



Contents lists available at ScienceDirect

## International Journal of Law and Psychiatry



## Psychometric properties and predictive validity of a police version of a violence risk screen – A pilot study

John O. Roaldset<sup>a,b,c,\*</sup>, Pål Hartvig<sup>b</sup>, Stål Bjørkly<sup>b,d</sup>

<sup>a</sup> Ålesund Hospital, Psychiatric Department, Box 1600, 6026 Ålesund, Norway

<sup>b</sup> Oslo University Hospital, Centre of Research and Education in Forensic Psychiatry, Box 4956, Nydalen, 0424, Oslo, Norway

<sup>c</sup> Department of Mental Health, The Norwegian University of Science and Technology (NTNU), Box 8905, 7491 Trondheim, Norway

<sup>d</sup> Molde University College, Box 2110, 6402 Molde, Norway

### ARTICLE INFO

#### Article history:

Received 2 March 2017

Accepted 14 June 2017

Available online xxxxx

#### Keywords:

Screen

Risk assessment

Police officers

Prospective

Violence recidivism

### ABSTRACT

The Violence Risk Screening – Police Version (V-RISK-POL) is a seven-item screening tool for use at police stations by police officers and law enforcement officials to assist in the process of decision making regarding release, restrictive measures or arrest for apprehended individuals where the risk of future violence must be considered. The screen is based on the V-RISK-10, originally developed for emergency psychiatry. We examined psychometric properties and the prospective predictive validity of future violent convictions for the V-RISK-POL in a sample of 111 persons arrested for suspicion of violent crimes. Seventeen persons were convicted for a new violent crime committed during the 24–40 months follow-up. The V-RISK-POL demonstrated good internal consistency; Cronbach's alpha = 0.81 (95% CI = 0.75–0.86) and moderate predictive validity; the area under the curve of the receiving operator characteristics (AUC) = 0.753 (95% CI = 0.644–0.843). Further research on larger and more heterogeneous samples is necessary to examine whether the screen may be useful in the police context.

© 2017 Elsevier Ltd. All rights reserved.

### 1. Introduction

The past three decades have seen a change from reactive responding to proactive approaches in the prevention of violence in society through community policing (Chappel & O'Brien, 2014). At the start of this process, Borum, Deane, Steadman, and Morrissey (1998) claimed that this change actually was a paradigm shift from traditional law enforcement to community-oriented models. Police officers who encounter potentially violent incidents are confronted with a series of different decisions they are required to make. First, they have to evaluate the likelihood of an imminent risk of violence. Second, they must make a judgment about what the future risk for violence may be. Finally, they need to decide on how to prevent future risk. Being a police officer requires taking urgent action to prevent *imminent* violence and making the necessary decisions involves demanding personal and professional challenges. Still, the issues involved in how to prevent *future* incidents may be even more complicated. Typically, this kind of decision will involve choosing between, for instance, release, prosecution, restrictive measures such as arrest or restraining orders, or recommendation for psychiatric treatment. For many years, police officers and officials have determined the risk levels of violent persons by intuition (Dayan, Fox, & Morag, 2013).

Up to the early 1980s, such unstructured and intuitive decisions also dominated violence risk assessment in mental health services. However, from that time on, actuarial and structured professional judgment (SPJ) approaches gradually replaced the dominant unstructured risk assessment paradigm. The most important factor in this change was the development and validation of structured assessment tools for a wide variety of subgroups of violence.

About a decade ago, a parallel process of replacing unstructured and intuitive violence risk assessment with structured tools started in policing. This process increased the demand for further development of efficient risk assessment tools, and the current use of such tools to inform efforts to prevent violence is steadily growing in police work (Belfrage et al., 2012). Still, this change has not produced a very large number of assessment tools adapted for use in policing. The great majority of these instruments have been created for identifying risk for intimate partner violence (IPV). Examples of such tools are the Brief Spousal Assault Form for the Evaluation of Risk (BSAFER) (Kropp, Hart, & Belfrage, 2010), the Spousal Assault Risk Assessment Guide (SARA) (Kropp, Hart, Webster, & Eaves, 1999), and the Ontario Domestic Assault Risk Assessment Guide (ODARA) (Hilton et al., 2004). Even though there are relatively few empirical studies on how these tools work in policing, a brief look at some findings may be informative.

As part of a multicenter study, Belfrage et al. (2012) tested the predictive validity of the 20-item Spousal Assault Risk Assessment Guide (SARA). Overall the results indicated that the association between

\* Corresponding author at: Ålesund Hospital, Psychiatric Department, Box 1600, 6026 Ålesund, Norway.

E-mail address: john.olav.roaldset@ntnu.no (J.O. Roaldset).

SARA assessment and recidivism was small in magnitude. Results from testing of the predictive validity of summary risk ratings (low, moderate or high risk) were very weak.

The Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER) comprises 10 perpetrator risk factors and 5 victim vulnerability factors. Storey, Kropp, Hart, Belfrage, and Strand (2014) examined the B-SAFER used by Swedish police officers in the assessment and management of IPV. Results were positive with total scores and overall risk ratings that predicted recidivism. However, validation of the use of police management plans based on the B-SAFER showed decreased recidivism in high-risk perpetrators but increased recidivism in low-risk perpetrators. The authors suggested that the B-SAFER may be better suited than the SARA for police work.

