

The Suicide Risk Assessment and Management Manual (S-RAMM) Validation Study 1

Atif Ijaz, Alexia Papaconstantinou, Helen O'Neill, Harry G Kennedy

Ir J Psych Med 2009; 26(2): 54-58

Abstract

Objective: There are validated tools for structured professional judgement of risk of violence, but few for risk of suicide. The Suicide Risk Assessment and Management Manual (S-RAMM) is a new structured professional judgement tool closely modelled on the HCR-20. This is the first validation study for the S-RAMM. We measured inter-rater reliability, internal consistency, concurrent validity with another validated risk instrument (HCR-20) and with a measure of psychopathology (PANSS). We tested whether the tool could distinguish between groups of patients clinically assessed as at varying levels of risk of suicide or self harm.

Method: Two researchers jointly interviewed 25 current in-patients for inter-rater reliability (Cohen's kappa) and internal consistency (Cronbach's alpha) and interviewed 81 of 83 current in-patients to assess whether the mean scores for different wards were significantly different (using ANOVA). Two other researchers made independent ratings of the HCR-20 and PANSS.

Results: Inter-rater reliability was acceptable for all items (Cohen's kappa >0.5 for all but three items) and all sub-scale and total scores (Spearman correlations all >0.8). Internal consistency was high, (Cronbach's alpha all sub-scales >0.6). Scores stratified significantly with high scores for admission and intensive care units and progressively lower scores in rehabilitation and pre-discharge units. The HCR-20 historical and S-RAMM background scores did not correlate but the dynamic sub-scales correlated significantly. PANSS scores also correlated significantly with S-RAMM scores.

Conclusion: The S-RAMM has better than minimum acceptable characteristics for use as a clinical or research tool. Prospective studies of sensitivity and specificity are now required.

Key words: Suicide, Violence, Risk, Assessment, Structured Professional Judgement, Validity, S-RAMM, HCR-20.

Atif Ijaz, MB, MRCPsych, Psychiatric registrar, **Alexia Papaconstantinou**, MB, MRCPsych, Psychiatric registrar, **Helen O'Neill**, MB, FRCPSych, Consultant Forensic Psychiatrist, Central Mental Hospital, Dundrum, Dublin 14, Ireland.
***Harry G Kennedy**, BSc, MD, FRCPI, FRCPSych, Consultant Forensic Psychiatrist, Clinical Professor of Forensic Psychiatry (University of Dublin, Trinity College, Dublin, Ireland), Clinical Director, Central Mental Hospital, Dundrum, Dublin 14, Ireland.
 Email: Harry.kennedy@ireland.com

*Correspondence

SUBMITTED: JANUARY 4, 2008. ACCEPTED: JANUARY 23, 2009.

Introduction

The assessment of risk of suicide and self harm is one of the most common clinical tasks for mental health professionals. As yet this is largely dependent on unstructured clinical judgement.

Unstructured clinical judgement is inherently difficult to address through training or evidence based changes of practice, it does not lend itself to research or audit and assessments are difficult to justify in a transparent way.

Structured professional judgement addresses these issues and lends itself to clear treatment planning by addressing the risks identified. Although there are well validated tools for the structured professional judgement of risk of violence, such as the HCR-20,¹ as yet there are few such tools for the assessment of the risk of suicide or self harm other than some mnemonics.

The S-RAMM² is the first structured professional judgement tool for this purpose that we are aware of. The handbook for the S-RAMM references published research studies, largely epidemiological in nature, which provide the evidential basis for the inclusion and definition of each item. The evidential integrity and applicability to European and North American populations can be assessed from the handbook and appears satisfactory.

We report the first validation study for the S-RAMM, a clinical tool for structured professional judgement concerning the risk of suicide. Structured professional judgement tools have been widely recognised as 'third generation' risk assessment tools, replacing traditional unstructured clinical judgement as the standard for such work, and increasingly also replacing or enhancing 'second generation' actuarial checklists.

Actuarial checklists claim high predictive accuracy but are constrained by a narrow focus on mainly static historical items, inability to respond to change and inability to include other factors relevant to individual cases, though some have been revised to meet some of these requirements.

The best known of the structured professional judgement tools is the HCR-20,¹ which distinguishes between historical, current clinical and future risk factors for violence. By distinguishing between historical and dynamic factors, the HCR-20 and related tools can be used to identify and prioritise treatment goals and to measure change. The S-RAMM has been drafted to follow the structure of the HCR-20. The Risk Management Authority of Scotland has recently published criteria by which risk assessment tools can be assessed.³ This work was carried out as a first stage in validation according to their criteria.

