ΥΦΥΠΟΥΡΓΕΙΟ ΚΟΙΝΩΝΙΚΗΣ ΠΡΟΝΟΙΑΣ

ΚΥΠΡΙΑΚΗ ΔΗΜΟΚΡΑΤΙΑ

Predicting violence risk: Clinical Judgment v. Structural Professional Judgment v. Actuarial Assessment

Metaxas, G.M.[1], Antoniou, A.[2], & Christodoulou S.C.[1]
[1] Deputy Ministry of Social Welfare
[2] Neapolis University

Introduction

Objective: To assess the predictive validity of two SPJ tools (HCR-20 and SAPROF) as actuarial tools as well as the predictive validity of the clinical judgment methodology.
Participants: N=42 professionals, (14 social workers & 14 psychologists trained in risk assessment tools & 14 social workers untrained using their clinical judgment)

Methodology

Methodology: Retrospective analysis. In particular, 30 vignettes were prepared from record files of persons with violent behaviour, including 15 persons with confirmed risk (murder/serious violence) and another 15 with a confirmed absence of risk (no record of repetition of violent behaviour). Participants, 42 professionals (28 social workers & 14 psychologists), were invited to evaluate each case without prior knowledge of the outcome, either by using the two tools (two Structured Professional Judgment tools (SPJ's), namely the HCR-20 (Historical, Clinical, and Risk Management - 20) and SAPROF (Structural Assessment of Protective Factors), alone (14 social workers & 14 psychologists) or as a team of two professionals (14 pairs of social worker and psychologist) or using their clinical judgment without any tool (14 social workers). Both tools are recommended for risk assessment by NHS in the UK (service specification no. C03/S/a; NHS Commissioning Board, 2013).

A retrospective analysis methodology is employed to investigate the **predictive validity** of the risk assessment in different conditions:
- One professional (i.e. social worker, psychologist) scores the tool.
- A couple of professionals (social worker and psychologist) score the tool.
- One professional assesses the risk using a clinical judgement (no tool).
- Predicting validity is compared in the condition when only risk factors or protective are assessed, and the condition when both tools are scored (Overall Risk and Protection).
- Comparison of the usage of the tools as actuarial and the usage of the tools as SPJ's or risk assessment as clinical judgment.

Analysis:
- t-test in order to investigate the possibility of the two assessment tools or clinical judgment to separate the vignettes into "high risk" and "low risk".
- ROC (Receiver Operating Characteristic) analysis was used to evaluate the predictive validity in various conditions.
- Intraclass Correlation Coefficient was used to investigate Inter-Rater Reliability in various conditions.

Research was approved by the National Bioethics Committee (ΕΕΒΚ ΕΠ 2018.01.84).

Results

28 evaluators * 15 cases *N* = 840 evaluations for HCR & SAPROF / 14 evaluators * 15 cases *N* = 420 evaluations for group evaluation & clinical judgment.

Cronbach α	
SAPROF 28 One Professional	α=.86
SAPROF 14 Two Professionals	α=.86
HCR-20 28 One Professional	α=.88
HCR-20 14 Two Professionals	α=.90

Internal consistency for both tools is considered high and the results agree with the relevant results in the literature. Improvement of the specific index (Cronbach α) is indicated for both tools in the condition "team of two professionals".

The **correlation** of SAPROF with HCR-20 is statistically significant at the *p*<.001 level and takes the value ***r*=-.52** and ***r*=-.64** when scoring was done by one or two professionals respectively.

t-test

HCR-20 (Individual Assessment): statistically ***significant*** difference between low risk (*M*=20.51, *SD*=9.43) & high risk (*M*=26.43, *SD*=7.91), *t*(838)=-10.48, *p*= 0.00, ***Cohen's d* = 0.68**
HCR-20 (Couples Assessment): statistically ***significant*** difference between low risk (*M*=17.50, *SD*=8.50) & high risk (*M*=28.57, *SD*=6.49), *t(418)*=-15.08, *p*=0.00, ***Cohen's d* = 1.47**
SAPROF (Individual Assessment): statistically ***significant*** difference between low risk (*M*=14.48, *SD*=6.15) & high risk (*M*=6.48, *SD*=4.44), *t*(838)=21.61, *p*= 0.00, ***Cohen's d* = 1.49**
SAPROF (Couples Assessment): statistically ***significant*** difference between low risk (*M*=14.97, *SD*=6.47) & high risk (*M*=6.58, *SD*=4.18), *t(418)*=15.78, *p*=0.00, ***Cohen's d* = 1.54**
Clinical Judgment Risk: statistically ***significant*** difference between low risk (*M*=3.53, *SD*=1.22) & high risk (*M*=4.35, *SD*=0.98), *t*(418)=-7.62, *p*= 0.00, ***Cohen's d* = 0.74**
Clinical Judgment Protection: statistically ***non significant*** difference between low risk (*M*=2.75, *SD*=0.98) & high risk (*M*=1.77, *SD*=0.97), *t*(418)=10.32, *p*>.05
Clinical Judgment indicates either no statistical differentiation between "high risk" and "low risk" vignettes (protection) either low effect size (Cohen's *d*). Improvement of the effect size is indicated for Risk Assessment (HCR-20) in the condition "team of two professionals". "High Risk" and "Low Risk" vignettes are differentiated with a high effect size for both tools.

