Should we Always Look on the Bright Side of Life?

The Role of Protective Factors in Forensic Risk Assessment and Management

Krista Ann Jansen, BA (Hons)

Thesis submitted to the University of Nottingham for the degree of Doctorate in Forensic Psychology (D.Foren.Psy)

January 2017

Abstract

This thesis will aim to identify the role of protective factors in clinical risk prediction and management of inpatients in secure services. Chapter one provides an introduction to the field of forensic risk assessment, and a justification for the importance of including protective factors to create a more balanced assessment of patient risk.

Chapter two presents a systematic review exploring the role of protective factors in various risk assessment tools. The aims of this review is to identify risk assessment tools that incorporate a measure of protective strengths, and the predictive validity of these tools for the outcome of inpatient aggression or violence.

Chapter three provides a critical review of the START, specifically focused on its strengths and weaknesses as a risk prediction tool, demonstrating good validity, reliability, in addition to excellent inter-rater reliability. Findings will be reported in terms of the clinical utility of this tool.

Chapter four presents the findings from a research project that aimed to identify the predictive validity of the START in female forensic inpatients, highlighting the efficacy of the specific risk estimates.

Chapter five utilises a case study to identify how promoting protective factors in the form of coping skills can reduce self-harming behaviours. This case study aims to demonstrate that

focusing treatment on promoting patient strengths can be effective in reducing adverse outcomes.

Finally, chapter six discusses the overall findings and clinical implications from the present thesis. Future research directions will also be discussed.

Preface

This thesis aims to provide an overview of the role of protective factors in forensic risk assessment and management. There are a number of fourth generation risk assessment tools employing the benefits of protective factors or measures of an individual's strengths in order to create a more balance assessment of an individual's risk. Each chapter is designed to provide unique evidence to support the clinical utility of incorporating protective factors into forensic risk assessment and management. An introduction to the thesis is provided, justifying the focus and overall aims.

A systematic review explores the literature surrounding existing risk assessment tools that include a measure of individual strengths. A review of this kind has not been conducted in recent years, and thus provides an update on the status of these tools and their predictive validity with regards aggressive or violent outcomes.

The Short-Term Assessment of Risk and Treatability (START) is a widely used risk assessment tool that uniquely requires raters to rate each item as both a strength and a vulnerability. This tool is critically evaluated to establish the reliability and validity of this method, as well as the clinical utility of the strength scale, vulnerability scale, and specific risk estimates.

An empirical research project investigates the predictive validity of the START. Much of the current research on the predictive validity of the START focuses on males, thus this research

project investigates whether the same conclusions of the effective predictive validity of the START can be applied to a female sample.

Whilst it may be clinically useful to assess an individual's strengths, it is important to consider how this informs risk management and treatment approaches. A case study design is utilised to demonstrate how promoting coping skills can reduce the number of self-harm incidents in a patient with borderline personality disorder. Thus demonstrating the effectiveness of reducing risk outcomes by promoting protective factors.

A final discussion draws together the overall findings and these are considered against the aims of the thesis. In light of the findings, clinical implications and future research directions are also commented upon.

Statement of Authorship and Publication

Chapter Two

This systematic review has been submitted for publication in the peer reviewed Aggression and Violent Behavior. The authorship of this article indicates collaborative working. I am the primary author of this article and Professor Vincent Egan (The University of Nottingham) is named as co-author.

Chapter Three

This critique has been submitted for publication in the peer reviewed Journal of Forensic Psychology Practice. The authorship of this article indicates collaborative working. I am the primary author of this article and Professor Vincent Egan (The University of Nottingham) is named as co-author.

Chapter Four

This research project has been submitted for publication in the peer reviewed International Journal of Forensic Mental Health. The authorship of this article indicates collaborative working. I am the primary author of this article; Paul Mooney (Partnerships in Care) and Professor Vincent Egan (The University of Nottingham) are named as co-authors.

The results of this research project were presented as a poster (Appendix I) at the following professional conferences:

Jansen, K., Mooney, P., & Egan, V. (2016, May). *Risk Incidents in Females in Medium Secure Forensic Mental Health Services: Predictive Validity of the Short-Term Assessment of Risk and Treatability (START).* CLAHRC East Midlands Institute of Mental Health Conference, Nottingham, UK.

Jansen, K., Mooney, P., & Egan, V. (2016, June). *Risk Incidents in Females in Medium Secure Forensic Mental Health Services: Predictive Validity of the Short-Term Assessment of Risk and Treatability (START).* Division of Forensic Psychology Conference, Brighton, UK.

Chapter Five

This case study has been submitted for publication in the peer reviewed Journal of Forensic Psychology Practice. The authorship indicates collaborative working. I am the primary author of this article and Professor Vincent Egan (The University of Nottingham) is named as co-author.

Acknowledgements

I would like to thank my supervisor, Professor Vincent Egan, for his continuous support and guidance throughout this process. Your knowledge and expertise have been invaluable and very much appreciated. Thank you.

Special thanks go to my clinical supervisor Paul Mooney, who gave me the opportunity to conduct this research and the support to keep going when I wanted to give up. Also to Ian McGraw for tirelessly extracting my data from the system and making my research possible.

To my colleagues along the way, your patience and support has not gone unnoticed. During stressful days in the office and numerous tearful research moments, I thank you for always putting things in perspective and keeping me grounded.

Finally to my friends and family, whose continuous support and encouragement helped to keep me smiling, even when drowning in statistics! I could not have done this without you. A special thank you to my Mum and Dad who have supported me both emotionally and financially throughout this doctorate; I owe you the world. And I promise to start paying back my debt!

Table of Contents

Abstract..... **i**

Preface..... **ii**

Statement of Authorship and Publication..... **iv**

Acknowledgements..... **vi**

List of Tables..... **xiii**

List of Figures..... **xiv**

List of Appendices..... **xv**

Rationale for Chapter One..... **1**

Chapter One: Introduction..... **2**

1.1 Forensic Risk Assessment Tools..... 3

1.2 The Role of Protective Factors..... 5

1.3 Aims of the Current Thesis..... 7

Rationale for Chapter Two..... **8**

Chapter Two: A Systematic Review of the role of Protective Factors within Risk Assessment of Inpatient Aggression or Violence: Examination of the SAPROF, SAVRY, and START..... **9**

2.0 Abstract..... 10

2.1 Introduction..... 11

2.1.1 Actuarial..... 11

2.1.2 Structured Professional Judgement..... 12

2.1.3 The Role of Protective Factors..... 12

2.1.4 The Structured Assessment of Violence Risk

in Youth (SAVRY)..... 13
2.1.5 The Inventory of Offender Risks, Needs, and Strengths (IORNS)..... 14
2.1.6 The Structured Assessment of Protective Factors for Violence Risk (SAPROF)..... 14
2.1.7 The Short-Term Assessment of Risk and Treatability (START) 15
2.1.8 Aims and Objectives..... 16
2.2 Method..... 16
2.2.1 Search Strategy..... 16
2.2.2 Study Selection..... 17
2.2.3 Quality Assessment..... 19
2.2.4 Data Extraction..... 20
2.3 Results..... 20
2.3.1 Quality Assessment of Included Studies..... 22
2.3.2 Characteristics of Included Studies..... 23
2.3.3 Risk Assessments Examined in Included Studies..... 29
2.3.4 Predictive Validity..... 30
2.3.4.1 SAPROF..... 37
2.3.4.2 SAVRY..... 37
2.3.4.3 START..... 37
2.4 Discussion..... 40
2.4.1 Limitations and Future Considerations..... 44
2.4.2 Conclusion..... 47

Rationale for Chapter Three..... **48**

Chapter Three: The Short-Term Assessment of Risk and Treatability: A Critique..... **49**

3.0 Abstract..... 50

3.1 Introduction..... 51

3.1.1 Overview of the START..... 53

3.1.1.1 Development of the START.... 54

3.1.1.2 Structure..... 54

3.2 Reliability..... 56

3.2.1 Internal Consistency..... 56

3.2.2 Correlations Between Scales..... 56

3.2.3 Inter-Rater Reliability..... 58

3.3 Validity..... 59

3.3.1 Construct Validity..... 59

3.3.2 Convergent Validity..... 59

3.3.3 Incremental Validity..... 60

3.3.4 Predictive Validity..... 60

3.4 Conceptual Issues and Clinical Application..... 62

3.5 Discussion..... 66

3.5.1 Conclusion..... 67

Rationale for Chapter Four..... **69**

Chapter Four: Risk incidents by females in medium-secure forensic Mental Health Services: Predictive validity of the Short-Term Assessment of Risk and Treatability (START)..... **70**

4.0 Abstract..... 71

4.1 Introduction..... 72
4.1.1 The Short-Term Assessment of Risk and Treatability..... 75
4.1.2 Aims of Current Study..... 77
4.2 Method..... 79
4.2.1 Ethics..... 79
4.2.2 Participants..... 79
4.2.3 Procedure..... 80
4.2.3.1 START assessment..... 80
4.2.3.2 Demographic and clinical information..... 82
4.2.3.3 Risk incidents..... 83
4.2.4 Data Analysis..... 84
4.3 Results..... 86
4.3.1 Inter-Rater Reliability..... 86
4.3.2 Participants..... 86
4.3.3 START Scores and SREs..... 87
4.3.4 Incident Data..... 88
4.3.5 Predictive Validity..... 89
4.3.6 START Scores and Number of Risk Incidents. 89
4.3.7 SREs and Number of Risk Incidents..... 92
4.3.8 Incremental Validity..... 93
4.4 Discussion..... 100
4.4.1 Limitations and Future Considerations..... 103
4.4.2 Conclusion..... 107
Rationale for Chapter Five..... **108**

Chapter Five: Reducing Risk Behaviours by Promoting Protective Factors: A Case Study on the Effectiveness of a Coping Skills programme for a Female with Borderline Personality Disorder and Self-Harming Behaviours..... **109**

5.0 Abstract..... 110

5.1 Introduction..... 111

5.1.1 Client Overview..... 112

5.1.2 Assessment and Analysis..... 113

5.1.3 Formulation..... 118

5.1.4 Case Conceptualisation..... 122

5.2 Intervention..... 124

5.3 Results..... 127

5.4 Discussion..... 139

Rationale for Chapter Six..... **145**

Chapter Six: Discussion..... **146**

6.1 Aim 1..... 147

6.2 Aim 2..... 147

6.3 Aim 3..... 148

6.4 Aim 4..... 149

6.5 Clinical Implications..... 150

6.6. Future Directions..... 151

6.7 Conclusion..... 152

References..... **154**

Appendices..... **217**

List of Tables

Chapter Two

2.1 Electronic Database Search

2.2 Included Studies

2.3 AUC Statistics for Risk Assessments

Chapter Three

3.1 Relationships between START Total Scores and Risk Estimates

3.2 Satisfaction of Clinicians using START in Practice

Chapter Four

4.1 Specific Risk Estimate frequencies by START rating

4.2 Predictive Validity of START scores and SREs

4.3 Spearman's rho correlations between START scores and number of Risk Incidents

4.4 One-way ANOVA for significance of SRE rating and number of Risk Incidents

4.5 Block-entry Linear Regression Analyses

Chapter Five

5.1 START Specific Risk Estimates

5.2 The 5 P's of CBT Formulation

5.3 Treatment Aims

5.4 Session Details

5.5 Coping Style Questionnaire Scores

List of Figures

Chapter Two

2.1 Systematic Review Search Strategy

Chapter Five

5.1 Case Formulation

5.2 START Strength scores

5.3 START Vulnerability scores

List of Appendices

Chapter Two

A Search Syntax

B Inclusion Form

C Quality Assessment Form

D Data Extraction Tool

E Excluded Studies

Chapter Four

F Ethics Approval Letter

G Invitation to Take Part and Research Proposal

Chapter Five

H Consent Form

I Risk Incidents in Females in Medium Secure Forensic Mental Health Services: Predictive Validity of the Short-Term Assessment of Risk and Treatability (START) Poster

Rationale for Chapter One

Chapter one provides an overview of the current approaches and tools in forensic risk assessment. The components of current risk assessments will also be introduced, as well as a beginning commentary on the potential role of protective factors in the assessment and management of risk. Finally, the overall aims of the thesis will be outlined.

Chapter One

Introduction

The ability to accurately predict and manage individual risk is a fundamental element of forensic mental health care. This does not just pertain to violence risk, but several other areas including self-harm, self-neglect, substance misuse and victimisation to name a few. In order to provide both effective and appropriate care within secure mental health settings, clinicians often rely on a risk assessment tool to reach an overall risk rating. During this process, it is vital that clinicians consider both the overall needs of the individual, and the risk to either themselves or others (Department of Health; DoH, 2009). Risk assessment and management should also be collaborative between disciplines, and adopt a positive approach to help to acknowledge the strengths of an individual (DoH, 2009). The specific process of risk assessment has developed over the years, from the types of variables considered predictive of risk, to the approach and tools utilised by professionals.