The Ontario Domestic Assault Risk Assessment Guide (ODARA) is a 13-item, cross-validated actuarial assessment tool designed to estimate the risk of spousal/partner assault recidivism. It comprises both static and dynamic risk factors and, in contrast to the SARA and B-SAFER, it was originally developed for use by police officers, service workers, courts, and so forth. Results from tests of predictive validity have varied from good to poor (Ulmer, 2015).

The magnitude of predictive validity estimates for SARA, B-SAFER and ODARA have varied within each tool in different investigations. Thus, although some recent publications have reported on tests of the predictive validity of risk assessment tools for IPV used by the police, no firm conclusions can be made concerning their efficiency (Belfrage et al., 2012; Dayan et al., 2013; Ulmer, 2015). Furthermore, these tools require training and are quite time-consuming because of the relatively high number of items (13–20) and because some of the necessary background information is not easy to get. Obtaining information from social networks and, in particular, from previous and potential future victims is necessary for scoring items such as “Recent relationship problems” (Item 4, SARA), “Victim concern about future assaults” (Item 7, ODARA), and “Intimate Relationship Problems” (Item 7, B-SAFER). It is not always easy to contact the relevant individuals, and some of them may be reluctant to provide any information at all. To rely only on perpetrator self-report is inadvisable.

The V-RISK-POL is a seven-item violence risk checklist developed for use at police stations by police officers and law enforcement officials. The checklist is based on the Violence Risk Checklist-10 (V-RISK-10) (Bjorkly, Hartvig, Heggen, Brauer, & Moger, 2009), an instrument that was developed for screening violence risk in inpatient settings and after discharge from emergency psychiatric units. The V-RISK-10 was developed in response to the need for a violence risk instrument that could easily be used in general and particularly in emergency psychiatric units. Such units are characterized by high turnover, high time pressure, often inexperienced professionals on duty, and being open around the clock all year round. The existing instruments at that time, such as the HCR-20 and others, were comprehensive, time-consuming and required qualified expertise. The V-RISK-10 was developed based on the results of a pilot project exploring 33 selected risk factors (Hartvig, Alfarnes, Østberg, Skjønberg, & Moger, 2006). It consists of 10 items: 5 historical items (previous or present physical violence, previous or present threats of violence, previous or present substance abuse, previous or present severe mental illness, and personality disorders), 3 clinical items (lack of insight, suspiciousness, and lack of empathy) and 2 future/risk management items (unrealistic plans and stress exposure).

Tests of interrater reliability have shown that the screen has good interrater reliability (intraclass correlation coefficient (ICC) = 0.87) (Bjorkly et al., 2009; Roaldset, Hartvig, & Bjorkly, 2011). Research has also shown that the tool has high predictive accuracy for inpatient violence (Area under the curve [AUC] of the receiver operator characteristics of 0.83) and of outpatient violence within 3 months (AUC = 0.80) and 12 months (AUC = 0.75) after discharge (Hartvig, Roaldset, Moger, Østberg, & Bjorkly, 2011; Roaldset et al., 2011). The checklist also significantly identified violent behavior within the first year after discharge from psychiatric hospital in persons without previously

known violent behavior and showed higher predictive accuracy for seriously violent acts compared with moderate or light violent acts and violent threats (Roaldset et al., 2011).

The V-RISK-10 was transformed into a seven-item police-version tool V-RISK-POL. This was done by combining the two items “physical violence” and “threats of violence” into one item (violence) and by excluding two items, “personality disorders” and “unrealistic plans”, from the original V-RISK-10 (Roaldset et al., 2011). There are some parallels between the use of V-RISK-POL in police work and the V-RISK-10 in acute psychiatry. Both services have a high turnover of service users, high time pressure for evaluation and assessment, and the obligation to be operative 24 h a day all year round. Furthermore, many arrested persons with violence issues show emotional instability or display conduct problems or psychiatric symptoms (Lamb, Weinberger, & DeCuir, 2002).

Still, there are also distinct differences between the competency of police officers and mental health professionals. Medical doctors and specialists in psychiatry and clinical psychology are qualified for diagnosing severe mental illness and personality disorders (V-RISK-POL, Item 3). They also have high competence in assessing dynamic factors, such as lack of insight (Item 4), suspiciousness (Item 5), and lack of empathy (Item 6). However, their information regarding previous violence (Item 1) from patients outside the treatment context is limited, and they often have to contact the police for further information. Because previous violence is the strongest predictor of future violence, this information is a cornerstone in risk assessment of violence. Police officers are clearly closer to this information than mental health professionals are. Another advantage for police officers is the availability of direct observations or information from colleagues concerning a given individual's behavior in very stressful interactions (Item 7) before, during, or after violent incidents. They can also often closely observe how the influence of substance abuse (Item 2) works on a person's cognitive, emotional and behavioral coping.

To our knowledge, validation research on violence risk judgment in policing has, so far, only used tools made for assessment of risk for IPV. The fact that police are confronted with other types of violence more often than they are with IPV was one of the main reasons for the development of the V-RISK-POL. We also wanted to create a screening instrument that is easy to use and time-efficient. The checklist does not require extensive training. The aims of this prospective, naturalistic investigation were to test the reliability and the predictive validity of the V-RISK-POL ratings for violent convictions during a follow-up period after initial arrest.