Inter-rater reliability			
Assessment Type	**Intraclass Correlation Coefficient**	***CI95%***	***p***
HCR-20 (Individual Assessment)	.86	.85-.87	*p*<.001
HCR-20 (Couples Assessment)	.90	.89-.92	*p*<.001
SAPROF (Individual Assessment)	.83	.81-.85	*p*<.001
SAPROF (Couples Assessment)	.86	.84-.88	*p*<.001
Clinical Judgment	.41	.33-.49	*p*>.001

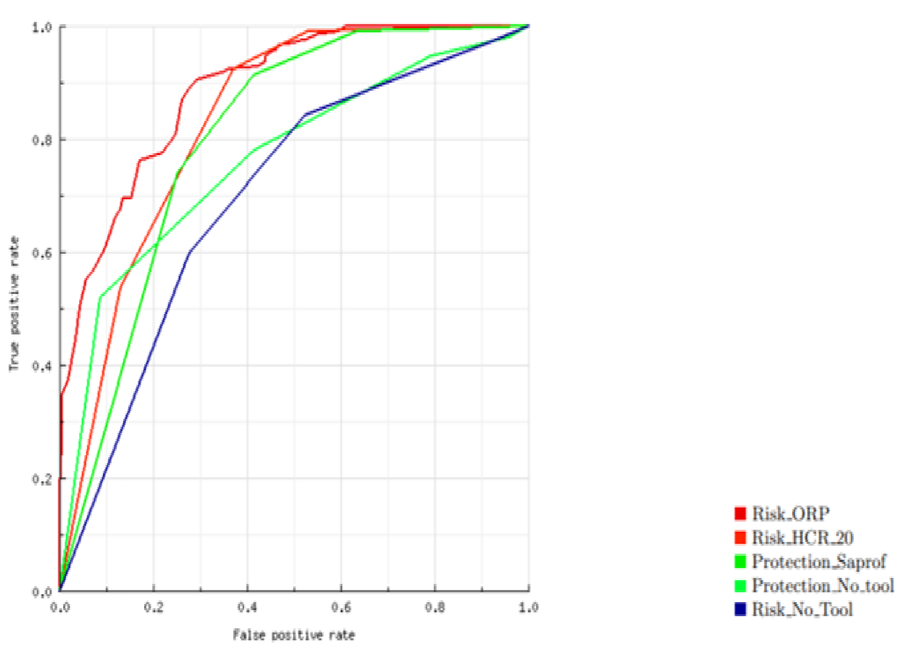
The results indicate that both tools are reliable and valid and both teams of professionals are able to use SPJ's for risk assessment in a community sample.

The condition "multi-professional group" improves all psychometric properties of the risk assessment.

Clinical Judgment indicates poor psychometric properties and low inter-rater reliability.

	AUC	CI95%	Tp	fp
HCR-20				
14 Social Workers	.68	.63-.73	52%	12%
14 Psychologists	.72	.67-.77	50%	10%
14 pairs of SW & Psy	.84**	.81-.88	64%	7%
SAPROF				
14 Social Workers	.84	.81-.88	96%	43%
14 Psychologists	.84	.81-.88	84%	33%
14 pairs of SW & Psy	.85	.82-.89	92%	41%
ORP (Actuarial)				
14 Social Workers	.79	.75-.83	58%	13%
14 Psychologists	.81	.77-.85	49%	2%
14 pairs of SW & Psy	**.89****	.86-.92	89%	27%
Final Protection (SPJ)				
14 Social Workers	.80	.76-.84	59%	9%
14 Psychologists	.82	.78-.87	69%	10%
14 pairs of SW & Psy	.81	.76-.85	64%	10%
Final Risk (SPJ)				
14 Social Workers	.83	.79-.87	92%	37%
14 Psychologists	.85	.81-.88	90%	40%
14 pairs of SW & Psy	.86	.82-.89	93%	38%
Risk - Clinical Judgment				
14 Social Workers	.70	.65-.75	60%	27%
Protection - Clinical Judgment				
14 Social Workers	.76	.72-.81	51%	9%

1 ROC Plots



Note: **AUC**: Area Under the Curve, CI95%: Confidence Interval 95%, **tp**: true positive, **fp**: false positive, **HCR-20**: total score of HCR-20, **SAPROF**: total score of SAPROF, **ORP**: Overall Risk Protection: Total Score of HCR-20 minus total of SAPROF, **Final Protection (SPJ)**: Structural Professional Judgment of protective factors, **Final Risk (SPJ)**: Structural Professional Judgment of risk factors / **: p<.001.

Discussion

ROC (Receiver Operating Characteristic) analysis was used to evaluate the predictive validity of the tools or that of the condition of clinical judgment. Results showed that there is a statistically significant AUC (Area Under the Curve) in all cases using a tool (social worker/psychologist, multidisciplinary group) indicating the tools' reliability in predicting the possibility of a person re-emerging violent behaviour in the future. Statistical analysis and comparison of AUC values suggest a **significant improvement in predictive validity for the condition "multidisciplinary group" – "Overall Risk Protection", i.e., the combined usage of both tools as actuarial-tools, over all other conditions ("only HCR-20", "psychologist", "social worker")**. **Clinical Judgment indicates poor psychometric properties and low inter-rater reliability.**

Methodology of the research has certain **limitations** and namely, since 42 professionals assessed the same 30 case vignettes, **results are embodied and there is a possibility of technically rising the psychometric properties of the tools**. In future research a longitudinal methodological design would be a better solution for the assessment of the psychometric properties of the tools.