1.1 Forensic Risk Assessment Tools

Over recent years, a variety of risk assessment tools have been developed to aid clinicians in predicting the frequency, immanency, and likelihood of patients exhibiting adverse behaviours. Traditionally risk assessment tools focused on static factors. These are historical factors that cannot be changed (DoH, 2009). This can create a bleak view of risk, as an individual's risk will never change if only historical factors are considered. However, we cannot discount the value of static factors, as these can give a valuable insight into the likelihood of an individual engaging in an adverse outcome. Therefore, whilst considering these historical

factors, there needed to be some way of acknowledging an individual's current level of risk. This change in approach to risk assessment started to see the inclusion of dynamic risk factors. These are factors that can change over time, and thus reflect the relative change in an individual's risk (DoH, 2009). The inclusion of these dynamic factors also helped to shape the current forensic risk assessment tools to not only predict risk, but also make steps toward managing risk and highlighting areas for intervention (DoH, 2009).

A recent survey conducted by Singh, Fazel, Gueorguieva, and Buchanan (2014) found that the most widely used risk assessments within the forensic field are the Historical, Clinical, and Risk-20 Version 3 (HCR-20 (v3); Douglas, Hart, Webster, & Belfrage, 2013) and the Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp, & Webster, 1997). These structured professional judgement tools utilise both static and dynamic factors in a combination of clinical judgement and actuarial methods, culminating in a final clinical judgement of the overall risk of the individual (Doyle & Dolan, 2002). This approach allows clinicians to create a risk-based formulation that can be used to develop appropriate risk management strategies. Whilst these tools are the most widely used within forensic populations, they do not incorporate a measure of protective strength. There is an argument that without the inclusion of potential risk-reducing factors, a risk-only assessment will inherently create an overestimation of an individual's risk, which

may impact on their treatment goals and result in a more punitive approach to the management of the individual (Rogers, 2000).

1.2 The Role of Protective Factors

The inclusion of protective factors within forensic risk assessment is a fairly new process, with few risk assessment tools incorporating an assessment of protective strengths (de Vries Robbe, de Vogel, & Stam, 2012). These include the Inventory of Offender Risks, Needs, and strengths (IORNS; Miller, 2006) the Structured Assessment of Protective Factors (SAPROF; de Vogel, de Ruiter, Bouman & de Vries Robbé, 2009), the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel & Forth, 2002), and the Short-Term Assessment of Risk and Treatability (START; Webster et al., 2009). These structured professional judgement tools incorporate a measure of protective factors, either within each item or incorporate a separate measure or scale of protective strengths. This subsequently means that the final clinical judgement of an overall rating of risk includes the consideration of protective strengths, and potential risk-reducing factors.

Incorporating protective factors into forensic practice is already evident in current treatment models. Ward and colleagues (Ward & Brown, 2004; Ward, Mann & Gannon, 2007) developed a strength-based treatment approach called the Good Lives Model. Ward proposed that the best way to reduce undesired outcomes, such as recidivism, is to provide individuals with the tools they need to live satisfying and rewarding lives, rather than continuing to

develop risk management assessments and strategies (Ward & Brown, 2004).

Whilst forensic services are obligated to do risk assessment, it may be that the shift of focus of these assessments needs to be towards a more balanced assessment of both risk and protective factors in order to create a more balanced view of an individual, this may provide insight into the protective factors that may tie in to the Good Lives model approach to offender rehabilitation. This is not a new idea, and recent literature has started to look at the effectiveness of combining traditional risk assessment measures with a measure of protective strength. For example de Vries Robbé, de Vogel and Douglas (2013) looked at the predictive validity of the HCR-20 combined with the SAPROF (which includes a measure of protective strengths). They found that adding a measure of patient strengths increased the predictive validity of the HCR-20 (de Vries et al., 2013).

Though the inclusion of protective factors does not replace risk factors, both components should be incorporated into the risk assessment process to create the most balanced assessment of patient risk. De Ruiter and Nicholls (2011) argue that to provide the most comprehensive assessment of patient risk, clinicians should acknowledge static, historical risk factors in order to establish a base rate of risk for future behaviour; and dynamic risk factors in order to diminish risk, and individual strengths and protective factors to increase engagement in the therapeutic process and

subsequently reduce patient risk by emphasising the positive aspects of the individuals life.

1.3 Aims of the Current Thesis

In recognition of the growing interest in incorporating protective factors into the process of forensic risk assessment and management, the current thesis aims to:

1. Identify current risk assessment tools that incorporate protective factors and their effectiveness at predicting risk behaviours within secure mental health services;
2. To critically evaluate the structure of such a tool;
3. To understand the specific role of a measure of protective strengths within a risk assessment tool;
4. And investigate if promoting said protective factors can reduce risk behaviours in an individual.

Rationale for Chapter Two

Chapter two introduces the different approaches of risk assessment, and focuses on those tools that incorporate a measure of strengths or protective characteristics. A systematic approach was adopted to search the existing literature in order to gain an insight on the current knowledge on the effectiveness of incorporating protective factors in the risk prediction of aggressive or violent incidences within secure mental health services.

Chapter Two

A Systematic Review of the role of Protective Factors within Risk Assessment of Inpatient Aggression or Violence: Examination of the SAPROF, SAVRY, and START

2.0 Abstract

Purpose: This study aimed to systematically review the role of protective factors within violence risk assessment tools in predicting inpatient aggression and violence.

Background: There are several actuarial and structured professional judgement tools currently used within secure services, however few of them include a measure of protective factors that may reduce the likelihood of an individual engaging in aggressive behaviour. This review aimed to build upon the existing literature and evaluate the overall quality of the research in this area.

Method: Six electronic databases were searched, including a manual search of relevant reference lists. Four experts were also contacted to obtain further relevant studies. Inclusion criteria were predefined in reference to study design, population, exposure, and outcome measures. Relevant studies were quality assessed prior to data extraction.

Results: Initial search yielded a result of 1488 publications. 319 duplicate publications were removed and a further 847 irrelevant publications were excluded. 305 publications were excluded for not meeting the initial inclusion criteria, quality assessment threshold, or were unobtainable. A final total of 17 studies were included.

Conclusions: The findings presented in this review support the inclusion of protective factors within risk assessment for inpatient aggression.

2.1 Introduction

Risk assessment is at the forefront of forensic services' priorities and responsibilities. The type, structure, and focus of risk assessment tools have changed dramatically over the years. This is reflective of the increase in knowledge of risk and the most effective way to predict risk in a variety of different areas relevant to the clinical management and treatment of individuals.

2.1.1 Actuarial

Actuarial risk assessments conceptualise risk in terms of probability of a future adverse event (Hart, Michie, & Cooke, 2007). This approach ignores other facets of risk including nature, severity, imminence, and duration (Hart, 2001, 2003). By utilizing fixed algorithms and predictor variables derived from the existing literature of known recidivistic and non-recidivistic groups, a probability is computed that an individual will engage in a specified adverse outcome (Hart, Michie, & Cooke, 2007). These subsequently produce a quantifiable outcome that can be compared across norms and within populations.

Whilst this mathematical approach to risk assessments has its advantages in that it reduces rater bias, there are also flaws. One of the major limitations of using a score to predict future violence is the 'margin of error' (Hart, Michie, & Cooke, 2007). As the probabilities derived from actuarial risk assessments are based on group data, these 95% confidence intervals are only reflective of a group, not an individual case (Hájek & Hall, 2002). When looking at individual cases, the margin of error for an individual committing or

not committing a violent act will not be the same as a group of violent individuals (Hájek & Hall, 2002). Therefore actuarial risk assessments have the potential to be an underestimation of individual risk.

2.1.2 Structured Professional Judgement

With the use of structured professional judgement becoming ever more popular amongst clinicians, both the quantity and quality of these tools have increased over recent years, and are commonly used across secure forensic mental health services (Khiroya, Weaver & Made, 2009). These risk assessment tools utilize the clinical relevance of professional judgement, whilst providing an appropriate structure on which to rate the applicability of the variables identified as predictive of a certain behaviour or outcome.

Unlike actuarial assessments, this procedure allows the use of dynamic risk factors that are likely to change throughout intervention (Otto & Douglas, 2011). This means risk can be reassessed more frequently and will theoretically reflect the progress the individual is making throughout intervention. This approach allows a more tailored risk assessment process, informed by both the structure of an actuarial method, with the freedom to individualize the assessment process through professional judgement.

2.1.3 The Role of Protective Factors

A recent shift in the focus of risk assessment tools has been to acknowledge the potential importance of protective factors in assessing and managing risk. Traditional risk assessments focus on

how discrete items increase the likelihood of an individual engaging in a risky behaviour. One-sided risk assessments focussing on problem behaviours create a biased negative view of individuals, as the protective factors that may reduce the likelihood of the risk behaviours are overlooked (Braithwaite, Charette, Crocker, & Reyes, 2010). This is particularly evident in forensic instruments, as very few incorporate protective factors into the assessment process. Instruments that include both risk and protective factors potentially create a more balanced view, aiding in risk management whilst acknowledging the progress an individual is making to reduce these risk behaviours.

Protective factors have been proposed to reduce the likelihood of an adverse outcome, either by reducing the impact of risk factors, or having an independent influence on the adverse outcome (Rogers, 2000). By ignoring protective factors, a balanced evaluation of risk assessment cannot take place. In response to this, several risk assessments now incorporate protective factors in their evaluation of the likelihood of an individual engaging in aggressive or violent behaviours.

2.1.4 The Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel & Forth, 2002)

The first violence risk assessment tool to incorporate protective factors was the SAVRY. This tool is designed for use with individuals aged 12-18 years, and was developed from the growing literature surrounding reducing the risk of juvenile delinquency. Alongside the 24 risk items, the SAVRY includes six protective

factors that are rated as either present or absent (Borum et al., 2002). These factors are prosocial involvement, strong social support, strong attachment and bonds, positive attitudes toward intervention and authority, strong commitment to school, and resilient personality traits (Borum et al., 2002).

2.1.5 The Inventory of Offender Risks, Needs, and Strengths (IORNS; Miller 2006)

The IORNS was the first violence risk assessment tool for adults to include protective factors. The IORNS is a 130-item self-report measure that assesses static and dynamic risk. Whilst the IORNS is primarily focused on the risks of an individual, this tool includes a 26-item protective strength index. This index assesses an individual's cognitive and behavioural regulation, anger regulation, and education and training (Miller, 2006). These protective factors are then used to provide an estimation of recidivism, as well as identifying treatment need and informing risk management strategies (Miller, 2006).

2.1.6 The Structured Assessment of Protective Factors for Violence Risk (SAPROF; de Vogel, de Ruiter, Bouman & de Vries Robbé, 2009)

The authors of the SAPROF identified the gap in the need for an exclusively focussed assessment of protective factors. This tool does not incorporate any risk factors, and was originally designed as an addition to existing risk focused SPJ tools. The SAPROF includes 17 protective factors rated on a three-point scale that produces an overall judgement of level of available protection in the individual

(de Vogel et al., 2009). Raters are also required to highlight "key items" and "goal items" for the individual to be used to inform risk management and treatment planning (de Vogel et al., 2009).

2.1.7 The Short-Term Assessment of Risk and Treatability (START; Webster et al., 2009)

Whilst the aforementioned risk assessments comprise separate risk and protective factors scales, the START offers a unique view of risk of violence by considering each item as both a risk ("vulnerability") and a protective factor ("strength"). The 20 items on the START are dynamic factors and rated on a three-point scale (0: *Minimally Present,* 1: *Moderately Present,* and 2: *Maximally Present*). Raters are also required to provide a specific risk estimate of either low, medium, or high for the seven risk domains of the START, which are violence risk to others, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, and victimisation (Webster et al., 2009). The summary risk estimate should be a consideration of all 20-item of the START and reflect both the vulnerabilities and strengths of the individual.

The START is the first risk assessment to rate risk and protective factors for the same items of the scale, rather than having additional or separate protective indices. It could be argued that this approach embeds the identification of protective factors into the process of risk assessment, creating a more balanced view of the individual.

2.1.8 Aims and Objectives

With an increasing interest in incorporating protective factors into forensic risk assessment, this review aimed to further the knowledge of the role of protective factors in violence risk assessment, examining the efficacy of these risk assessment tools in relation to the prediction of violent or aggressive outcomes in inpatient secure mental health settings. This review employed rigorous quality assessment, helping to inform mental health practitioners select an appropriate risk assessment tool that incorporates protective factors to create a more balanced assessment of inpatient risk of aggressive or violent behaviour.

2.2 Method

2.2.1 Search Strategy

The PRISMA guidelines (Moher et al., 2009) were followed when carrying out this systematic review. The search was conducted across six databases (see Table 2.1). The publishing dates were not predefined to ensure all literature published on protective factors in risk assessment were captured in the search. A hand search of reference lists of relevant papers was also carried out. In order to obtain further publications, four experts in the field were contacted. However, none responded within the given two-month timeframe to be included in the current review.