## 2. Methods

### 2.1. Design

This was an observational prospective study conducted in a naturalistic context. Violence risk was assessed at the time of arrest (baseline) and compared with convictions for violent crimes during the 24 to 40 months follow-up period.

### 2.2. Setting

From August 2012 to December 2013, persons who were arrested and suspected of violence were screened with the V-RISK-POL checklist by police officers and law enforcement officials at Sunnmøre police district. The police district consisted of four police stations and covered about 130,000 inhabitants in a semirural and small-town area on the west coast of Norway.

### 2.3. Sample

The target population consisted of all persons arrested for violence or threats of violence ( $N = 133$ ). Seven persons arrested as suspects

of other crimes were also included because the police were very concerned about their risk of future violence. There were 10 duplicates (the same person had two arrests/assessments) for which the second assessment was excluded. The rationale for this was to avoid enhanced results due to repeat assessments for these individuals. Nineteen persons were excluded due to incomplete data ( $n = 11$ ), being under investigation as suspects of crime ( $n = 4$ ), or being unable to locate ( $n = 4$ ). The final sample consisted of 111 persons, 95 men and 16 women.

#### 2.4. Characteristics of the sample at baseline

The mean age for men was 29 years, which was significantly lower than for women (35 years) ( $p = 0.023$ ). Of the arrested persons, 54 were suspected of violence against strangers; 32, of violence against victims they knew (but not family); and 25, of violence against family members. Two were suspected of violence against both unknown and known (not family) victims. At the time of the arrest, 57 offenders (51%) were under the influence of narcotics or alcohol but their victims were not, and in 35 situations, (32%) both offenders and victims were under the influence. Only 19 offenders (17%) and their victims were not under the influence. Fourteen persons (13%) were not Norwegian citizens. At baseline, restrictive measures had been imposed on 35 persons (32%): custody ( $n = 9$ , 8%), psychiatric assessment ( $n = 11$ , 10%), restraining order ( $n = 4$ , 4%), or other measures ( $n = 11$ , 10%), such as a security alarm for the victims. Restrictive means were imposed on 10 persons (19%) suspected of violence against strangers; on 11 persons (34%), of violence against victims they knew (but not family); and on 13 persons (52%), of violence against family members.

#### 2.5. Procedure

During the inclusion period, 52 police officers or law enforcement officials at the police stations scored one or more forms containing the baseline measures ( $mean = 2.3$ ,  $median = 1$ ,  $range 1–9$ ). Scores were subsequently transferred to SPSS using a computer program to scan the forms. At the end of the follow-up period, registry data from the Norwegian police records (STRASAK) containing data on arrests and convictions was coupled with the baseline data to identify convictions of violence during follow-up.

#### 2.6. Baseline measures

##### 2.6.1. V-RISK-POL

As described above, V-RISK-POL is a short violence risk checklist, drafted as a police version of the V-RISK-10. The target user group is police officers and law enforcement officials who are in the “first line” in police stations when deciding upon release, arrest, or the use of other restrictive means for potentially violent apprehended persons. The purpose of the check-list is to provide a duty officer or lawyer (or another police officer) with a broader basis for determining whether action should be taken to prevent future incidents of violence. V-RISK-POL is designed to be used by both experienced and inexperienced police officers and law enforcement officials around the clock throughout the year. Rating of the check-list should be based on the information present at the time, and collection of further information is not required. In cases with insufficient information, “Do not know” is the correct choice for the actual item in the check-list. At police stations (as in psychiatric emergency departments) it is very difficult, or rather impossible, that all shifts on duty include a police-officer or law enforcement official that has been trained to use a violence risk checklist. Because of that, the V-RISK-POL was designed as a self-explanatory checklist with all necessary instructions for use in the coding sheet. There are no requirements for training or further education before use.

V-RISK-POL consists of seven items. Each item has four scoring categories: (i) *No – not present*, (ii) *Moderate - partly present*, (iii) *Yes –*

*definitely present*, and (iv) *Do not know if present, insufficient information*. The seven items are as follows:

- (1) Physical violence or threats of violence, previous or current.  
*Moderate:* Medium severe aggressive acts; such as kicks, blows and shoving that do not cause severe harm to the victim. Verbal or physical threats with low to moderate severity.  
*Yes:* Severe physical attack (including use of dangerous objects) toward another person with intent to inflict severe physical harm. Severe verbal or physical threats.  
*Verbal threats:* Statements, yelling and the like, that involve threats of inflicting physical harm on another person.  
*Physical threats:* Movements and gestures that warn of a physical attack.
- (2) Substance abuse, previous or current.  
The person has a history of abusing alcohol, medication and/or other substances (e.g. amphetamine, heroin, cannabis). Abuse of solvents or glue should be included.
- (3) Serious mental illness, previous or current.  
The person is or has been in a psychotic condition characterized by hallucinations, delusions, severe confusion, etc.
- (4) Lack of insight concerning own behavior and/or mental illness.  
The degree to which the person lacks insight into his or her deviant behavior or having a severe mental disorder, see item 3.
- (5) Suspiciousness.  
The person expresses suspicion toward other individuals either verbally or nonverbally. The person appears to be “on guard” toward the environment.
- (6) Lack of empathy.  
The person appears emotionally cold and without sensitivity to others' thoughts or emotional experiences. This is particularly serious if it concerns a recent or former victim of his or her acts of violence.
- (7) Exposure to future stress.  
The person is vulnerable to future stress and stressful situations, and his or her ability to cope with stress without resorting to violence is insufficient. Examples: Other persons setting limits, proximity to potential victims of violence, substance abuse, homelessness, stays in a violent environment, easy access to weapons.