We set out to assess psychometric properties, inter-rater reliability, concurrent validity and a proxy for criterion validity of the S-RAMM. In part II we will report the results of a prospective study.⁴

Table 1: Cohen's kappa statistic for inter-rater reliability; two researchers, 25 current forensic inpatients

Item	Description	Kappa	Significance (p)
B1	History of deliberate self harm	0.787	0.001
B2	Seriousness of previous attempts	0.768	0.001
B3	Previous hospitalisation	0.712	0.001
B4	Mental disorder	1	Inf
B5	Substance abuse	0.792	0.001
B6	Personality	0.717	0.001
B7	Childhood adversity	0.816	0.001
B8	Suicide in family	0.850	0.000
B9	Age, gender, marital status	0.902	0.000
C1	Suicidal ideation	0.570	0.001
C2	Hopelessness	0.810	0.000
C3	Psychological symptoms	0.634	0.000
C4	Treatment adherence	0.278	0.05
C5	Substance use	1	Inf
C6	Psychiatric admission	0.785	0.001
C7	Psychosocial stress	0.369	0.015
C8	Problem solving deficits	0.691	0.001
F1	Access to preferred means	0.769	0.001
F2	Future service contact	0.567	0.001
F3	Future response to drug tx	0.568	0.001
F4	Future response to psychosocial tx	0.339	0.011
F5	Future stress	0.875	0.001
	Global Risk	0.525	0.001

Methods

Clinical sample

The Central Mental Hospital is the only forensic mental health facility for Ireland, providing high, medium and low security on one campus, with a coherent pathway through care for those admitted from prisons, courts or transferred from civil mental health facilities. At the time of this survey, there were 83 beds in operation, including eight for women, a 12 bed male admission/high secure unit, a six bed high secure unit for persistent challenging behaviour, a 16 bed male medium secure unit, a second 16 bed male medium secure/rehabilitation unit, a 16 bed low secure/rehabilitation unit and a 10 bed pre-discharge unit.

Researchers

Interviews using the S-RAMM were carried out by two post membership psychiatric trainees (equivalent to US fellows). Interviews using the HCR-20 and PANSS⁵ were carried out by two psychology assistants at masters level, working independently of the two S-RAMM assessors. Each pair of assessors was blind to the work of the other. All research interviewers were trained in the use of the research instruments by two consultant psychiatrists, trained as trainers.

Consent

All participants participated as part of routine quarterly and monthly assessments of risk and a range of outcome measures. The research protocol was approved by the research ethics committee.

Study design

We first measured inter-rater reliability, based on joint interviews with patients by two clinicians (post membership psychiatric trainees, equivalent to North American 'fellows'), who rated each subject. Inter-rater reliability refers to the extent of convergence of judgements about individual risk factors and overall risk classifications, of different assessors using the tool on the same patient.

We tested concurrent validity with the HCR-20 because the S-RAMM is designed to have similar form and structure to the HCR-20, and because of the conceptual relationship between violence to others and violence to self, confirmed in some previous studies with the HCR-20.⁶ We examined concurrent validity with the Positive and Negative Symptoms Scale (PANSS) because of the known relationship between risk of suicide and severity of mental state abnormalities.

Sensitivity and discriminatory capacity is a tool's ability to differentiate within each of the risk levels or classifications, between individuals and the levels of risk they present. We have previously reported that the HCR-20 distinguished between forensic patients in admission/high secure, rehabilitation/medium secure and pre-discharge/low secure wards in a forensic hospital.⁷

We therefore used this stratification of patients as a proxy for criterion measures of dynamic risk using the S-RAMM, along with a test for correlation between the HCR-20 scores and S-RAMM scores. We hypothesised that the static historical background scores would correlate poorly but dynamic current clinical and future risk subscales should correlate to some extent.

Statistics

All data were entered in SPSS-12.⁸ The kappa statistic for inter-rater reliability was calculated for pairs of researchers jointly interviewing 25 inpatients.

The same data allowed us to calculate the Cronbach's alpha statistic for internal consistency of the subscales (background, current and future risk factors) and total scores, a measure of content concordance, which is a proxy for content validity.

Results

Inter-rater reliability: joint ratings of patients

Table 1 shows that when two researchers rated the same interviews with 25 current inpatients, the kappa statistic was at an acceptable level for all nine background items.