Table 2.1

Electronic Database Search

Database	**Date Range**	**Date Completed**
OVID:		
PsycINFO	1806 – June Week 5 2016	12th July 2016
MEDLINE	1946 – June Week 5 2016	12th July 2016
EMBASE	1980 – 2016 Week 27	12th July 2016
Web of Science	1900 – 2016	12th July 2016
PROQUEST:		
ASSIA	1987 – 2016	12th July 2016
Cochrane Library	1997 – 2016	12th July 2016

The following is a simplified version of the search terms used across all electronic databases. A full search syntax is provided in Appendix A.

(violence/aggression/hostility)

AND

(risk assessment/names of known violence risk assessment tools with protective factors)

AND

(prediction/area under curve/validity)

2.2.2 Study Selection

An inclusion criteria tool was created to assess eligibility for inclusion in the current systematic review (see Appendix B). Criteria were defined by study design, population, exposure, and outcomes.

If all criteria were met, the publication was included in the quality assessment process. Opinion or commentary papers were excluded from the review as these are not empirical in nature of their evaluation.

The population was restricted to inpatients within a mental health setting. This is because being able to accurately predict inpatient aggression is at the forefront of mental health services. With inpatient aggression producing a high emotional and financial cost for staff, patients, and institutions (Duffy, Doyle, & Ryan, 2004), a valid and reliable method of predicting aggression is highly sought after. From a practical perspective, this data is routinely collected in mental health settings, and thus the majority of the research utilises this setting. Similarly, these tools are predominantly designed for use within mental health settings, so from a theoretical perspective, these tools are designed for this population. Therefore it was deemed clinically useful to limit this review to secure mental health services.

Only violence risk assessment tools that incorporate some measure of protective factors were included in this review. Recent literature suggests that the predictive validity of other violence risk assessments such as the HCR-20 (Douglas, Hart, Webster, & Belfrage, 2013) is increased with the addition of a measure of protective factors (de Vries Robbé, de Vogel, & de Spa, 2011). The current systematic review aimed to investigate the predictive validity of tools that incorporate a protective factor index or scale.

Aggressive and violent incidents were chosen as the outcome measures. These were defined in the literature as inclusive of both physical and verbal aggression. No limit was put on the target of the violence, i.e. towards others or towards objects.

There was no restriction on publication language set so as to avoid selection bias. Both English and American spellings were included in the search terms.

2.2.3 Quality Assessment

A quality assessment tool was created based on the Critical Appraisal Skills Programme (CASP; 2006) and the QUADAS (Whiting et al., 2003; see Appendix C). This tool was applied to all publications that met the initial inclusion/exclusion criteria. The following threshold criteria were applied during the quality assessment:

- Clear identification of protective factors
- Clear identification of aggressive or violent behaviours
- Use of Area Under Curve (AUC) analysis

Publications that did not meet these criteria were excluded for not meeting the quality threshold.

In order to enhance accuracy of quality assessment, an independent second rater (forensic psychologist in training) was recruited to quality assess 20% of the studies that met the initial inclusion criteria. Inter-rater reliability was calculated.

2.2.4 Data Extraction

A data extraction tool was created to extract the data from the studies that reached the quality assessment threshold prior to synthesis (see Appendix D).

2.3 Results

The full search strategy is displayed in Figure 2.1. The electronic database search yielded a total of 1488 publications. Three hundred and nineteen duplicate publications were removed and a further 847 irrelevant publications were excluded. Three hundred publications were excluded for not meeting the initial inclusion criteria (see Appendix E for reasons for exclusion). Four publications did not meet the quality threshold criteria as they did not include AUC analyses. One publication was unobtainable despite contacting the authors. A final total of 17 publications were included in the review.

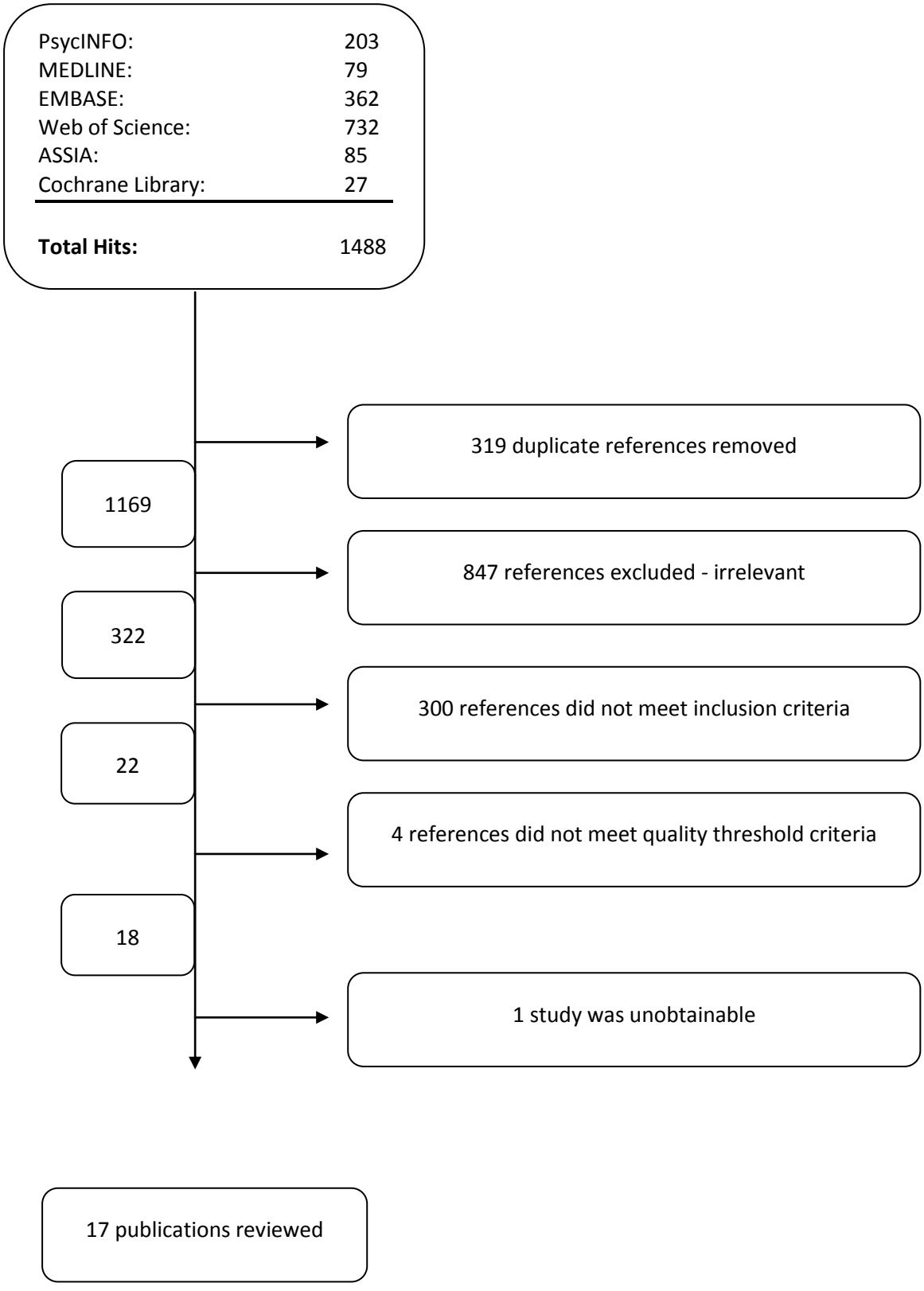


Figure 2.1 Systematic Review Search Strategy

2.3.1 Quality Assessment of Included Studies

Inter-rater reliability analyses revealed substantial agreement between raters, with Kappa=.79 (p<.001; Landis & Koch, 1977). Instances where raters did not agree on quality assessment score were discussed between raters until a consensus was reached between both parties.

CASP quality scores for the included studies ranged from 21 (Singh et al., 2011) to 34 (Dickens & O'Shea, 2015). All samples were considered to be recruited in an acceptable way and representative of the chosen population. However, generalizability outside of the study varied due to small sample sizes and studies including exclusively male participants or a minority of females (e.g. Braithwaite et al., 2010).

Only six of the included studies measured inter-rater reliability (Abidin et al., 2013; Braithwaite et al., 2010; Desmarais et al., 2012; Hogan & Olver, 2016; Nicholls et al., 2006; Wilson et al., 2013). With the exception of Braithwaite et al., (2010) these studies demonstrated Intra-class correlations or Spearman's r values of 0.8 or above, with the exception of START strength score in Abidin et al., (2013) and Hogan and Olver (2016).

Five of the included studies reported on subject attrition and reasons for non-inclusion in the study (Chu et al., 2011; Dickens & O'Shea, 2015; Gray et al., 2011; O'Shea & Dickens, 2016; Viljoen et al., 2016). These included patients discharged before the study completed, not having been admitted to the service for a long

enough period of time, and missing data from risk assessments (e.g. Dickens & O'Shea, 2015).

2.3.2 Characteristics of Included Studies

The characteristics of the included studies are displayed in Table 2.2. The overall number of participants from the 17 studies included in this review is 2945. As two studies utilised the same sample (O'Shea & Dickens, 2015; O'Shea & Dickens, 2016), these participants were only counted once. Similarly, O'Shea and Dickens (2014) conducted a meta-analysis whose sample included participants from other studies included in this review (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Gray et al., 2011; Nonstad et al., 2010; Wilson et al., 2013). Again, these participants were only counted once. As a result there are actually 2322 unique participants within this review.

Study sample sizes ranged from 27 (Inett et al., 2014) to 915 (Singh et al., 2011). Mean age of participants ranged from 31.86 years (O'Shea & Dickens, 2015) to 46.59 years (Viljoen et al., 2016). Ten of the included studies had both male and female participants (Abidin et al., 2013; Chu et al., 2011; Dickens & O'Shea, 2015; Hogan & Olver, 2016; Nicholls et al., 2006; Nonstad et al. 2010; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Singh et al., 2011; Viljoen et al., 2016). The remaining studies utilised only male participants, with the exception of O'Shea and Dickens (2014) where it was unclear which studies were included in this meta-analysis, meaning sample gender cannot be commented upon.

Table 2.2

Included Studies

Study	Quality Score	Study Type	Risk Assessment	n	Gender	Mean Age (SD)	Setting	Outcome
Abidin et al., (2013)	26	Cohort	START & SAPROF	98	Male & Female	40.45 (12.8)	High, medium, & low secure forensic mental health service	Harm to others
Braithewaite et al., (2010)	27	Cohort	START	34	Male	40.40 (13.38)	Locked rehabilitation service	Aggression towards others
Chu et al., (2011)	25	Cohort	START	50	Male & Female	34.66 (13.92)	High secure forensic psychiatric hospital	Inpatient aggression Interpersonal violence Verbal threat
Desmarais et al., (2012)	30	Cohort	START	120	Male	37.97 (11.74)	Secure forensic psychiatric hospital	Any aggression Verbal aggression Physical – objects

								Physical – others
Dickens & O'Shea (2015)	34	Cohort	START	217	Male & Female	34.10 (15.0)	Low and medium secure mental health service	Physical aggression Verbal aggression
Gray et al., (2011)	32	Cohort	START	44	Male	40.20 (14.7)	Medium secure and civil psychiatric hospital	Violence to others Verbal aggression
Hogan & Olver (2016)	27	Cohort	START	99	Male & Female	36.70 (12.8)	High secure forensic psychiatric hospital	Any aggression
Inett et al., (2014)	24	Cohort	START	27	Male	39.00	Low secure hospital	Harm to others Overt aggression Verbal aggression Physical aggression
Nicholls et al., (2006)	31	Cohort	START	137	Male & Female	M = 38.31 (11.25) F = 39.61 (11.44)	Forensic psychiatric hospital	Verbal aggression – others Physical aggression – objects

								Physical aggression - others
Nonstad et al., (2010)	23	Cohort	START	47	Male & Female	36.00	High secure hospital	Physical violence
O'Shea & Dickens., (2014)	26	Meta-analysis	START	543				Any aggression Physical – others Physical – objects Verbal aggression
O'Shea & Dickens., (2015)	26	Cohort	START	200	Male & Female	M = 35.15 (15.8) F = 31.86 (13.0)	Medium and low secure mental health service	Any aggression Physical aggression Verbal aggression
O'Shea et al., (2016)	32	Cohort	START	200	Male & Female	34.30 (15.2)	Medium and low secure mental health service	Any aggression Physical aggression Verbal aggression
Quinn et al., (2013)	25	Cohort	START	80	Male & Female	38.35 (10.06)	Medium secure forensic mental health	Any aggression

							service	
Singh et al., (2011)	21	Systematic review	SAVRY	915				Violent offending
Viljoen et al., (2016)	31	Cohort	START & SAPROF	102	Male & Female	M = 46.59 (12.68) F = 46.59 (12.13)	Psychiatric hospital	Physical aggression against others Verbal aggression
Wilson et al., (2013)	29	Cohort	START	30	Male	37.10 (12.07)	High, medium, and low secure forensic psychiatric hospital	Any aggression

The majority (15) of the publications were cohort designs (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Dickens & O'Shea, 2015; Gray et al., 2011; Hogan & Olver, 2016; Inett et al., 2014; Nicholls et al., 2006; Nonstad et al. 2010; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Viljoen et al., 2016; Wilson et al., 2013). One systematic review was included (Singh et al., 2011) and one meta-analysis (O'Shea & Dickens, 2014).