Settings with high turnover and limited time to gather necessary information may generate a fair amount of “Do not know if present” scores. It is neither correct to let these count as “No, not present” in the statistics nor to exclude them. Recent research on the V-RISK-10 in emergency psychiatry has shown that the “Don't know if present” option has significant predictive validity for violent behavior (Eriksen et al., 2016; Roaldset & Bjørkly, 2010). According to this, the sum score for all items in V-RISK-POL was calculated in two ways, with the following values for each item: First with *No = 0*, *Do not know = 1*, *Moderate = 2*, *Yes = 3*, with a sum-score range of 0–21; and second with *No = 0*, *Do not know = 0*, *Moderate = 1*, *Yes = 2*, with a sum-score range of 0–14.

##### 2.6.2. Overall risk assessment

After completing the screen, the overall presence and relevance for future violence are considered for all 7 items. According to this routine, law enforcement officials or police officers scored an “overall risk assessment” categorized into “low, moderate, or high” risk for future violence. The ratings were based on all information available to the duty officers at the time, including scores on the V-RISK-POL. For obvious ethical reasons, the duty officers could not limit basing their judgments on the information from the V-RISK-POL alone nor determine further action without taking into consideration potential risk identified by using the V-RISK-POL. Including overall risk assessment estimates and considering relevant risk management strategies comply with the

procedure of the structured professional judgment tradition (Douglas, Ogloff, & Hart, 2003).

### 2.7. Outcome measures

The outcome variable, *violence*, was defined as having at least one conviction for physical violence, violent threats, or arson. Further specifications and operational definitions of physical violence and violent threats were given in 2.6.1 V-RISK-POL. Solved cases of violence committed during follow-up that were dismissed due to insanity were also included. All types of recorded violence were combined into one dichotomous variable (*any violence*). This was done in order to obtain a larger number of outcome episodes and higher statistical power per analysis.

### 2.8. Statistics

Chi-square tests were used to compare categorical variables and *t*-tests were used to compare continuous variables. Internal consistency was analyzed by Cronbach's alpha, and the single item's contribution by "Cronbach's alpha if Item Deleted". Analysis of the area under the curve (AUC) of the receiver operating characteristic (ROC) was performed to assess the overall predictive accuracy of the V-RISK-10 concerning violent behavior during follow-up. AUC is sensitivity/(1 - specificity) plot of all the values of the tool. The AUC value ranges from 0 to 1, where an area of 0.5 equals chance and 1.0 equals a perfect prediction. Binary logistic regression analyses were conducted to obtain univariate and multivariate effect sizes (odds ratio, OR). To make a calculation of the model fit (explained variance); how much of the total variance of violent convictions that could be explained by the variable(s) in the analysis, two  $R^2$  substitute estimates (Cox & Snell and Nagelkerke) were used as the lower and the upper limits (Brace, Kemp, & Snelgar, 2006).

Statistically significant AUCs and ORs do not necessarily mean that a decision-making method is useful in clinical practice. In line with previous recommendations (Buchanan, 2008; Deeks & Altman, 2004; Singh, Yang, & Mulvey, 2015), the following predictive validity estimates of violence risk during follow-up were also computed: (1) sensitivity (the proportion of the violent persons that had a positive test), (2) specificity (the proportion of non-violent persons who had a negative test), (3) positive predictive value (PPV, the proportion of those judged to be at risk of violence who actually were violent), (4) negative predictive value (NPV, the proportion of those judged not to be at risk of violence who actually did not exhibit violence), (5) number needed to assess (NNA, how many persons would need to be assessed to identify one truly violent person), and (6) the likelihood ratio (LR). The LR determines the extent to which the odds of an outcome (e.g., violence) increase when a test is positive (LR+) and decrease when a test is negative (LR-) (Deeks & Altman, 2004). Data were analyzed using SPSS version 23.0.

### 2.9. Ethics

The study was approved by the Regional Committee for Medical and Health Research Ethics (REK) and the Norwegian Police Directorate. The approval granted exemption from informed consent to participate.

## 3. Results

### 3.1. Prevalence of violence

During follow-up (24–40 months), 17 persons (15%) committed a violent crime that led to a conviction or was dismissed due to insanity, 14 men (14%) and 3 women (18%). Six men were convicted for violent threats, five men for violent assault, one man for arson, and two men for homicide. Two women were convicted for violent assault and one for

violent threats. Mean age of the convicted men (27 years) and women (30 years) was not significantly different. All of the convicted persons during follow-up were Norwegian citizens. Every person convicted of a violent crime during follow-up had been under the influence of drugs or alcohol at baseline. Influence of drugs or alcohol at baseline was a significant risk factor for committing a violent crime during follow-up ( $\chi^2 = 4.2$ ,  $df = 1$ ,  $p = 0.042$ ). Five (14%) of the 35 persons given restrictive measures were convicted for violent crimes, while 12 (16%) of the 76 persons without restrictive measures were convicted.

