective assessments. The current research indicates that assessing for risk factors alone and neglecting protective factors is not assisting risk prediction. Such an approach simply ensures that the accuracy of our predictions of aggression fail to improve.

Of further note is clear value in asking for an overall clinical judgement on risk that is *directly informed by* the completion of risk and protective structured clinical assessments. Making risk statements using *unstructured* clinical judgement (i.e., those made without the benefit of a structured clinical guide) is contentious in the field and not generally preferred (Quinsey et al, 2006). However, what the current research suggests is value in the concept of an "*informed* Integrative Final Risk Judgement," namely that following completion

of structured risk and protective assessments. There was clear predictive value in the overall Integrative Final Risk Judgement at both time points. This was adding positively to incremental predictive value for aggression. The findings promote value in allowing for a degree of professional discretion when arriving at an Integrative Final Risk Judgement. Indeed, even the Final Risk Judgement (based solely on the HCR-20^{V3} completion) and the Final Protection Judgement (based solely on the SAPROF-LC-pilot) were consistently predicting risk, although effect sizes were consistently larger for the Final Risk Judgement. This suggests that completing the HCR-20^{V3}/SAPROF-LC-pilot and then making a judgement of overall risk holds value.

Interestingly, how such decisions were reached by a rater and, specifically, how they integrated risk and protective judgements remains unknown. It is accepted that there is diversity in how the association between risk and protective factors are conceptualised; we have no means of ascertaining if the raters were considering protective factors as interacting factors serving to reduce/nullify risk and/or as factors that predicted low risk probability (Loeber et al., 2008; Rutter, 1987; Ttofi et al., 2016), or if this altered across individual patient ratings. However, what the findings demonstrate is value in an integrative final judgement. What is not yet revealed is the exact mechanism by which this is achieved. This falls beyond the scope of the current study but represents a valuable direction for future research. Connected to this, there could be clear value in adopting methodological designs for exploring the different conceptualisations of protective factors (i.e., interactive protective factors vs. those for risk-based protective factors: Ttofi et al., 2016), an approach future research could consider.

We certainly recognise limitations to the current study, which should temper absolute conclusions. For example, the HCR-20^{V3} and SAPROF-LC-pilot were rated retrospectively from file information, a criticism that has been levelled at previous research (Burghart et al., 2023). Although the acquired information was comprehensive, two SAPROF-LC-pilot items had to be omitted due to lack of information. This was not considered a significant limitation as all remaining SAPROF-LC-pilot items could be rated. In addition, caution should be applied when attempting to generalise results beyond a secure forensic population, and to women. The latter remains an under-researched area in particular (Burghart et al., 2023). Finally, replication of the current findings will be important in clarifying the predictive validity of the risk and protective assessment approach, with the current study intending to spark research interest in this important but neglected area.

CONFLICT OF INTEREST STATEMENT

Jane L. Ireland and Simon Chu are employed by the Trust where the data were collected. Jane L. Ireland provides training on risk assessment guides.

DATA AVAILABILITY STATEMENT

The participants of this study did not give written consent for their data to be shared publicly. Consequently, supporting data is not available.

ORCID

Jane L. Ireland http://orcid.org/0000-0002-5117-5930
Sören Henrich http://orcid.org/0000-0002-4792-8725

ENDNOTES

[1] The terms violence and physical aggression have been used interchangeably in the risk literature, although violence has been used to refer specifically to physical aggression. The term aggression, however, has been used to refer to the wider range of direct aggressive acts that can take place (e.g., verbal) but has also been used by some researchers to capture physical aggression. Consequently, for clarity, the term aggression is adopted in the current research to include all forms of direct aggressive acts.

[2] The mean number of items omitted due to insufficient information across the HCR-20^{V3} and the SAPROF-LC-pilot was 2.27 (SD = 0.68) per patient. The most frequently omitted items were "intelligence" (n = 82; 90.1%) and "work" (n = 90; 98.9%), both part of the SAPROF-LC-pilot (internal and motivational items respectively). No issues were noted regarding the HCR-20^{V3}. In addition, the final SAPROF-LC now includes six more additional factors that were derived from insights working with the LC in long term care practice. These factors overlap markedly with the suggested additional factors for ID (intellectually disable populations: Cappon et al., 2023).

[3] Risk assessments were completed to cover the period May 1, 2021 to November 1, 2021, with the incidents then rated prospectively.

[4] Op. cit. endnote 2.