Two of the eight current risk items had a kappa below 0.5, C4 'treatment adherence' and C7 'psychosocial stress'. One of the five future risk items was below 0.5, F4 'future response to psychosocial stress'. For the overall score, Spearman's rank correlation co-efficient between the two researchers was 0.951 ($p < 0.001$), for the background items subscale, $r = 0.978$ ($p < 0.001$), for the current risk item subscale $r = 0.868$ ($p < 0.001$) and for the future items subscale $r = 0.843$ ($p < 0.001$).

Table 2: Content coherence (Cronbach's alpha statistic), 25 current inpatients

Item	Description	Full scale		Sub scales	
		Researcher A	Researcher B	Researcher A	Researcher B
Overall score	Cronbach's alpha for overall score	0.844	0.829		
Subscale Background Risk	Cronbach's alpha for subscale			0.822	0.807
		Scale score if this item is omitted			
B1	History of deliberate self harm	0.838	0.825	0.782	0.769
B2	Seriousness of previous attempts	0.824	0.814	0.768	0.759
B3	Previous hospitalisation	0.846	0.829	0.795	0.793
B4	Mental disorder	0.836	0.821	0.827	0.812
B5	Substance abuse	0.832	0.822	0.788	0.785
B6	Personality	0.841	0.821	0.814	0.788
B7	Childhood adversity	0.833	0.813	0.790	0.767
B8	Suicide in family	0.840	0.823	0.819	0.788
B9	Age, gender, marital	0.844	0.829	0.834	0.820
Subscale Current Risk	Cronbach's alpha for subscale			0.728	0.666
		Scale score if this item is omitted			
C1	Suicidal ideation	0.845	0.831	0.689	0.656
C2	Hopelessness	0.833	0.812	0.633	0.558
C3	Psychological symptoms	0.838	0.818	0.645	0.514
C4	Treatment adherence	0.848	0.829	0.715	0.685
C5	Substance use	0.846	0.831	0.743	0.680
C6	Psychiatric admission	0.841	0.825	0.757	0.664
C7	Psychosocial stress	0.831	0.820	0.672	0.641
C8	Problem solving deficits	0.838	0.819	0.704	0.612
Subscale Future Risk	Cronbach's alpha for subscale			0.751	0.732
		Scale score if this item is omitted			
F1	Access to preferred means	0.824	0.826	0.793	0.820
F2	Future service contact	0.845	0.825	0.744	0.689
F3	Future response to drug tx	0.842	0.824	0.656	0.659
F4	Future response to psychosocial tx	0.840	0.819	0.665	0.603
F5	Future stress	0.821	0.809	0.658	0.614

Content concordance

Cronbach's alpha statistic measures the extent to which each item fits into the subscale or overall scale to which it is allocated. This is a measure of content coherence – do all items in the overall scale or subscale measure the same thing? If not, then the omission of an item would improve the Cronbach alpha statistic.

Table 2 shows that for the S-RAMM total score, comparing both observers' Cronbach's alpha was 0.844 for one researcher and 0.829 for the other. When omitted, only items C1 'suicidal ideation', and C5 'substance use' led to small improvements in the alpha score for both researchers. No item if omitted led to substantial improvements in the alpha statistic.

Table 3: S-RAMM mean scores (SD): stratified across hospital units.

MMSU1 = Male Medium Secure Unit1, MMSU2 = Male Medium Secure Unit2, MLSU = Male Low Secure Unit. SABU = Selective Adaptive Behaviour Unit

	Number	S-RAMM total	S-RAMM Background	S-RAMM Current	S-RAMM Future	S-RAMM Dynamic (C+F)
Admission unit	11	27.3 (4.9)	11.1 (2.7)	7.6 (2.8)	8.5(1.1)	16.2(2.9)
MMSU1	15	22.8 (4.9)	10.3 (3.5)	5.7(2.2)	7.4(2.7)	13.1(4.0)
MMSU2	15	18.9 (4.9)	7.4 (2.4)	5.8(1.8)	5.9(2.1)	11.5(3.5)
MLSU	15	17.3 (5.0)	8.9 (3.2)	3.7(1.5)	4.7(2.2)	8.4(2.9)
Hostel Ward	10	13.4 (4.7)	7.2 (2.4)	2.9(1.5)	3.3(1.8)	6.2(3.1)
SABU	6	27.7 (5.0)	11.8 (3.9)	6.7(1.2)	9.2(0.9)	15.8(1.9)
Total	81	21.0 (6.6)	9.5 (3.5)	5.3(2.4)	6.3(2.6)	11.6(4.5)

Table 4: HCR-20 means scores (SD): stratified across hospital units

MMSU1 = Male Medium Secure Unit1, MMSU2 = Male Medium Secure Unit2, MLSU = Male Low Secure Unit. SABU = Selective Adaptive Behaviour Unit