### 3.2. V-RISK-POL

#### 3.2.1. Sum-scores

The mean V-RISK-POL sum-score was 9.1 (95% CI 8.1–10), and there was no significant score difference between men ( $M = 9.2$ , 95% CI 8.2–10) and women ( $M = 8.1$ , 95% CI 5.2–10). The sum-score for those convicted of violent crime during follow-up was significantly higher compared with those not convicted for violent crime; 12.0 (95% CI 10–14) versus 8.6 (95% CI 7.6–9.5),  $p = 0.002$ . The sum-score for those on whom restrictive measures at arrest/baseline were imposed was significantly higher than for those on whom restrictive measures were not imposed, respectively 12.0 (95% CI 9.7–14.0) versus 7.8 (95% CI 7.0–8.6),  $p < 0.001$ . There was no significant difference between the sum-scores of those who were influenced by narcotics or alcohol at arrest/baseline and those not influenced, respectively 8.9 (95% CI 7.9–9.9) versus 8.5 (95% CI 6.0–11.0),  $p = 0.561$ .

The V-RISK-POL sum-scores of the Overall risk ratings were 5.5 (95% CI 4.7–6.6) for "Low risk", 9.6 (95% CI 8.5–11.0) for "Moderate risk", and 16.0 (95% CI 14.0–17.0) for "High risk". All were significantly different at the  $p < 0.001$  level. There were no significant differences between the sum-scores for those who at arrest/baseline were suspected of violence against unknown victims, 8.6 (95% CI 7.2–10.0), those who were suspected of violence against known victims (but not family) 8.0 (95% CI 6.4–9.6), and those who were suspected of violence against family members, 11.0 (95% CI 8.8–13.0).

#### 3.2.2. Internal consistency

The internal consistency (Cronbach's alpha) for the scores on the seven items of V-RISK-POL was 0.811 (95% CI 0.752–0.861),  $p < 0.001$ . Cronbach's alpha for single measures was 0.411 (95% CI 0.334–0.497),  $p < 0.001$ . When one item of the scale was deleted and the other entered, ("Cronbach's alpha if Item Deleted"), the highest value (alpha = 0.85) was obtained for V1 (violence) and V2 (substance abuse). The lowest value = 0.78 was obtained for V5 (suspiciousness). The value for all the other items, V3 (serious mental illness), V4 (lack of insight), V6 (lack of empathy) and V7 (exposure to future stress), was about 0.79.

#### 3.2.3. Predictive validity of single items

The odds ratios for the single items for future violence are displayed in Table 1. The low values of the reference category "No" for Item V1 caused it not to fit the logistic regression model, so, for this item, the reference category was computed by adding the scores for the "No" and "Don't know" categories.

#### 3.2.4. Predictive validity of sum-scores

The predictive validity (ROC-AUC) of the V-RISK-POL screening is displayed in Table 2. We found no significant difference between using the "Don't know = 1" with a maximum score of 21 and "Don't know = 0" with a maximum score of 14. The "Don't know = 1" was used in the further analyses because it increased the score variance.

When suspected violence at arrest was categorized into violence against strangers ( $n = 54$ ), violence against known persons (not family,  $n = 32$ ), and violence against family members ( $n = 25$ ), the AUC values of V-RISK-POL for convictions of any violent crime during follow-up were 0.765 (95% CI = 0.599–0.932),  $p = 0.018$ , for violence against strangers (8 convictions), 0.745 (95% CI = 0.579–0.911),  $p = 0.040$ ,

**Table 1**  
V-RISK-POL: Predictive values of the single items for future violence, adjusted for age and gender.

Scoring Item	No			Don't know			Moderate			Yes		
	n (reference)	n	OR (95% CI)	p	n	OR (95% CI)	p	n	OR (95% CI)	p		
V1: Violence <sup>a</sup>	7	22	–		53	0.7 (0.16–34)	0.679	36	2.3 (0.53–10)	0.270		
V2: Severe mental illness	27	32	1.7 (0.14–21)	0.676	14	8.3 (0.92–97)	0.091	45	8.3 (0.94–73)	0.057		
V3: Substance abuse	40	56	7.4 (0.88–63)	0.065	9	15.0 (1.3–176)	0.030	13	21 (2.0–219)	0.012		
V4: Lack of insight	43	45	1.5 (0.33–7.1)	0.584	16	6.2 (1.2–32)	0.029	14	3.7 (0.63–21)	0.150		
V5: Suspicion	52	40	5.3 (1.0–28)	0.049	17	11.0 (1.9–69)	0.008	9	10.0 (1.1–92)	0.042		
V6: Lack of empathy	49	40	2.0 (0.51–8.3)	0.311	13	2.6 (0.39–17)	0.327	14	7.6 (1.4–42)	0.020		
V7: Future stress	23	48	1.4 (0.25–8.1)	0.695	27	1.6 (0.26–10)	0.606	18	3.4 (0.52–23)	0.201		

<sup>a</sup> Due to low “No” numbers, the reference group was composed of the No + Don't know groups.

for violence against victims they knew (not family, 8 convictions), and 0.663 (95% CI = 0.311–1.00),  $p = 0.453$ , for violence against family members (2 convictions).