Of the included studies, seven were located in the UK (Dickens & O'Shea, 2015; Gray et al., 2011; Inett et al., 2014; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013). Six studies were located in Canada (Braithwaite et al., 2010; Desmarais et al., 2012; Hogan & Olver, 2016; Nicholls et al., 2006; Viljoen et al., 2016; Wilson et al., 2013). One study was located in Ireland (Abidin et al., 2013), and one study in Australia (Chu et al., 2011). It was not possible to identify the location of the systematic review by Singh et al., (2011), as the authors did not comment on the location of their included studies. Not all of the data reported in this study were included in the present review. Therefore it was unclear as to which studies included by Singh et al., (2011) contributed to the relevant data extracted for the present review.

All studies utilised inpatients in secure mental health hospitals. The level of security varied across studies, including locked rehabilitation, low, medium, and high secure settings. Only two studies did not comment on the level of security of the

psychiatric hospital in question (Nicholls et al., 2006; Viljoen et al., 2016). Length of follow-up period varied across studies, ranging from one month (e.g. Dickens & O'Shea, 2015) to one year (e.g. Viljoen et al., 2016). Gray et al., (2011) recorded number of incidents per 100 days as several participants left before the end of their allocated follow-up period. Hogan and Olver (2016) did not report the length of follow-up period. Due to the nature of the systematic review and meta-analysis included in this review, it is not possible to comment on the follow-up period due to their inclusion of multiple studies with varying timeframes (O'Shea & Dickens, 2014; Singh et al., 2011).

Outcome measures varied across the included studies. These comprised of any aggression, harm to others, physical aggression/violence, verbal aggression/threat, severe aggression, overt aggression, inpatient aggression, and interpersonal violence (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Dickens & O'Shea, 2015; Gray et al., 2011; Hogan & Olver, 2016; Inett et al., 2014; Nicholls et al., 2006; Nonstad et al. 2010; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Singh et al., 2011; Viljoen et al., 2016; Wilson et al., 2013).

2.3.3 Risk Assessments Examined in Included Studies

The majority (16) of the included studies investigated the predictive validity of the START (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Dickens & O'Shea, 2015; Gray et al., 2011; Hogan & Olver, 2016; Inett et al.,

2014; Nicholls et al., 2006; Nonstad et al. 2010; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Viljoen et al., 2016; Wilson et al., 2013). Two studies investigated the predictive validity of both the START and the SAPROF (Abidin et al., 2013; Viljoen et al., 2016). One study investigated the predictive validity of the SAVRY (Singh et al., 2011). In instances of publications reporting on additional risk assessment tools that do not incorporate protective factors (e.g HCR-20), these data are not reported due to the focus of this review.

2.3.4 Predictive Validity

AUC values for all included studies are presented in Table 2.3. Predictive accuracy is measured by the area under a Receiver Operating Characteristic (ROC) curve. ROC curves are created by plotting the true-positive rate against the false-positive rate (Dolan & Doyle, 2000). The area under the curve measures discrimination, that is, the ability of the tool to accurately categorize those who have engaged in a risk behaviour and those who have not (Dolan & Doyle, 2000). AUC values range from 0 to 1, where .5 represents chance (Dolan & Doyle, 2000). AUC values can be compared to Cohen's *d* effect sizes, where AUC values of .556, .638, and .714 are equivalent to Cohen's *d* of .1 (small), .5 (moderate), and .8 (large; Rice & Harris, 2005). Significant AUC values across the included studies ranged from .14 to .91 (Quinn et al., 2013).

Table 2.3

AUC Statistics for Risk Assessments

		Outcome AUC Data							
Risk Assessment	**Study Author/Year**	**Any Aggression**	**Harm to Others**	**Physical Aggression/ violence**	**Verbal Aggression/ threat**	**Severe Aggression**	**Overt Aggression**	**Inpatient Aggression**	**Interpersonal Violence**
SAPROF	Abidin et al., (2013)		6M=.85**						
	Viljoen et al., (2016)			6M=.52; .56(M); .46(F) 12M=.54; .58(M); .48(F)	6M=.60; .70*(M); .42(F) 12M=.63*; .71*(M); .51(F)	6M=.53; .74(M); .31(F) 12M=.50; .67(M); .27(F)			
Protective judgement	Viljoen et al., (2016)			6M=.46; .54(M); .35(F) 12M=.51; .55(M); .46(F)	6M=.55; .63(M); .44(F) 12M=.60; .64(M); .54(F)	6M=.47; .63(M); .31(F) 12M=.54; .66(M); .38(F)			
SAVRY	Singh et al., (2011)			.71[++]					

START	Nicholls et al., (2016)			12M=.69*(objects) 12M=.65**(others)	12M=.67**				
Strength	Abidin et al., (2013)		6M=.78**						
	Braithwaite et al., (2010)			2M=.65*(others)					
	Chu et al., (2011)				1M=.64			1M=.71*	1M=.75*
	Desmarais et al., (2012)	12M=.76**		12M=.77**(objects) 12M=.80**(others)	12M=.75**				
	Dickens & O'Shea, (2015)			1M=.63* 2M=.67** 3M=.65**	1M=.64* 2M=.66** 3M=.69**				
	Gray et al., (2011)[b]			.21*(others)	.28*				
	Hogan & Olver, (2016)	.69*							
	Inett et al., (2014)		1M=.61 3M=.58	1M=.25** 3M=.29**	1M=.34** 3M=.43		1M=.59 3M=.72**		
	Nonstad et al., (2010)			3M=.77**					
	O'Shea & Dickens, (2014)	.71[+]		.70(objects)[+] .75(others)[+]	.71[+]				
	O'Shea & Dickens, (2015)	3M=.68*(M) 3M=.68(F)		3M=.65*(M) 3M=.62(F)	3M=.66*(M) 3M=.72*(F)				

	O'Shea et al., (2016)	3M=.69**		3M=.66**	3M=.67**			
	Quinn et al., (2013)	1M=.14* 3M=.32 6M=.43						
	Viljoen et al., (2016)			6M=.65*; .71*(M); .57(F) 12M=.60; .63(M); .57(F)	6M=.61; .64(M); .54(F) 12M=.64*; .64(M); .65(F)	6M=.62; .67(M); .56(F) 12M=.66*; .65(M); .66(F)		
	Wilson et al., (2013)[a]	12M=.84*						
Vulnerability	Abidin et al., (2013)		6M=.82**					
	Braithwaite et al., (2010)			2M=.66**(others)				
	Chu et al., (2011)				1M=.77*		1M=.76*	1M=.78*
	Desmarais et al., (2012)	12M=.79**		12M=.80**(objects) 12M=.77**(others)	12M=.79**			
	Dickens & O'Shea, (2015)			1M=.60* 2M=.58* 3M=.61*	1M=.60* 2M=.57 3M=.57			
	Gray et al., (2011)[b]			.68*(others)	.74*			
	Hogan & Olver, (2016)	.76**						

	Inett et al., (2014)		1M=.42 3M=.38	1M=.71** 3M=.69**	1M=.66** 3M=.57		1M=.41 3M=.32*	
	Nonstad et al., (2010)			3M=.77**				
	O'Shea & Dickens, (2014)	.74^{+}		.77(objects)$^{+}$ .73(others)$^{+}$	.78^{+}			
	O'Shea & Dickens, (2015)	3M=.59(M) 3M=.74*(F)		3M=.59(M) 3M=.73(F)	3M=.57(M) 3M=.75*(F)			
	O'Shea et al., (2016)	3M=.64**		3M=.64**	3M=.61*			
	Quinn et al., (2013)	1M=.85* 3M=.74 6M=.67						
	Viljoen et al., (2016)			6M=.50; .51(M); .48(F) 12M=.52; .54(M); .47(F)	6M=.59; .61(M); .59(F) 12M=.61; .61(M); .62(F)	6M=.54; .76(M); .35(F) 12M=.56; .71*(M); .39(F)		
	Wilson et al., (2013)[a]	12M=.82*						
SRE	Braithwaite et al., (2010)			2M=.52 (others)				
	Chu et al., (2011)							1M=.69
	Desmarais et al., (2012)	12M=.80**		12M=.84**(objects) 12M=.85**(others)	12M=.78**			

Dickens & O'Shea, (2015)		1M=.71** 2M=.68** 3M=.68**	1M=.65* 2M=.62* 3M=.62*	
Gray et al., (2011)[b]		.65	.70*	
O'Shea & Dickens, (2014)	.74[+]	.76(others)[+]	.76[+]	
O'Shea & Dickens, (2015)	3M=.68*(M) 3M=.83*(F)	3M=.68*(M) 3M=.85*(F)	3M=.59(M) 3M=.71(F)	
O'Shea et al., (2016)	3M=.69**	3M=.70**	3M=.64*	
Quinn et al., (2013)	1M=.91* 3M=.68 6M=.58			
Viljoen et al., (2016)		6M=.69*; .69*(M); .69*(F) 12M=.67*; .64(M); .70*(F)	6M=.68*; .68*(M); .68*(F) 12M=.63*; .62(M); .66(F)	6M=.64; .64(M); .65(F) 12M=.71*; .73*(M); .69(F)
Wilson et al., (2013)[a]	12M=.89**			

Note. *$p \leq .05$; **$p \leq .001$; SRE = Specific Risk Estimate; 1M = one month follow-up; 2M = two month follow-up; 3M = three month follow-up; 6M = six month follow-up; [+] mean weighted AUC value (no *p* reported) ; [++] median AUC value (no *p* reported); (M) = Male; (F) = Female

[a] Additional data excluded due to lack of clarity regarding categorization of incidents
[b] Data recorded as number of incidents per 100 days

2.3.4.1 SAPROF Two studies investigated the predictive validity of the SAPROF. Abidin et al., (2013) found a large effect size for harm to others. Viljoen et al., (2016) found significant predictive validity for males in relation to verbal aggression at both six-month and 12-month follow-up. When the protective judgement aspect of the SAPROF was separated out and its predictive validity examined independently, no significant results were found (Viljoen et al., 2016).

2.3.4.2 SAVRY Only one study explored the predictive validity of the SAVRY. Singh et al., (2010) identified that the SAVRY was predictive of physical aggression. However this was a median AUC value drawn from a systematic review of eight studies, and there was no significance value reported.

2.3.4.3 START 16 studies examined the predictive validity of the START. One study found that START total scores were predictive of both physical and verbal aggression (Nicholls et al., 2016). 15 studies revealed that the START strength scale was predictive of inpatient aggression, whether defined as any aggression, harm to others, physical aggression/violence, verbal aggression/threat, severe aggression, overt aggression, inpatient aggression, or interpersonal violence (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Dickens & O'Shea, 2015; Gray et al., 2011; Hogan & Olver, 2016; Inett et al., 2014; Nonstad et al. 2010; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Viljoen et al., 2016; Wilson et al., 2013). However, some studies

showed that the predictive validity of the strength scale varies across follow-up periods. For example, Inett et al., (2014) demonstrated that the predictive validity of the strength scale was not significant one month after assessment, but increased to a large significant effect size (AUC=.72) at a three-month follow-up for severe aggression. Two studies separated the predictive validity of the strength scale by gender (O'Shea & Dickens, 2015; Viljoen et al., 2016). These studies demonstrated that the only gender difference was for any aggression and physical aggression, with the strength scale only being predictive for males (O'Shea & Dickens, 2015; Viljoen et al., 2016).

These 15 studies also found that the START vulnerability scale was predictive of inpatient aggression, howsoever defined (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Dickens & O'Shea, 2015; Gray et al., 2011; Hogan & Olver, 2016; Inett et al., 2014; Nonstad et al. 2010; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Viljoen et al., 2016; Wilson et al., 2013). Quinn et al., (2013) demonstrated that the vulnerability scale was predictive of any aggression after one month follow-up but not after three months. Similarly, Inett et al., (2014) showed no change in predictive validity of the vulnerability scale from one-month follow-up (AUC=.71) to three-month follow-up (AUC=.69) for physical aggression. A similar finding was also reported for verbal aggression (Inett et al., 2014). Gender differences were also highlighted in two of the included studies. O'Shea & Dickens (2015)

established that the vulnerability scale was predictive for females in relation to any aggression and verbal aggression, but not for males. Additionally Viljoen et al., (2016) showed that after twelve months the vulnerability scale was predictive of severe aggression for males but not females.