	Number	HCR-20 total	HCR-20 Historical items	HCR-20 Clinical Items	HCR-20 Risk items	HCR-20 Dynamic (C+R)
Admission unit	11	28.6 (12.9)	19.1 (10.1)	7.1 (2.2)	3.9 (1.9)	10.6 (3.9)
MMSU1	15	24.0 (8.9)	13.8 (5.7)	6.3 (2.4)	3.9 (2.9)	10.2 (4.6)
MMSU2	15	20.0 (4.4)	13.9(3.2)	3.7 (2.5)	2.5 (2.2)	6.1 (4.1)
MLSU	15	16.1 (4.2)	12.5 (3.2)	2.4 (2.8)	1.2 (2.4)	3.6 (3.7)
Hostel Ward	10	12.7 (6.9)	10.8 (4.6)	1.6 (2.7)	0.3 (0.5)	1.9 (2.9)
SABU	6	30.3 (4.6)	18.3 (1.9)	7.8 (1.7)	4.2 (1.8)	12.0 (3.1)
Total	81	21.1 (9.2)	14.3 (5.8)	4.5 (3.3)	2.5 (2.6)	6.9 (5.2)

Table 2 also shows that for the 'background' subscale, the overall Cronbach's alpha score was 0.822 for one researcher, 0.807 for the other. Items B4 'mental disorder', and B9 'age, gender and marital status' by omission would lead to very small improvements in the overall alpha statistic. For the 'current risk' subscale, the alpha statistic (0.728 and 0.666 for the two researchers) improved marginally for both researchers if item C5 'substance use' was omitted. For the 'future risk' items, (Cronbach's alpha 0.751 and 0.732 respectively for the two researchers) only the omission of item F1 'access to preferred method' led to a consistent but small increase in the alpha statistic.

Concurrent validity with HCR-20

To establish a form of concurrent validity, HCR-20 scores were correlated with S-RAMM scores for 81 current inpatients. S-RAMM total scores correlated moderately with HCR-20 total scores (Spearman $r = 0.480$, $df = 80$, $p < 0.001$). The HCR-20 historical risk items did not correlate with the S-RAMM background items (Spearman $r = 0.108$, $df = 80$, $p = 0.34$).

The combined clinical and future items of each scale (dynamic items) correlated best (Spearman $r = 0.621$, $df = 80$, $p < 0.001$). The component subscales also correlated significantly for matching pairs of current risk and future risk sub-scales (HCR-C -v- S-RAMM C Spearman $r = 0.502$; HCR-R -v- S-RAMM F Spearman $r = 0.442$); and also for crossed pairs of current and future risk subscales (HCR-C

-v- S-RAMM F Spearman = 0.613; HCR-R -v- S-RAMM C Spearman $r = 0.512$).

Concurrent validity with PANSS total score

The PANSS total score correlated poorly with the S-RAMM background subscale (Spearman $r = 0.227$, $p = 0.044$) but correlated better with the S-RAMM future items sub-scale ($r = 0.551$, $p < 0.001$), the S-RAMM current items sub-scale ($r = 0.570$, $p < 0.001$) and correlated best with the S-RAMM dynamic scale (current and future items combined) ($r = 0.610$, $p < 0.001$). The PANSS total score correlated with the S-RAMM total score 0.504 ($p < 0.001$).

Stratification of S-RAMM scores

We have previously shown that HCR-20 scores stratify significantly across the hospital units in keeping with the pathway through care, from admission to pre-discharge. We tested whether this also held true for the S-RAMM. Table 3 shows that the S-RAMM total score differed significantly across units (ANOVA $F = 11.3$, $df = 6$, $p < 0.001$); S-RAMM background subscale ($F = 4.3$, $df = 6$, $p = 0.001$); S-RAMM current risk subscale ($F = 7.6$, $df = 6$, $p < 0.001$); and S-RAMM future risk sub-scale ($F = 10.1$, $df = 6$, $p < 0.001$). The S-RAMM dynamic score ie. current and future combined also stratified significantly (ANOVA $F = 12.7$, $df = 6$, $p < 0.001$).

Table 4 shows a similar pattern of stratification for the HCR-20 scores.

Discussion

Study limitations

Criterion validity (predictive accuracy, predictive power) is the most difficult to assess and can only really be assessed by a prospective study in which suicide or self-harm is the outcome. This paper reports the steps antecedent to such a study. In an accompanying paper, we report the results of a prospective study in the same population.⁴ This paper describes the use of the S-RAMM in a high risk inpatient population. Replication in other populations would be helpful, but high risk populations may be the most appropriate for use of such tools.