**3.2.5. Multivariate logistic regression analyses**

Results of univariate regression analyses and multivariate logistic regression analysis including age, gender, restrictive means, the Overall risk assessment score and the V-RISK-POL sum-scores, are displayed in Table 3. V-RISK-10 and restrictive measures were the only variables that remained significant in the multivariate analyses. V-RISK-10 yielded an increase in OR for each point of the scale (0–21) of 1.28 (95% CI = 1.01–1.58,  $p = 0.034$ ). The incremental validity estimate (how many more of the violent crimes that could be explained by adding the variable) increased from 5 to 10% (from 11 to 19% to 16–29%),  $\chi^2 = 5.1$  ( $df = 1$ ),  $p = 0.024$ , when V-RISK-POL was added to the other four variables. The explained variance of all the five variables in the analysis was 16 to 29%, with  $\chi^2 = 16.3$ , ( $df = 5$ ),  $p = 0.006$ .

All persons convicted of a violent crime during follow-up had been under the influence of narcotics or alcohol at arrest at baseline. Hence, it was not possible to dichotomize the “influenced by alcohol or narcotics at arrest” variable for the multivariate logistic regression analysis.

**3.2.6. Other predictive validity estimates**

Examining all the coordinates on the sensitivity – specificity plot of the AUC-curve (i.e., the sensitivity and specificity of each of the 0–23 possible sum-scores) revealed that cut-off at 8.5 for V-RISK-POL sum-score gave the best combination of sensitivity and specificity. The best combination was defined as the highest sum of sensitivity + specificity. With this cut-off, the following predictive measures were obtained: sensitivity = 82%, specificity = 63%, positive predictive value (PPV) = 28%, negative predictive value (NPV) = 95%, number needed to assess (NNA) = 3.6, positive likelihood ratio (LR+) = 2.2, and negative likelihood ratio (LR-) = 0.23.

**Table 2**  
ROC-AUC values of V-RISK-POL.

	AUC	(95% CI)	p-value
V-RISK-POL (score 0–21) <sup>a</sup>	0.753	0.644–0.863	0.001
V-RISK-POL (score 0–14) <sup>b</sup>	0.733	0.616–0.850	0.002
V-RISK-POL 6 items (V1 violence excluded) <sup>a</sup>	0.747	0.633–0.862	0.001
V-RISK-POL 4 items (Dynamic items, V4–V7) <sup>a</sup>	0.700	0.574–0.825	0.009
<i>Single items</i>			
V1 previous or current violence <sup>a</sup>	0.600	0.452–0.747	0.191
V2 previous or current substance abuse <sup>a</sup>	0.712	0.592–0.833	0.005
V3 previous or current severe mental illness <sup>a</sup>	0.724	0.605–0.843	0.003
V4 lack of insight <sup>a</sup>	0.657	0.519–0.795	0.040
V5 Suspiciousness <sup>a</sup>	0.706	0.585–0.827	0.007
V6 Lack of empathy <sup>a</sup>	0.629	0.486–0.778	0.089
V7 exposure for future stress <sup>a</sup>	0.619	0.470–0.768	0.119

<sup>a</sup> Sum-score 0–21 based on: No = 0, Don't know = 1, Maybe/Moderate = 2, Yes = 3.

<sup>b</sup> Sum-score 0–14 based on: No = 0, Don't know = 0, Maybe/Moderate = 1, Yes = 2.

**3.3. The overall risk assessment of low, moderate, or high risk**

The baseline assessment of “low-moderate-high” risk for future violent crimes showed that 42 persons were assessed as “low risk”, 53 persons as “moderate risk”, and 16 persons as “high risk”. The ROC-AUC for the Overall risk assessment was 0.704 (95% CI = 0.546–0.863),  $p = 0.011$ .

**4. Discussion**

The main finding indicates that the V-RISK-POL had good internal consistency and moderate to good predictive validity. Results were promising when taking into consideration the naturalistic, prospective design and that the raters had no training in using the screen and were not experts on risk assessment of violence. The findings indicate that the screen was more accurate for general violence than for domestic violence. Current or prior violence (Item 1) was not a significant predictor. This departs from other research that has found past violence to be a very strong risk factor. Our results may be due to the fact that violence was a characteristic of the sample (suspected for violence was inclusion criterion) and that this limited the predictive value of this item. Still, the significant findings for substance abuse, severe mental illness, lack of insight, suspiciousness, and lack of empathy suggest that other common factors had good explanatory power. Further research is needed to confirm this result.

The use of restrictive measures might have had an impact on violent behavior during follow-up. The scores of the V-RISK-POL (Table 1) indicate that restrictive means might have prevented violence. In this way a possible true positive prediction may have been turned into a false positive, resulting in reduced predictive power of the tool. This preventive effect may particularly explain the non-significant results for individuals suspected of family violence. More than half of them were given restrictive measures at baseline and only two were convicted of a new violent crime.