11 studies explored the predictive validity of the Specific Risk Estimate (SRE) for violence towards others (Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Dickens & O'Shea, 2015; Gray et al., 2011; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Quinn et al., 2013; Viljoen et al., 2016; Wilson et al., 2013). These studies demonstrated the SRE was predictive of 'any aggression', 'physical aggression', 'verbal aggression', and 'severe aggression'. Quinn et al., (2013) revealed the biggest difference in predictive validity across follow-up periods, with a decrease in predictive validity from one month (AUC=.91) to three months (AUC=.68) to six months (AUC=.58), with these final two AUC values not being significant. Dickens and O'Shea (2015) showed a slight decrease in the predictive validity of the SRE in relation to physical and verbal aggression. Conversely, Viljoen et al., (2016) exhibited an increase in predictive validity of the SRE from six to twelve month follow-up for severe aggression. In terms of gender differences O'Shea & Dickens (2015) demonstrated a stronger predictive validity for the SRE for females over males for both 'any aggression' and 'physical aggression'. Similarly, Viljoen et al., (2016) found that the SRE was predictive after twelve months of physical aggression in females but not males (after 12 months) for

physical aggression. However it was also demonstrated that SRE was predictive of severe aggression in males but not females (Viljoen et al., 2016).

2.4 Discussion

The findings from this systematic review show little support for the inclusion of protective factors in inpatient risk assessment for aggressive or violent outcomes. Whilst all three risk assessments demonstrated a level of predictive validity, it was difficult to ascertain the specific utility of the protective factors in the risk assessments highlighted in this review. The SAPROF data revealed that whilst the total score of the measure was predictive of inpatient aggression, when the protective judgement aspect of the SAPROF was looked at in isolation, this did not produce significant results for either six-month or 12-month follow-up (Viljoen et al., 2016). However, as this is only one study, it is difficult to draw conclusions and thus more prospective research on the SAPROF is needed to determine the true predictive validity of the protective judgement component.

Similarly, only one study in the current review examined the SAVRY. Singh et al., (2010) demonstrated that the SAVRY predicted physical aggression. However this was a median AUC value taken from a systematic review. Unlike the SAPROF study, the six protective factors included in the SAVRY were not separated out to determine their unique influence on predictive validity. Therefore whilst the SAVRY has demonstrated predictive validity for physical

aggression, the unique contribution of the protective factors cannot be determined.

The majority of the results from this review pertain to the START, and thus will be the main focus of this discussion. The findings from the START clearly show the utility of including a strength scale. All fifteen studies that investigated this scale found significant results for the prediction of aggressive or violent outcomes. Within the studies, the predictive validity of the strength scale either remained consistent or increased (in terms of effect) with length-of follow-up period (Dickens & O'Shea, 2015; Viljoen et al., 2016). The exception to this was for Inett et al., (2014) and Quinn et al., (2013), who demonstrated a loss of significance from one to three months for verbal aggression and any aggression respectively. Viljoen et al., (2016) also had one variable (physical aggression) that was significantly predicted by the strength scale at six months, but not at twelve months. The two studies that investigated gender differences found that the strength scale was not predictive for females for any aggression or physical aggression (O'Shea & Dickens, 2015; Viljoen et al., 2016). However, in both of these studies females made up a smaller proportion of the investigated sample, which may have influenced these findings.

The START vulnerability scale demonstrated a high level of predictive validity for the same aggressive or violent outcomes as the START Strength scale. In comparison to the strength scale, whilst six studies showed the vulnerability scale to be more predictive of aggressive or violent outcomes, another six studies

demonstrated minimal or no difference in effect size for the predictive validity of aggressive outcomes between the strength and vulnerability scales (Abidin et al., 2013; Braithwaite et al., 2010; Chu et al., 2011; Desmarais et al., 2012; Gray et al., 2011; Hogan & Olver, 2016; Inett et al., 2014; Nicholls et al., 2006; Nonstad et al. 2010; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; Quinn et al., 2013; Singh et al., 2011; Wilson et al., 2013). In addition to this, O'Shea et al., (2016) demonstrated that the vulnerability scale had a smaller effect size for verbal aggressive than for the strength scale of the START. Dickens and O'Shea (2015) found that effect sizes were smaller for the vulnerability scale predicting physical and verbal aggression. These studies support the predictive validity of protective factors in comparison to aggravating risk factors for inpatient aggression.

In terms of predictive validity across follow-up periods, the vulnerability scale did not perform as well as the strength scale. Dickens and O'Shea (2015) showed that the vulnerability scale performed worse across follow-up periods than the strength scale, only significantly predicting verbal aggression at the one-month follow-up. While Inett et al., (2014) demonstrated an increase in effect size for the strength scale across follow-up periods, their findings showed that the vulnerability scale either exhibited a reduction in effect size or a loss of significance. These findings suggest that the strength scale performs better over time. This implies that protective factors may be more stable predictors of inpatient aggression.

In relation to gender differences, the vulnerability scale produced similar findings to the strength scale (O'Shea & Dickens, 2015; Viljoen et al., 2016). This suggests that gender does not influence the predicative validity of either scale.

The SRE data also showed significant predictive validity across outcome measures. The SRE judgements outperformed both the strength and vulnerability scales across all outcomes, genders, and follow-up periods (with the exception of verbal aggression; Desmarais et al., 2012; O'Shea & Dickens, 2014; O'Shea & Dickens, 2015; O'Shea et al., 2016; Viljoen et al., 2016) and any aggression (O'Shea & Dickens., 2014; O'Shea et al., 2016). In relation to follow-up periods, the SRE still outperformed the strength and vulnerability scale for the majority of outcomes, although appeared to reduce in predictive validity across time for physical and verbal aggression (Dickens & O'Shea, 2015; Quinn et al., 2013; Viljoen et al., 2016). Again, gender did not contribute to the predictive validity of the SREs, except for physical aggression in men at the six-month follow-up (Viljoen et al., 2016). The fact that SREs appeared to outperform both the strength and vulnerability scale on most outcomes is perhaps not surprising. This professional judgement aspect of the START is generated through a consideration of all of the items on the START assessment (rated as both a strength and a vulnerability). Therefore it is logical to find that this is the most predictive aspect of the START scale, as it is a consideration of both risk and protective factors, so leading to a more balanced risk assessment for inpatient aggression.

The most predictive aspect of the START was the SRE scale, which is derived from a consideration of both the strength and the vulnerability scale within the START assessment. Therefore it may be less important that a risk assessment separates out protective factors, but rather that the overall score incorporates these, and thus increases the predictive validity of the assessment, as demonstrated by the START SRE data being more predictive than either the strength of vulnerability score individually. This combination effect highlights the importance of incorporating both risk and protective factors to create a more balanced and subsequently more accurate assessment of individual risk.

2.4.1 Limitations and Future Considerations

One major limitation to this systematic review is the lack of variety of risk assessments included. The majority of the studies examined the START, with only two investigating the SAPROF, and one the SAVRY. The IORNS was not represented at all in this review. This may be illustrative of the use of these tools within secure services, and thus the amount of literature dedicated to investigating the psychometric properties of said assessments. However with this lack of variety, it is difficult to draw concrete conclusions with regard the predictive validity of protective factors within risk assessment.

This lack of variety in risk assessment tools identified in the current review may also reflect the limitations of the search strategy. It may be that the SAPROF and, SAVRY, and IORNS are more likely to be used in a community setting, rather than secure

mental heath. Therefore these studies would not have been captured under the current search strategy.

Similarly, the inclusion of a youth measure amongst several adult measures may have created conceptual issues and thus it may not be accurate to draw overall conclusions from the findings of this review. Future reviews should focus on either adult or youth measured to address this issue.

The inclusion of a systematic review and a meta-analysis may have also impacted on the findings of the current systematic review. Due to the nature of these designs their inclusion may have limited the amount of information that could be interpreted. For example, within these studies the follow-up period could not be commented on due to the inclusion of multiple studies with varying timeframes. Similarly, information may have been lost in relation to the outcome measures. For example, the outcome measure in Singh et al., (2011) was violent offending, however it may be that if the individual studies included in this article were identified as individual studies, there could have been more detailed information available for comparison, rather than the general term "violent offending". This means the inclusion of the review rather than the individual studies themselves has impacted on the detail of the data used in the present review, thus limiting the subsequent conclusions made.

The use of Singh et al., (2011) in particular may have created some methodological issues. Whilst this looked at nine different risk assessment tools in total, only one tool (SAVRY) utilised protective factors. Therefore this was the only information used in the current

review. However, this meant that a median value for AUC was obtained, and it would have been beneficial to seek out the individual studies included in Singh et al., (2011) pertaining to the SVARY and include them individually. This would have yielded more information for comparison with the other cohort studies, and potentially provided more accurate conclusions.

Another limitation is the inconsistency of follow-up period. Direct comparisons across studies regarding follow-up period and subsequent changes in predictability could not be made due to different sample characteristics. This needs to be addressed in future research in order to reach a more robust conclusion about the predictive validity of these risk assessments over time.

Similarly, the studies included in the current review measured a variety of different outcome variables. This made direct comparisons difficult across studies. Uniformity of outcome variables would allow more concrete comparisons and subsequently improve the support for the inclusion of protective factors in risk assessment tools.

Settings were not commented on in results because several studies recruited participants across security levels and it was not possible to determine these as separate sub-samples. Future research should focus on individual security levels in order to gain more information about the impact this may have on the predictive validity of these risk assessment tools.

2.4.2 Conclusion

Whilst in the case of the SAPROF and SAVRY, the clinical utility for the inclusion of protective factors is underwhelming, it may be that simply having the protective factors incorporated into the assessment influences raters perceptions of the risk items, thus creating a more balanced risk assessment. However, until the research is conducted to separate out the protective components of these assessments, these conclusions cannot be made.

The majority of the research in this review was conducted on the START. This demonstrates both the frequency of use within secure mental health services, but also the importance of determining the role of protective factors in risk assessment. Whilst individual studies reported larger AUC statistics for the strength scale than the vulnerability scale, the SREs produced the highest level of predictive validity. As the SRE's are formed from a consideration of both the protective and risk factors of the START, these findings demonstrate the importance of including both risk and protective factors in the risk assessment process, and the combination of these two scales produced the most accurate representation of patient risk. Therefore the findings from this review demonstrate some support for a consideration of protective factors within the risk assessment process for inpatient aggression in mental health services, albeit underwhelming.

Rationale for Chapter Three

The findings from the systematic review highlighted the potential efficacy of incorporating protective factors in risk assessment. The majority of the literature focused on the Short-Term Assessment of Risk and Treatability (START), which may be a reflection of the frequency of use within secure services. Therefore, in order to establish the scientific properties of the START, chapter three critically evaluates the START to determine its clinical utility as a risk assessment tool.

Chapter Three

The Short-Term Assessment of Risk and Treatability:

A Critique

3.0 Abstract

Purpose: This critique aimed to determine the scientific properties of the Short-Term Assessment of Risk and Treatability (START), in order to draw conclusions about the clinical utility and applicability of this structured professional judgement tool.

Background: The development of risk assessment tools has seen four distinct generations. Moving from unstructured clinical judgement, to actuarial and structured professional judgement. The fourth generation aims to encompass protective factors in order to aid both management and treatment of individual risk. The START is one such tool that encompasses a measure of both strengths and vulnerabilities.

Method: The structure, reliability, and validity of the START were examined against existing risk assessment tools to determine the clinical utility of the START.

Results: The START was shown to be both valid and reliable, with excellent inter-rater reliability. Good correlations between scales were exhibited, in addition to high convergent validity with other risk assessment tools such as the HCR-20. Predictive validity of the START varied between the strength scale, vulnerability scale, and specific risk estimates, depending on the outcome of interest.

Conclusions: The clinical utility of the START may lie more in treatment planning rather than risk prediction. However, most of the literature has only looked at violent outcomes, therefore more research is needed into the other six risk domains to make a conclusive comment of the clinical utility of the START.

3.1 Introduction

The process of risk assessment has developed considerably in recent years. From clinical judgement to structured tools, the focus of risk assessment and definitions of what makes a good risk assessment tool have changed dramatically. Andrews, Bonta and Wormith (2006) identify four distinct generations of offender risk assessment, which have been shaped by years of theoretical and practical research.

The first generation of risk assessment relied upon an unstructured process of clinical judgement to determine the probability of challenging or offending behaviour (Andrews et al., 2006; Burman et al., 2007). Whilst this is the starting point for many new approaches, the likelihood of bias and inconsistency means this method is considered unreliable (Burman et al., 2007).

Second generation tools are empirically based instruments. These tools have a focus on static, historical risk factors such as number of previous convictions (Andrews et al., 2006; Burman et al., 2007). Whilst this generation is seen as more objective than the previous generation, there are still some limitations. The reliance on static factors creates a rigid risk assessment, and implies that an individual's risk cannot change (Burman et al., 2007). Whilst the static factors may contribute to an individual's likelihood to engage in offending behaviour, the immanency or seriousness may be more accurately determined by dynamic factors (Burman et al., 2007).