Main findings

This is the first validation study for the S-RAMM, and we believe the S-RAMM to be the first structured professional judgement tool for suicide risk assessment.

We have established that the S-RAMM achieves satisfactory levels of convergence, (inter-rater reliability) particularly for the summated scores of subscales and total score.

We have established that the sub-scales of the S-RAMM have internal consistency and the total S-RAMM score also has satisfactory internal consistency. Those items for which inter-rater reliability was relatively low nonetheless contributed significantly to the internal consistency (Cronbach's alpha) of the full scale and subscale scores and if omitted, the Cronbach's alpha score would have deteriorated or would not have improved.

The 'historical' item subscale of the HCR-20 does not correlate significantly with the 'background' or historical subscale of the S-RAMM, but the current and future subscales, collectively measuring dynamic or changeable risk, do correlate significantly between the HCR-20 and S-RAMM.

The correlation of the PANSS total score with the S-RAMM dynamic and S-RAMM total scores confirms a common underlying factor measured by both. This is in keeping with earlier findings confirming a correlation between PANSS scores and HCR-20 scores.⁷ It is worth noting that the PANSS total score includes subscales for positive and negative symptoms but also a 'general' subscale made up of symptoms of depression and anxiety.

Finally, we have shown that the S-RAMM score and dynamic subscale scores stratify as expected along the pathway through care towards recovery in a therapeutically secure hospital.

Future validation studies

This study is a preliminary step in the validation of the S-RAMM. The next step in validating a structured professional judgement tool for assessing risk of suicide would be to carry out a prospective study in which the outcome

is suicide or self-harm. There are obvious ethical considerations in such a study. Those identified as at high risk would have to be offered appropriate interventions. This would mean that any such study would be blunted in its ability to detect an improvement in service over current practice using unstructured clinical judgement. In an accompanying paper we describe such a study.⁴ The definitive study would probably have to screen very large numbers of high-risk patients in order to demonstrate an effect.

Mindful of the criteria suggested by the Risk Management Authority of Scotland for an evidence based risk assessment tool,³ it may be necessary to replicate studies such as this in different populations and cultures.

Advantages of structured professional judgement in clinical practice

A substantial proportion of all suicides are by people who are or have recently been in contact with psychiatric services.⁹ Structured professional judgement has many advantages over unstructured professional judgement.¹⁰ Of greatest practical value, because risk factors are identified as background or dynamic, it facilitates the prioritisation of treatment plans aimed at reducing and managing risk of suicide. For this reason alone, structured professional judgement tools should be used increasingly by clinicians in all areas of practice.

Structured professional judgement is a clearly defined record of the process of risk assessment which is valuable because it allows review of grounds for compulsory detention before statutory mental health tribunals. As a clearly defined process it also lends itself to quality improvement in this common and important area of decision making.

Declaration of Interest: None.

References

1. Webster CD, Douglas KS, Eaves D, Hart SD. HCR-20: Assessing risk for violence, version 2. Burnaby, British Columbia: Simon Fraser University, 1997.
2. Bouch J, Marshall JJ. S-RAMM Suicide Risk Assessment and Management Manual (Research Edition) Cardiff: The Cognitive Centre Foundation, 2003.
3. Risk Management Authority of Scotland. Risk Assessment Tools Evaluation Directory (RATED) version 1. Scotland: Astron. 2006. www.rmascotland.gov.uk
4. Fagan J, Ijaz A, Papaconstantinou A, Lynch A O'Neill H, Kennedy HG. The suicide risk assessment and management manual (S-RAMM) validation study II: prospective study of a structured professional judgement tool for suicide risk assessment. Ir J Psychol Med 2009; 26(3).
5. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schiz Bull 1987; 13: 261-277.
6. Daffern M, Howells K. The prediction of imminent aggression and self-harm in personality disordered patients of a high security hospital using the HCR-20 clinical scale and the dynamic appraisal of situational aggression. Int J Forensic Ment Health 2007; 6: 137-143.
7. Pillay SM, Oliver B, Butler L, Kennedy HG. Risk stratification and the care pathway. Ir J Psychol Med 2008; 25(4): 123-127.
8. Norusis M. SPSS for Windows version 12.0.1 (11 November 2003). Chicago: SPSS Inc. 2004.
9. Appleby L. Safety First: Five-Year Report of the National Confidential Inquiry into Suicide and Homicide by People with Mental Illness. London: Department of Health, 2001.
10. Maden T. Treating violence: a guide to risk management in mental health. Oxford: Oxford University Press, 2007.