The likelihood ratio analyses showed that persons who scored above cut-off had a 2.2 fold increased probability of being convicted for a violent crime (“ruling in effect”) during follow-up compared with those scoring below cut-off. Persons scoring below cut-off had a 4 times higher probability of *not* being convicted (“ruling out” effect). The likelihood ratios, the high sensitivity, low specificity, low positive predictive

**Table 3**  
Uni- and multivariate analysis of V-RISK-POL and other baseline measures.

	Univariate			Multivariate		
	OR	(95% CI)	p-Value	OR	(95% CI)	p-Value
Age	0.97	(0.91–1.03)	0.299	0.96	(0.88–1.03)	0.237
Female gender	1.4	(0.34–5.4)	0.668	5.1	(0.85–31)	0.074
Restrictive measures	0.90	(0.29–2.8)	0.838	0.17	(0.03–1.002)	0.050
Overall risk assessment	3.6	(1.5–8.4)	0.004	2.3	(0.61–8.8)	0.219
V-RISK-POL	1.2	(1.1–1.3)	0.002	1.3	(1.0–1.6)	0.034

power and high negative predictive power indicate that the screen is better at predicting who will not be violent than who will be violent. This finding is in line with results from other risk assessment tools. High sensitivity, which means that the greater part of the violent cases are identified, is a precondition for being an efficient checklist or screen. The low specificity and positive predictive power results were a weak point in this study. However, the consequence of this lack of assessment accuracy is mainly that restrictive measures may be considered for more persons than necessary. However, a screen should always be seen as a first step in a stepwise assessment process. Hence, the negative consequence of high false positive rates is that further assessment should be done before a decision is taken on restrictive measures.

Finally, we found a significant correlation between V-RISK-POL scores and the Overall assessment scores (low-moderate-high). However, AUC values for the Overall assessments did not improve the predictive validity over the sum-scores of V-RISK-POL in this study. This may indicate that the V-RISK-POL covers relevant risk factors and that the police officers did not find any other information that would have improved the accuracy of their risk judgments. One should also bear in mind that a possible preventive effect of restrictive measures, turning true positives into false positives, might, in particular, have affected the results of the Overall assessment. Still, our investigation did not generate data to conclude whether the V-RISK-POL scores had a significant impact on restriction decisions or not.

#### 4.1. Strengths

One advantage of our pilot study is the prospective design and the naturalistic design. The high number of raters in the study, randomly selected depending on who was on duty when the screening was done, increases the external validity of the findings. The low sample attrition rate (21%) and a strict definition of having committed violence (base rate 15%) was another asset. One major asset of this type of naturalistic design is its strong external validity. There were no extra time or financial resources allocated to the police officers and law enforcement officials for participation in this research. It is very likely that this new routine can be implemented in most policing contexts that carry out unstructured violence risk screening.

#### 4.2. Limitations

The small sample size in this pilot project may have caused low statistical power and increased likelihood of Type-II errors. Large confidence intervals (in particular for results from logistic regression analysis) and only one police district being involved limit the external validity of the investigation. There was a lack of information concerning the duration of the initial detention for the nine who were taken into custody at baseline or whether some of the participants in the study had been detained during follow-up. This may have resulted in a shorter time at risk for committing a violent crime for some participants. Participants for whom restrictive measures had been imposed at baseline had higher scores on V-RISK-POL compared to participants without restrictive measures. Hence, this methodological limitation may have reduced rather than enhanced the results from testing the accuracy of the screen.

Another limitation is the predictive study design. Even though the use of check-lists as part of security procedures has become increasingly common in many fields of society and has shown some promising results, the effect of the use of such check-lists is still inconclusive. More important than having good measures with predictive accuracy is having risk assessment check-lists, and more comprehensive risk assessment tools, that actually contribute to mitigating risk of violence. Our literature search yielded no such “pre – post” studies from policing or law enforcement, and we found only three studies from psychiatry. All three showed 30 to 50% reduction of violent episodes and in the use of coercive measures in psychiatric departments after implementing daily use of a short 6-item check-list recording “warning signals” in

the individual patient (Brøset Violent Checklist, BVC) (Abderhalden et al., 2008; Hvidhjelm, 2015; van de Sande et al., 2011). Despite the challenges posed both in designing reliable measures of significant “pre-post” effects and in implementing “pre-post” projects, which entail significant resources, more research with this design is recommended.

## 5. Conclusion

Police officers are necessarily given a broad mandate to protect individual safety and public order. They must make decisions in the moment about how to respond to threatening situations. Obviously, a checklist such as the V-RISK-POL cannot provide immediate aid to calm down a violent escalation. Still, after a situation is physically under control, police officers are confronted with very challenging decisions in response to questions like: Can we send the individual back to his family? Should we keep him until tomorrow? Does he need a restrictive order or should he be incarcerated? If supported by further research, it is our view that the V-RISK-POL may inform such decisions by offering a structured approach to violence risk assessment.

In this study the V-RISK-POL seemed to be a user-friendly screen, indicating that it may easily be implemented in other police settings. Future research may want to involve large-scale studies involving more police districts and international/foreign samples. Controlling for the effect of implemented risk management strategies, expanding operationalization of violence beyond that of convictions (such as self-report and collateral reports of violence) may be important improvements. However, the most important issue is whether the use of checklists can have an impact on reducing rates of violence. Obviously, studies of predictive validity will not bring us to this end. Still, accurate measurement of violence risk by police officers may be an important first step in this direction.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Acknowledgments

We are thankful to police chief Ingar Bøen, police inspector Terje Rishaug, and all the law enforcement officials and police officers at Sunnmøre police district who took part in the project.