Attempting to address these limitations, the third generation of risk assessments utilised both clinical judgement and actuarial

methods (Andrews et al., 2006; Burman et al., 2007). These structured professional judgement (SPJ) tools utilise both static and dynamic factors identified through empirical research to form an overall judgement of an individual's risk (Andrews et al., 2006; Burman et al., 2007). This acknowledgement of criminogenic needs creates an individualised risk assessment that maintains the predictive validity of more actuarial methods (Burman et al., 2007). Subsequently SPJ tools can be used to guide intervention and capture changes in an individual's risk over time (Andrews et al., 2006; Burman et al., 2007). However, by including an element of clinical judgement, there is inherently an increase in the risk of bias and issues with consistency (Burman et al., 2007).

Finally, the fourth generation tools aim to inform treatment planning and management in addition to predict risk. Rather than simply predicting an individual's risk, these tools aim to bring together information about an individual's history and criminogenic needs in order to tailor the intervention to the individual (Andrews et al., 2006; Burman et al., 2007). These risk assessments also include an acknowledgment of an individual's strengths or protective factors that may help to reduce the likelihood of an adverse outcome (Andrews et al., 2006; Burman et al., 2007). This new era of risk assessments aim not only to predict risk, but support treatment planning and identification of key responsivity characteristics (Burman et al., 2007).

One example of a fourth generation risk assessment tool is the Short-Term Assessment of Risk and Treatability (START;

Webster et al., 2009). This review examines the efficacy of the START, in terms of its scientific properties and clinical utility.

3.1.1 Overview of the START

The primary aim of the START is to predict risk in inpatient settings in order to aid in the management and treatment planning of patients. The START is an SPJ tool, which focuses on dynamic risk factors. The START identifies acute risk of individuals engaging in adverse outcomes within a three-month period. These risk areas include risk to others, risk of self-harm, risk of suicide, risk of unauthorised leave, risk of substance abuse, risk of self-neglect, and risk of being victimized (Webster et al., 2009). The START is unique to other risk assessment tools as it rates each item as both a strength and a vulnerability, which allows clinicians to acknowledge protective factors as well as risk factors that will impact patient behaviour.

There are 20 risk factors included in the START. These factors are dynamic, meaning they can change over time. The aim of including dynamic risk factors is to acknowledge that patient risk can change over a relatively short period of time. By using dynamic factors, the START assessment can be regularly updated to reflect patient change, making the START more able to predict risk in the short-term. This is reflected in the START guidelines, which state that the START should be updated after every incident or change in the patient, and is ideally used to predict risk over a three-month period (Webster et al., 2009).

3.1.1.1 Development of the START

The drive to develop a new means of risk assessment came from St. Joseph's Healthcare Hamilton (Webster et al., 2009). Before the development of the START, professionals were using the Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962) to provide a context for clinical discussions and decision-making. However several of the professionals had 'lost touch' with the definitions of items on the BPRS, therefore reducing the inter-rater reliability of the tool to an unacceptable level (Webster et al., 2009). Additionally, while the items on the BPRS mapped on to the symptomatology of psychiatric diagnoses, they did not sufficiently reflect the risk posed to self and others (Webster et al., 2009). Therefore, it was decided that a new evidence-based tool was needed that would address both of these issues.

It was at this time that the developers acknowledged the utility of recognising protective factors within a risk assessment. By engaging in a self-reflective process, 50 or so variables were compiled and through weeks of clinically defining definitions with a multi-disciplinary team, these were reduced and refined to the 20 items we now see on the START (Webster et al., 2006).

3.1.1.2 Structure

As with the majority of risk assessment tools, the START utilises a Likert-type scale response style. Each of the 20 items is rated as minimally present, moderately present, or maximally present. The use of Likert and Likert-type scales has been discussed to some extent in the literature. These findings suggest that whilst

these scales provide an easy and accessible way of conducting assessments, they can be somewhat arbitrary. Despite being an ordinal scale, the interval between values on a Likert-type scale should not be presumed equal (Blaikie, 2003). The use of a Likert-type scale produces a subjective answering style, which could influence the reliability of the tool in question. Although subjectivity may be reduced with a three-point scale rather than a five-point scale, the definitions of "minimally" vs. "maximally" present are still open to interpretation.

In addition to the individual ratings for the 20-items, the START includes specific risk estimates for seven risk domains, comprising: violence risk to others, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, and victimisation by others. However, these specific risk estimates are not derived from a total score from the 20 items, but rather a consideration of these items. This is the professional judgement aspect of the risk assessment, and allows clinicians to take into account the total strength and vulnerability scores and come to a professional decision regarding the seven risks.

The START also allows the addition of case specific risk estimates that give clinicians the opportunity to add any risks specific to the individual that are not already incorporated into the START. This allows a unique opportunity for clinicians to adapt a risk assessment to the individual to allow a greater accuracy of risk prediction.

3.2 Reliability

3.2.1 Internal Consistency

Measures of internal consistency reported in the existing START literature use two different measures: Cronbach's alpha and mean inter-tem correlations (MICs). Whilst unidimensionality cannot be assumed from high internal consistency, it has been suggested that Cronbach's alpha values greater than .80 and MICs ranging from .15 to .50 represent unidimensional scales (Clark & Watson, 1995).

Nicholls et al., (2007) reported Cronbach's alpha of .80 for the START as a whole, while Abidin et al., (2013) report Cronbach's alphas of .95 for both the strength and the vulnerability scale ($p<.001$). Mean inter-item correlations ranged from .18 to .20, reflecting good item homogeneity (Nicholls et al., 2006). These findings are comparable to other risk assessment measures including the HCR-20(v3) and the PCL:SV demonstrating an acceptable level of homogeneity within the START (Douglas et al., 2013; Hart, Hare, & Cox, 1995; Nicholls et al., 2006).

High levels of internal consistency may be surprising for the START as it is designed to measure seven different risk domains. However, as risk assessments are designed to measure outcomes rather than an underlying construct that is related to risk, this may be irrelevant (Rosenfeld & Penrod, 2011).

3.2.2 Correlations Between Scales

Desmarais, Nicholls, and Brink (as cited in Webster et al., 2009) reported the association between total strength and

vulnerability scores and specific risk estimates. These are displayed in Table 3.1. This research demonstrates significant negative correlations between strength scores and specific risk estimates, and significant positive correlations between vulnerability scores and specific risk estimates. The correlation between strength score and self-harm was found to be in the expected direction, however was not statistically significant.

Table 3.1

Relationships between START Total Scores and Risk Estimates

	START Total Scores	
Risk Estimates	**Strength**	**Vulnerabilities**
Violence	-.54***	.56***
Self-harm	-.10	.13*
Suicide	-.13*	.16*
Unauthorised leave	-.48***	.48***
Substance abuse	-.25***	.24***
Self-neglect	-.45***	.44***
Victimisation	-.31***	.30***

Note. Adapted from Desmarais, Nicholls, & Brink (2008). Values represent Pearson correlations; n=142; *p<.05; **p<.01; ***p<.001

Similarly, Wilson, Desmarais, Nicholls, & Brink (2010) demonstrated that strength and vulnerability scale totals have significant negative correlations (r=-.90; p,<.01). Braithwaite, Charette, Crocker, and Reyes (2010) identified that individual item scores on the START such as emotional state, relationships, and

coping appear to be readily reflected as both a strength and a vulnerability (*r*=-.53, -.36, and -.54 respectively). Conversely other items such as material resources (*r*=.-.90) are more mutually exclusive, thus questioning the utility of having this item on both scales.

This research shows that the START has significant negative correlations between the strength and vulnerability scales, which raises the question of the clinical utility of rating each item as both a strength and vulnerability.

3.2.3 Inter-Rater Reliability

When considering the clinical application of the START, it is important to consider the inter-rater reliability. Inter-rater reliability is a measure of decision agreeability between two raters assessing the same individual. It is not appropriate to assume that in a clinical setting the START will be completed by any one individual, therefore inter-rater reliability is crucial for a tool of this nature.

The current literature demonstrates that the START has high levels of inter-rater reliability, with mean intraclass correlation coefficients (ICCs) ranging from .85 to .95 (Desmarais, Nicholls, Wilson, & Brink, 2012; Nicholls et al., 2006). Nicholls et al., (2006) also found excellent inter-rater reliability between professional disciplines including nursing (ICC=.88; $p<.001$), social work (ICC=.92; $p<.001$), and psychiatry (ICC=.80; $p<.001$).

These findings demonstrate excellent inter-rater reliability rates between disciplines, thus supporting the clinical application of the START within organizations with a multi-disciplinary team.

3.3 Validity

3.3.1 Construct Validity

Construct validity is the degree to which a tool measures what it is supposed to measure. Desmarais, Nicholls and Brink (as cited in Webster et al., 2009) looked at the clinical utility of the START across three levels of security. Their findings demonstrated good construct validity, as total strength scores increased and vulnerability decreased as security level decreased (Webster et al., 2009). Similar findings were demonstrated for specific risk estimates, with the exception of self-harm (Webster et al., 2009).

3.3.2 Convergent Validity

In relation to other existing risk assessments, Desmarais et al., (2012) found that the strength and vulnerability scores correlated moderately to strongly with the HCR-20 subscale and total scores for violence risk; vulnerability scores demonstrated a correlation range from .31 to .88, and strength scores demonstrated a correlation range from -.43 to .77.

Abidin et al., (2013) compared the START with the Structured Assessment of Protective Factors for violence risk (SAPROF; de Vogel et al., 2012). A significant positive correlation was found between the total score on the strength scale of the START and the SAPROF (r=.69). This study also demonstrated good convergent validity with the Suicide Risk Assessment and Management Manual (SRAMM; Bouch & Marshall, 2003).

These findings suggest that the START has good convergent validity with existing measures, justifying the use of the START in a

similar clinical setting.

3.3.3 Incremental Validity

Desmarais et al., (2012) found that the strength scale demonstrated a unique contribution to the prediction of physical aggression, whereas vulnerability scores showed a unique contribution to the prediction of verbal aggression, in comparison to the HCR-20 total and subscale scores (Desmarais et al., 2012; Douglas et al., 2013). In contrast, Wilson et al., (2013) found that whilst the vulnerability scale was able to demonstrate unique contribution above that of the Historical HCR-20 scale, the strength scale did not. Neither scale demonstrated incremental validity for aggression towards objects, but the specific risk estimate for violence towards others demonstrated a unique contribution over strength and vulnerability scores combined with HCR-20 and scores on the screening version of the Psychopathy Checklist (PCL:SV; Hart, Hare, & Cox, 1995) in relation to all types of aggressive outcomes.

The current literature on the START has had a main focus on the prediction of aggressive and violent outcomes, therefore incremental validity data on the other risk outcomes of the START is limited. It would appear that strength scores do offer a unique contribution to the prediction of risk, however more research is needed in this area.

3.3.4 Predictive Validity

Area Under Curve (AUC) analyses have been used to identify the predictive validity of the START. Both strength and vulnerability

scores have shown to significantly predict any aggressive outcome, although some studies have reported AUC values only slightly above that of chance; Strength AUC values ranging from .63 to .84 (p<.05) and vulnerability AUC values ranging from .63 to .82 (p<.05; Desmarais et al., 2012; Gray et al., 2011; O'Shea et al., 2015; Wilson et al., 2013). More specifically, both the strength and vulnerabilities scales have demonstrated predictive of verbal aggression, physical aggression against others and physical aggression towards objects. Additionally Nonstad et al., (2010) found that when strength scores are equal to or greater than vulnerability scores, there were no violent incidents over the following 90 days. This demonstrates the value of considering protective factors in risk assessment.

To put this in context, the HCR-20 demonstrates AUC values ranging from .69 to .76 for predicting violent outcomes (Wilson, Desmarais, Nicholls, & Brink, 2010). This means that the predictive validity of the START for violent outcomes is comparable to the HCR-20, a well-established risk assessment tool.

Whilst the research has demonstrated that the START predicts violent and aggressive outcomes, data on the remaining risk outcomes is minimal. Additionally, what literature there is on these outcomes are inconsistent in their findings. Braithwaite et al., (2010) found that both strength and vulnerability scales significantly predicted substance abuse and unauthorized leave (AUC's range from .63 to .67; p<.05), but did not significantly predict self-harm, suicidality, or victimization. Conversely, Gray et al., (2011) found

significant AUC values for self-harm, suicidality, and victimization, and O'Shea et al., (2015) demonstrated that strength scores have a higher predictive validity for self-neglect (AUC=.66).

3.4 Conceptual Issues and Clinical Application

The specificity of risk is a long-standing issue among all risk assessments. Early on, Mulvey and Lidz (1995) stated that risk assessment tools should be context-specific for the type, probability, frequency, seriousness and imminence of risk. Whilst the START considers the type and probability of risk, it is questionable as to whether the strength and vulnerability total scores, or specific risk estimates, consider frequency, seriousness or imminence of risk. The START manual dictates the settings the START is designed for use within, fulfilling the context-specific criteria outlines by Mulvey and Lidz (1995), though the other areas remain unspecified.