REFERENCES

Abbiati, M., Golay, P., Gasser, J., & Moulin, V. (2020). Protective factor assessments: What are we measuring?—An investigation of the internal validity of the structured assessment of protective factors for violence risk. *Criminal Justice and Behavior*, *47*(4), 383–398. https://doi.org/10.1177/0093854819901157

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry*, *13*(1):197. https://doi.org/10.1186/1471-244X-13-197

Brookstein, D. M., Daffern, M., Ogloff, J. R. P., 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*, *28*(3), 325–342. https://doi.org/10.1080/13218719.2020.1775152

Burghart, M., de Ruiter, C., Hynes, S., Krishnan, N., Levtova, Y., & Uyar, A. (2023). The structured assessment of protective factors for violence risk (SAPROF): A meta-analysis of its predictive and incremental validity. *Psychological Assessment*, *35*(1), 56–67. https://doi.org/10.1037/pas0001184

Cappon, L., Jentsch, A., Roggeman, S., & de Vries Robbé, M. (2023). Protective factors in forensic practice: The added value of the SAPROF-extended version pilot in relation to aggressive incidents. *Criminal Justice and Behavior*, *50*(11), 1603–1622. https://doi.org/10.1177/00938548231196572

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20v3*: Assessing Risk for Violence: User guide. Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., Hart, S. D., Webster, C. D., Belfrage, H., Guy, L. S., & Wilson, C. M. (2014). Historical-clinical-risk management-20, version 3 (HCR-20v3): Development and overview. *International Journal of*

Forensic Mental Health, *13*(2), 93–108. https://doi.org/10.1080/14999013.2014.906519

Fazel, S., Burghart, M., Fanshawe, T., Gil, S. D., Monahan, J., & Yu, R. (2022). The predictive performance of criminal risk assessment tools used at sentencing: Systematic review of validation studies. *Journal of Criminal Justice*, *81*, 101902 https://doi.org/10.1016/j.jcrimjus.2022.101902

Ferro, C. A. T. (2007). Comparing probabilistic forecasting systems with the Brier score. *Weather and Forecasting*, *v22*, 1076–1088.

Ferro, C. A. T. (2014). Fair scores for ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*. *140*(683), 1917–1923 https://doi.org/10.1002/qj.2270

Fleiss, J. L. (1986). *The Design and Analysis of Clinical Experiments*. Wiley.

Green, D., Schneider, M., Griswold, H., Belfi, B., Herrera, M., & De Blasi, A. (2016). A comparison of the HCR-20V3 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

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

Loeber, R., Farrington, D. P., Stouthamer-Loeber, M., & White, H. R. (2008). *Violence and serious theft: Development and prediction from childhood to adulthood*. Routledge. https://doi.org/10.4324/9780203933237

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

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-20V3, and SAPROF. *International Journal of Forensic Mental Health*, *16*(2), 117–129. https://doi.org/10.1080/14999013.2016.1266420

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). American Psychological Association. https://doi.org/10.1037/11367-000

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615–620. https://doi.org/10.1007/s10979-005-6832-7

Robbé, M. V., de Vogel, V., & Stam, J. (2012). Protective factors for violence risk: The value for clinical practice. *Psychology*, *03*(12), 1259–1263.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry*, *57*, 316–331. https://doi.org/10.1111/j.1939-0025.1987.tb03541.x

Smith, K. J., O'Rourke, S., & Macpherson, G. (2020). The predictive validity of the HCR20V3 within Scottish forensic inpatient facilities: A closer look at key dynamic variables. *International Journal of Forensic Mental Health*, *19*(1), 1–17. https://doi.org/10.1080/14999013.2019.1618999

Ttofi, M. M., Farrington, D. P., Piquero, A. R., Lösel, F., DeLisi, M., & Murray, J. (2016). Intelligence as a protective factor against offending: A meta-analytic review of prospective longitudinal studies. *Journal of criminal justice*, *45*, 4–18. https://doi.org/10.1016/j.crimjus.2016.02.003

Tuente, S. K., Bogaerts, S., & Veling, W. (2021). Mapping aggressive behavior of forensic psychiatric inpatients with self-report and structured staff-monitoring. *Psychiatry Research*, *301*:113983. https://doi.org/10.1016/j.psychres.2021.113983

de Vogel, V., De Beuf, T., Shepherd, S., & Schneider, R. D. (2022). Violence risk assessment with the HCR-20V3 in legal contexts: A critical reflection. *Journal of Personality Assessment*, *104*(2), 252–264. https://doi.org/10.1080/00223891.2021.2021925

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2012). *SAPROF, guidelines for the assessment of protective factors for violence risk*. Forum Educatief.

de Vries Robbé, M., van den Nagel, A., Bohle, A., & Veldhuizen, A. (2017). *SAPROF long-term care: Additional guidelines for the assessment of protective factors for violence risk in intensive (forensic) psychiatry care*.

Yudofsky, S. C., Silver, J. M., Jackson, W., Endicott, J., & Williams, D. (1986). The Overt Aggression Scale for the objective rating of verbal and physical aggression. *The American Journal of Psychiatry*, *143*, 35–39. https://doi.org/10.1176/ajp.143.1.35

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ireland, J. L., Levtova, Y., Abi Semaan, C.-M., Steene, L. M. B., Henrich, S., Gaylor, L., Driemel, L., Volz, S., Homann, J., Dickopf, M., Greenwood, L., & Chu, S. (2024). Risk and protective factors in risk assessment: Predicting inpatient aggression in adult males detained in a forensic mental health setting. *Aggressive Behavior*, *50*, e22150. https://doi.org/10.1002/ab.22150