## References

- Abderhalden, C., Needham, I., Dassen, T., Halfens, R., Haug, H. J., & Fischer, J. E. (2008). Structured risk assessment and violence in acute psychiatric wards: Randomised controlled trial. *British Journal of Psychiatry*, 193(1), 44–50. <http://dx.doi.org/10.1192/bjp.bp.107.045534>.
- Belfrage, H., Strand, S., Storey, J. E., Gibas, A. L., Kropp, P. R., & Hart, S. D. (2012). Assessment and management of risk for intimate partner violence by police officers using the spousal assault risk assessment guide. *Law and Human Behavior*, 36, 60–67.
- Bjorkly, S., Hartvig, P., Heggen, F. A., Brauer, H., & Moger, T. A. (2009). Development of a brief screen for violence risk (V-RISK-10) in acute and general psychiatry: An introduction with emphasis on findings from a naturalistic test of interrater reliability. *European Psychiatry*, 24, 388–394.
- Borum, R., Deane, M. W., Steadman, H. J., & Morrissey, J. (1998). Police perspectives on responding to mentally ill people in crisis: Perceptions of program effectiveness. *Behavioral Sciences & the Law*, 16, 393–405.
- Brace, N., Kemp, R., & Snelgar, R. (2006). *SPSS for psychologists*. New York: Palgrave Macmillan.
- Buchanan, A. (2008). Risk of violence by psychiatric patients: Beyond the “actuarial versus clinical” assessment debate. *Psychiatric Services*, 59, 184–190.
- Chappel, D., & O'Brien, A. (2014). Police responses to persons with a mental illness; international perspectives. *International Journal of Law and Psychiatry*, 37, 321–324.
- Dayan, K., Fox, S., & Morag, M. (2013). Validation of spouse violence risk assessment inventory for police purposes. *Journal of Family Violence*, 28, 811–821.
- Deeks, J. J., & Altman, D. G. (2004). Diagnostic tests 4: Likelihood ratios [comment]. *BMJ*, 329, 168–169.
- Douglas, K. S., Ogloff, J. R., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services*, 54, 1372–1379.

- Eriksen, B. M. S., Bjørkly, S., Færden, A., Friestad, C., Hartvig, P., & Roaldset, J. O. (2016). Gender differences in the predictive validity of a violence risk screening tool: A prospective study in an acute psychiatric ward. *International Journal of Forensic Mental Health, 15*(2), 176–187.
- Hartvig, P., Alfames, S., Østberg, B., Skjønberg, M., & Moger, T. (2006). Brief checklists for assessing violence risk among patients discharged from acute psychiatric facilities: A preliminary study. *Nordic Journal of Psychiatry, 60*, 243–248.
- Hartvig, P., Roaldset, J. O., Moger, T. A., Østberg, B., & Bjørkly, S. (2011). The first step in the validation of a new screen for violence risk in acute psychiatry: The inpatient context. *European Psychiatry, 26*, 92–99.
- Hilton, N. Z., Harris, G. T., Rice, M. E., Lang, C., Cormier, C. A., & Lines, K. J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario domestic assault risk assessment. *Psychological Assessment, 16*, 267–275.
- Hvidhjelm, J. (2015). *Prevention of aggression and violence on psychiatric wards via structured risk assessment – Evaluation of the Brøset violence checklist in Denmark*. (Doctoral dissertation) Faculty of Health and Medical Sciences, University of Copenhagen, 113–128.
- Kropp, P. R., Hart, S., & Belfrage, H. (2010). *Brief spousal assault form for the evaluation of risk (B-SAFER), version 2: User manual*. Vancouver, Canada: Proactive Resolutions Inc.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1999). *Spousal Assault Risk Assessment Guide (SARA)*. Toronto: Multi-Health Systems, Inc.
- Lamb, H. R., Weinberger, L. E., & DeCuir, W. J. (2002). The police and mental health. *Psychiatric Services, 53*, 1266–1271.
- Roaldset, J. O., & Bjørkly, S. (2010). Patients' own statements of their future risk for violent and self-harm behaviour: A prospective inpatient and post-discharge follow-up study in an acute psychiatric unit. *Psychiatry Research, 178*, 153–159.
- Roaldset, J. O., Hartvig, P., & Bjørkly, S. (2011). V-RISK-10: Validation of a screen for risk of violence after discharge from acute psychiatry. *European Psychiatry, 26*, 85–91.
- Singh, J. P., Yang, S., & Mulvey, E. (2015). Reporting guidance for violence risk assessment predictive validity studies: The RAGEE statement. *Law and Human Behavior, 39*, 15–22. <http://dx.doi.org/10.1007/s10979-011-9278-0>.
- Storey, J. E., Kropp, R., Hart, S. D., Belfrage, H., & Strand, S. (2014). Assessment and management of risk for intimate partner violence by police officers using the Brief spousal assault form for the evaluation of risk. *Criminal Justice and Behavior, 41*(2), 256–271. <http://dx.doi.org/10.1177/0093854813503960>.
- Ulmer, J. (2015). *The Ontario domestic assault risk assessment (ODARA): A validation and comparison study for an Oregonian law enforcement agency*. Doctor of psychology (PsyD) (Paper 177, <http://digitalcommons.georgefox.edu/psyd/177>).
- van de Sande, R., Nijman, H. L., Noorthoorn, E. O., Wierdsma, A. I., Hellendoorn, E., van der Staak, C., & Mulder, C. L. (2011). Aggression and seclusion on acute psychiatric wards: Effect of short-term risk assessment. *British Journal of Psychiatry, 199*(6), 473–478. <http://dx.doi.org/10.1192/bjp.bp.111.095141>.