It is also important to acknowledge the reciprocal nature of strengths and vulnerabilities in the START. As each item is rated as both a strength and vulnerability, this may encourage opposite answering styles, for example if a rater as marked "social skills" as a risk, they may not mark it as a strength. However Webster et al., (2009) state that each item should be considered separately as both a strength and a vulnerability, as "social skills" could be a strength in one situation but a risk under different circumstances.

When considering the clinical application of a risk assessment tool, it is important to look at user-friendliness. Collins, Demarais, Nicholls, and Brink (as cited in Webster et al., 2009) evaluated

feedback from clinicians using the START in a secure forensic psychiatric hospital about satisfaction and perceived utility of the START within this setting. These results are displayed in Table 3.2. These findings suggest that the START is user-friendly, and easy to apply in a clinical setting. One criticism highlighted was in relation to making the finer decisions of rating each individual item. However the START developers followed the same rating format as the HCR-20 in an attempt to eliminate confusion (Webster et al., 2009). Overall the findings of this study support the clinical utility of the START within a secure setting.

Table 3.2

Satisfaction of Clinicians using START in Practice

	% Agree
User-Friendliness	
Information required for completing START was readily available	88.0
Straightforward to assess item strengths and vulnerabilities independently	80.0
Easy to make the finer distinctions (i.e., deciding between 0 vs. 1 vs. 2)	56.0
Sufficient time to score the instrument	81.6
Clinical Utility – Last Assessment	
Items were readily applied to the circumstances of this patient	81.6
Instrument provided a comprehensive risk summary for this patient	83.7
Separate ratings for strengths & vulnerabilities were useful for this patient	83.3
Inclusion of the following information was useful for this patient:	
Signature risk signs	70.8
Specific risk estimates	87.2
Historical ratings of specific risk estimates	79.2
Key strength items	73.9
Critical vulnerability items	87.5

T.H.R.E.A.T ratings	72.3
Risk formulation	83.3
Clinical Utility - General	
Focus on dynamic or changeable items is clinically useful	81.1
Focus on vulnerabilities and strengths is clinically useful	92.5
Overall, START is a useful clinical measure	92.5
Assessor Confidence – Last Assessment	
Confident in ratings of:	
Individual START items	83.7
"Risk to Others" risk estimate	79.6
"Self-Harm" risk estimate	83.3
"Suicide" risk estimate	81.6
"Unauthorised Leave" risk estimate	81.6
"Substance Abuse" risk estimate	81.6
"Self Neglect" risk estimate	81.6
"Being Victimised" risk estimate	81.6
Confident in the overall accuracy of the assessment	82.6

Note. Adapted from Collins et al., (as cited in Webster et al., 2009).

3.5 Discussion

The data presented in this critique demonstrate mixed findings on the scientific properties and subsequent clinical utility of the START. The findings demonstrated excellent inter-rater reliability, where the highest agreements between professionals were in relation to the vulnerability scale. This may be reflective of the nature of risk assessment within a secure service, and that professionals have more experience identifying risks as opposed to protective factors.

The START demonstrates significant correlations between scales, in the expected negative direction, questioning the clinical utility of rating each item as both a strength and a vulnerability. It may be that this approach is more useful for treatment planning rather than risk assessment, as it can highlight strengths an individual already has, and areas in need of improvement. This may subsequently reduce the number of risk outcomes exhibited.

There is also an argument to say that if both these scales are predictive of the same outcomes, or are simply the opposite of each other, should they not be collapsed into one? However as the SREs are derived from a consideration of both the strength and vulnerability scales, including these as separate items promotes the acknowledgement of protective factors resulting in a more balanced assessment of patient risk (Rogers, 2000).

The START also demonstrated convergent validity with established violence risk assessment tools including the HCR-20 and the SAPROF, adding to the justification for clinical use. Incremental

validity is also demonstrated, again reinforcing the clinical use of the START.

The predictive validity of the START varies depending on the risk outcome in question. The strength scores, vulnerability scores, and specific risk estimates demonstrated stronger predictive validity for violence risk to others and aggressive risk outcomes than other risk outcomes. However self-harm was only significantly predicted by its specific risk estimate, not the strength or vulnerability scores. There was also insufficient evidence to determine the predictive validity of the START for suicide, unauthorized leave, and victimization by others. This may not be a statement that these risk outcomes are not predicted by the START, so much as a gap in the research as the literature has concentrated on violent risk outcomes.

Whilst the future of risk assessment is becoming more inclusive of protective factors, this may be more useful for treatment planning and risk management as opposed to risk prediction. However, the majority of the START research has focused on violent risk outcomes, therefore further research is required looking specifically at the other six risk domains before a judgment can be made on the predictive validity of protective factors with regard risk outcomes.

3.5.1 Conclusion

Whilst the START appears to show good face value in creating a balanced risk assessment, this critique has demonstrated a potential lack of clinical utility of rating all items as both a strength

and a vulnerability. Additionally, although one might suggest that the Likert-type scale utilized by the START may reduce subjectivity, it may produce additional confusion as each score has (potentially) multiple meanings. Finally, there is a significant lack of research on the predictive validity of the START for risk outcomes other than violence risk makes it difficult to support the clinical utility of the START for all seven risk domains. Therefore, based on the findings of this critique, the value of the START is not supported by all aspects of what make a clinically useful risk assessment tool.

Rationale for Chapter Four

The START appears to be an effective tool for risk assessment, although the clinical utility of the strength scale may lie more in risk management and treatment planning than risk prediction. However much of the current literature around the START presented in chapters two and three utilise a male sample. Chapter four investigates the predictive validity of the START for risk incidents in females in levels of medium security.

Chapter Four

Risk incidents by females in medium-secure forensic Mental Health Services: Predictive validity of the Short-Term Assessment of Risk and Treatability (START)

4.0 Abstract

Purpose: To determine the predictive validity of the Short-Term Assessment of Risk and Treatability (START) within a female sample.

Background: The START is a structured professional judgement tool that assesses both vulnerabilities and strengths for 20 items pertaining to seven risk domains. Each risk domain is given a specific risk estimate.

Method: Utilising a pseudo-prospective design, START assessments were collated for 101 female adults in medium-secure forensic services. Incident data for the three months following the START assessment were coded in line with the seven risk domains of the START.

Results: AUC values revealed neither strength nor vulnerability scores predicted any outcome, with self-harm the only risk estimate to predict its relative outcome. There was no association between either strength or vulnerability scores and number of risk incidents. One-way ANOVA identified significant progressive between-group effects for self-harm risk estimates. Block-entry linear regression analyses revealed that the self-harm risk estimate demonstrated incremental validity above and beyond strength and vulnerability.

Conclusions: The START demonstrated clinical utility for self-harm only. Limitations and future considerations are also discussed.

4.1 Introduction

Within secure forensic services, individualised risk assessment is a critical process undertaken by mental health professionals for both the prediction of risk and treatment planning. Structured professional judgement (SPJ) tools have become the most widely used risk assessment tools within forensic services. These tools comprise empirically derived risk factors, subject to the flexibility of professional judgement relative to interpretation and scoring guidelines (Guy, Packer, & Warnken, 2012). These tools have been criticised in recent years for their focus on deficits in an individual, and their lack of appreciation for protective strengths that may help to reduce an individual's risk (Rogers, 2000).

Some researchers believe that a risk–only assessment is inherently inaccurate. Rogers (2000) argues that these assessments create a biased representation of an individual's risk and pessimism among mental health professionals, potentially resulting in negative consequences for the individual. For example, labeling forensic patients as aggressive is likely to incur restrictions on personal freedom, and without a consideration of the protective factors that stop them becoming aggressive an unbalanced view is created. The overreliance on risk factors does not facilitate a balanced evaluation of risk argues Rogers (2000). In response to such opinions, research is starting to focus on the impact of incorporating protective factors into this process to identify if the addition of such items creates a more accurate representation and subsequent treatment of individuals belonging to a forensic population.

The concept of protective factors in the literature is somewhat ambiguous. Serin, Chadwick, & Lloyd (2016) identify that there are commonly two frameworks proposing different mechanisms of how protective factors can influence recidivism. The first model states that risk is cumulative, and a risk factor increases individual risk, and a protective factor decreases individual risk (Serin et al., 2016). Therefore overall risk is an accumulation of both risk (additive) factors and protective (subtractive) factors (Serin et al., 2016). However this model does not incorporate the salience of a protective factor to the individual, and assumes that all risk and protective factors carry the same "weight" in relation to overall risk. This undermines the importance of a powerful protective factor (such as a high-quality relationship) compared to more salient risk factors such as a pro-criminal attitude (Serin et al., 2016).

The second framework suggests that the model is more interactive in nature, such that protective factors are only relevant in the presence of risk factors (Serin et al., 2016). This means that the impact of protective factors is reliant on the level of risk, such that risk is reduced if both risk and protective factors are highly present (Serin et al., 2016). Alternatively, when risk is already low, protective factors have little to no impact (Serin et al., 2016).

In addition to this research, an alternative school of thought suggests that a protective factor does not have to have a corresponding risk factor (Johnson & Morris, 2008). For example, research has demonstrated that religiosity can reduce the likelihood

of delinquency in youth, though non-religiosity is not a risk factor for this behaviour (Johnson & Morris, 2008). Protective factors are most commonly conceptualized as variables that serve to moderate the effect of existing risk factors on the likelihood of an adverse outcome (de Ruiter & Nicholls, 2011). Research has demonstrated that protective factors can have a buffering effect against the negative impact of risk factors and subsequently reduces the likelihood of reoffending in adolescents (Lodewijks, de Ruiter, & Doreleijers, 2010).

The uncritical acceptance of risk assessment in forensic practice and the need for the consideration of an individual's strengths has sparked the development of several tools which incorporate protective factors. One such tool is the Structured Assessment of Protective Factors for violence risk (SAPROF; de Vogel, de Ruiter, de Bouman, & de Vries Robbé, 2009). The SAPROF is one of a handful of tools to consider protective factors, providing a measure of protective strength, and the utility of this has been widely validated in the literature (de Vogel, de Vries Robbé, de Ruiter, & Bouman, 2011; de Vries Robbé, de Vogel, & de Spa, 2011). Other such tools include the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2006) and the Inventory of Offender Risk, Needs, and Strengths (IORNS; Miller, 2006). The SAVRY includes 24 risk factors, plus an additional six protective factors including prosocial involvement and resilient personality traits (Borum, Bartel, & Forth, 2006). The IORNS includes a 26-item protective strength index which assesses

cognitive/behavioural regulation, anger regulation, and education/training (Miller, 2006). All three of these tools have incorporated protective factors, however these factors are stand-alone items within these tools and act as an additional component of the assessment for consideration by the rater.

4.1.1 The Short-Term Assessment of Risk and Treatability

The Short Term Assessment of Risk and Treatability (START; Webster et al., 2009) addresses the criticisms of negative bias in SPJs. This SPJ tool looks at 20 dynamic factors in predicting adverse outcomes in the short-term (recommended duration, three months). The START allows the rater to take into account both risk and protective factors by rating each of the 20 items as both a strength and a vulnerability. The START also requires a Specific Risk Estimate (SRE) for each of the seven risk domains, which is derived from consideration of all 20-items of the tool. The risk domains of the START are violence risk to others, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, and victimisation. The authors of the START describe this tool as an *aide-mémoire* rather than a diagnostic test, and thus its primary use is to inform mental health status evaluations and treatment planning, as well as communicating risk-related information (Webster et al., 2009).

The START has been well validated for violence risk. Desmarais, Nicholls, Wilson, and Brink (2012) looked at the predictive validity of the START for inpatient aggression. Using a male sample, they found that the START was more predictive of

inpatient aggression than HCR-20 subscale scores or PCL-R total scores. This study demonstrated that the use of *both* risk and protective scores allowed for a more accurate prediction of inpatient aggression (Desmarais et al., 2012). These findings have been replicated across different settings, including different levels of security, demonstrating the clinical utility of the START (Desmarais, Nicholls, & Brink as cited in Webster et al., 2009). Research has also demonstrated good inter-rater assessor reliability between individuals and across disciplines, a vital feature of any risk assessment tool used in a clinical setting (Desmarais, Nicholls, Wilson, & Brink, 2012; Nicholls et al., 2006). Whilst the START is well validated for violence risk, the remaining risk domains of the START have received less focus, and subsequently are less validated in the current literature.

The majority of current literature has focused on the predictive validity of the START in relation to adult males, with females either making up a small proportion of the sample (usually around 10% or fewer), or being overlooked all together. Addressing this gap in the literature, Viljoen, Nicholls, Greaves, de Ruiter, and Brink (2011) investigated the ability of the START to predict successful reintegration into the community of female forensic psychiatric patients. The findings demonstrated that women who successfully returned to the community had higher START strength scores and lower vulnerability scores (Viljoen et al., 2011). Whilst this study identified the predictive validity of the START in relation to community reintegration, this study did not specifically look at

predictive validity in terms of adverse outcomes. Success in the community was operationalized as a lack of a readmission to the hospital in combination with an absolute discharge (Viljoen et al., 2011). Therefore these findings do not shed light onto the predictive validity of the tool for the seven risk domains the START specifically identifies.

O'Shea and Dickens (2015) investigated gender differences in the predictive validity of the START for aggression and self-harm. The findings suggest that the START is more predictive of these incidents in women than in men, with SREs predicting the largest effect size. These findings support the clinical utility of the START for these behaviours. However, only two of the risk domains are captured in this research, and thus further exploration of the use of the START with females is needed in relation to the remaining five risk domains.

4.1.2 Aims of Current Study

The majority of the START research has not only utilised a male sample, but has also focused on either violent/aggressive outcomes or incidents of self-harm, with the remaining five risk domains often being overlooked. Although these are arguably the outcomes of most concern to secure services, the remaining five risk domains of the START are not trivial. The current study examines all seven risk outcomes for adult females in medium-secure care.

The unique feature of the START is its consideration of risk areas as both a vulnerability and a strength. This study therefore aimed to extend the current knowledge of the START by:

i) examining the predictive validity of the START strength scores, vulnerability scores, and SREs for risk incidents over a three-month period,
ii) examining the relationship between START strength scores and vulnerability scores for number of risk incidents over a three-month period,
iii) examining if there is a significant difference between SRE rating and number of risk incidents, and
iv) identifying whether strength scores, vulnerability scores, or SREs have incremental validity over one another.

The following hypotheses will be tested during the present study:

i) START strength scores, vulnerability scores, and SRE's will significantly predict risk incidents over a three-month period.
ii) START strength scores will be negatively correlated with number of risk incidents, and vulnerability scores will be positively correlated with number of risk incidents.
iii) There will be a significant difference between SRE rating (high, medium, low) and number of risk incidents.
iv) SREs will demonstrate incremental validity over strength and vulnerability scores.

4.2 Method

4.2.1 Ethics

This study was approved by University of Nottingham School of Medicine Ethics Committee (ethics reference number: K17092015 SoM DFP PAP; see Appendix F). The study was also approved by the head of research within the healthcare company from which data was gathered. As this was deemed a service evaluation, and data were routinely collected within the service, it was not felt necessary to seek ethical approval from the National Health Service England. Additionally, hospital directors were contacted individually for consent, and were provided with a research proposal with details of the study (see Appendix G). Initially, five hospitals were contacted for participation, however one hospital was excluded for not providing consent within the allotted time constraint. In total, four hospital directors gave their consent to take part in the study.

4.2.2 Participants

Eligible participants were gathered from four medium secure female forensic hospitals across the United Kingdom. Current residents of these hospitals who had completed an initial START assessment and had been a resident of the hospital for at least three months were included in the study.

Patients were excluded from the study if they held a diagnosis of learning disability. There is limited research in the use of START with this population, and despite one study yielding moderate predictive validity, the small sample size (n=28) means that these results are deemed ungeneralisable outside of the study

itself (Inett, Wright, Roberts, & Sheeran, 2014). Therefore a decision was made to exclude this population from the current study. Additionally, in line with the START guidelines, patients were excluded if they had in excess of five missing strength or five missing vulnerability ratings (Webster et al., 2009).

4.2.3 Procedure

This study utilised a pseudo-prospective cohort design. Initial START assessments were gathered from existing file data following admission. These assessments were completed by various members of the multidisciplinary team instructed to follow the manual guidelines, however due to the nature of utilising pre-existing data the consistency of START raters and their following of the manual guidelines could not be guaranteed. All data were present at the beginning of the study, as this data was routinely collected by the service. As this was pre-existing data, raters were blind to the purpose of the study, as the START assessments had already been completed prior to the commencement of the current study.

Additionally, demographic and clinical information, including age and ethnicity were gathered from patient records. Risk incidents were extracted from the incident recording system for the three months following the initial START assessment. All information was anonymised and linked using a unique coding number.

4.2.3.1 START assessment The START is a 20-item risk assessment based on structured clinical judgement (Webster et al., 2009). Each item is dynamic and scored on two scales: vulnerabilities and strengths. These items are rated on a three-point

scale (0: *Minimal,* 1: *Moderate,* and 2: *Maximal*). Additionally, risk rating on seven domains (violence risk to others, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, and victimisation by others) are given a Specific Risk Estimate (SRE) as *Low, Moderate,* or *High*. This risk is not derived from a total score, but rather a consideration of the 20-items.

Under the institutional service policy, the START is nursing-lead, and is completed by the multidisciplinary team where a consensus is reached. This is concordant with the START guidelines which state that the START is intended to be completed by a number of mental health specialists in order to incorporate a range of opinions from a variety of disciplines (Webster et al., 2009). As this study is looking at the clinical application of the START, and these assessments are completed by several different multidisciplinary teams across the hospitals, it was not possible to calculate inter-rater reliability here. However, previous research has demonstrated that the START has high inter-rater reliability (ICC's range from .81 to .92) in a variety of clinical settings, including secure mental health services (Braithwaite et al., 2010; Desmarais, Nicholls, Wilson, & Brink, 2012; Nicholls, Brink, Desmarais, Webster, & Martin, 2006).

All staff are required to attend START training as standard. This is a three-hour training session delivered by psychologists within the service who are experienced in utilising the START. This means staff are adequately trained above the recommendations from the START manual (which states that the START should be

completed by an experienced clinician, and specific training is not necessarily required; Webster et al., 2009). However, as START raters were anonymous, it could not be identified as to the level of experience of raters, potentially leaving this requirement of 'experienced clinician' unmet.

The START is recommended for completion every three months, which is reflected in the current service application of the tool (Webster et al., 2009). In line with the START manual, individuals were excluded if they had five or more missing items from the START (Webster et al., 2009). Individuals with four or fewer missing items were prorated using the following formula as recommended when using the START for research purposes (Webster et al., 2009):

[(raw total score/40) x number of missing items] + raw total score

It is of note that "raw total score" refers to total strength or vulnerability score, not combined total score, such that each subset is prorated individually.

The START also allows the rater to identify 'Key Items' and 'Critical Items' that seem particularly important in the development of treatment plans (Webster et al., 2009). As these items are more relevant to treatment planning rather than risk prediction, 'Key Items' and 'Critical Items' were not considered in the present study.

4.2.3.2 Demographic and clinical information Data were gathered regarding patient age, gender, ethnicity and Mental Health Act (1983, 2007) section from existing file data.

4.2.3.3 Risk incidents For the purpose of this study, risk incidents were characterised according to the START definitions of each of the seven risk domains. Violence was defined as any actual, attempted, or threatened harm to others or self. Self-harm was defined as an act causing superficial, moderate, or severe damage to self. Suicide was defined as either the completed act, or an incident where suicidal intent was explicitly expressed by the individual. Unauthorised leave was defined as absconding from hospital, either leaving the grounds or secure area without approval, or beaching terms of leave. Substance abuse was defined as use of substances resulting in failure to fulfil expectations or obligations. Self-neglect was defined in terms of diet, shelter, medical care, and personal hygiene. Victimisation was defined as victimisation of any type, including neglect and violence.

Service policy requires all incidents be recorded using the Incident Recording Information System (IRIS). These incidents are recorded by a member of the clinical team who was present at the time of incident. The current study was interested in the clinical application of the START risk assessment for all seven risk domains of the instrument. Risk incident data was extracted from IRIS for the three-month period following the initial START assessment, based on the recommended time frame in the START manual (Webster et al., 2009). These risk incidents were coded in line with the seven risk domains. Those incidents which did not fall into any of the seven risk categories of the START were excluded from analyses. These included medication errors made by staff, or

security breaches that were recorded in relation to the hospital as a whole rather than an individual patient. A measure of inter-rater reliability for coded risk incidents was conducted to ensure reliable coding.

4.2.4 Data Analysis

Power calculations were conducted for correlations, ANOVA, and linear regression, based on the recommendations from Cohen (1992), with α=.05, and β=.8. Linear regression produced the largest required sample size from these calculations, thus these results have been reported in order to provide sufficient power for all type of analyses. The results suggest that 114 participants were needed to detect a small effect size (r=.1), 41 participants were needed to detect a medium effect size (r=.3), and 27 participants were needed to detect a large effect size (r=.5). Descriptive statistics were calculated to identify characteristics of the sample, distribution of START strength and vulnerability scores, SRE scores, and risk incidents.

The predictive validity of the START strength scores, vulnerability scores, and SRE's was examined using area under the curve (AUC) values using Receiver Operating Characteristic (ROC) analyses. This was applied for each type of incident relating to the seven risk outcomes. AUC values range from 0 to 1, where .5 represents chance (Dolan & Doyle, 2000). Rice and Harris (2005) report AUC values of .556, .638, and .714 respectively are equivalent to Cohen's *d* values of .1 (small), .5 (moderate), and .8 (large), which are the most often used as measures of effect size.

ROC analysis was used as it is relatively unaffected by base rates, and it is not dependent upon how the risk assessment is coded (Mossman, 1994; Pepe et al., 2004). Additionally, the majority of the existing literature examining predictive validity of the START utilises this analysis, facilitating easy comparison.

Correlations were used to determine the relationship between strength scores and vulnerability scores to number of incidents that occurred over the following three-month period for each of the seven risk domains. Coefficients of .10, .30, and .50 are deemed small, moderate, and large respectively (Cohen, 1992).

A one-way analysis of variance (ANOVA) was used to determine if there was a significant difference between number of risk incidents for high, medium, and low SREs. Each SRE was matched to its relative risk incident. Levene's statistic was used to identify homogeneity of variance. When this assumption was violated, the Welch statistic was used (Field, 2013). Similarly, post-hoc analyses in these cases utilised the Games-Howell Procedure, as this does not rely on the assumption of equal variances (Field, 2013).

Block-entry linear regression analyses were conducted in order to determine incremental validity of strength scores and vulnerability scores over each other when predicting number of risk incidents. Additionally, SRE's were included to identify incremental validity of these ratings above strength and vulnerability scores. For the purpose of these analyses SRE data were treated as ordinal numerical as it was considered reasonable to assume the distances

between each rating were equal and meaningful. Model one saw the total vulnerability scores entered into Step 1 and total strength scores entered into Step 2. Model two saw this process reversed in order to identify if vulnerability scores held incremental validity over and above total strength scores. Model three saw vulnerability and strength scores entered into Step 1 and SRE's entered into Step 2. Significant changes in *F* values demonstrate a statistically significant improvement in the number of risk incidents explained by the model (Field, 2013).

In order to account for multiple comparisons, and the subsequent increased risk of making a Type 1 error, *p* values were adjusted using Bonferroni corrections, such that the original alpha value was divided by the number of analyses on the dependent variable (Field, 2013). This resulted in an adjusted *p* value of .007 for each hypothesis, with the exception of hypothesis three, where fewer outcomes were included, therefore resulting in an adjusted *p* value of .01.

4.3 Results

4.3.1 Inter-Rater Reliability

To ensure accurate categorisation of risk incidents, an independent rater was used to rate 10% of the risk incidents. An inter-rater reliability analysis was conducted using Cohen's Kappa statistic to determine consistency between raters of this categorical data. In line with the recommendations from Landis and Koch (1977) substantial agreement was found between raters, with Kappa=.72 ($p<.001$).

4.3.2 Participants

Of the 112 participants initially recruited, 11 were excluded due to missing START data, leaving a final sample of 101. Of these remaining participants, 25 (25%) had prorated strength scores, and 26 (26%) had prorated vulnerability scores. The mean age of the participants was 34.3 years (*SD*=11.9). The majority of the sample were White British (*n*=82, 81%), 10 (10%) were White Other, 2 (2%) had no recorded ethnicity, 2 (2%) were Black British Caribbean, 2 (2%) were Mixed White and Black African, 1 (1%) was Black British African, 1 (1%) was Asian British Indian, and 1 (1%) was White Irish. All patients were detained under the Mental Health Act (1983, 2007).

4.3.3 START Scores and SREs

Total strength scores ranged from one to 36 (*M*=21.67, *SD*=7.53) and the total vulnerability scores ranged from zero to 38 (*M*=23.88, *SD*=6.79). SRE frequencies are reported in Table 4.1.

Table 4.1

Specific Risk Estimate Frequencies by START Rating

	Low		Medium		High	
Outcome	**n**	**%**	**n**	**%**	**n**	**%**
Violence	36	35.6	45	44.6	20	19.8
Self-harm	41	40.6	47	46.5	13	12.9
Suicide	61	60.4	33	32.7	7	6.9
Unauthorised Leave	72	71.3	24	23.8	5	5.0
Substance Abuse	80	79.2	16	15.8	5	5.0
Self-neglect	57	56.4	33	32.7	11	10.9
Victimisation	42	41.6	51	50.5	8	7.9

4.3.4 Incident Data

The most common risk incidents were of violence, which 62 (61%) of the sample engaged in. Self-harm was evident in 52 (52%) of the sample. Victimisation was evident in 26 (26%) of the sample. Self-neglect was evident in 18 (18%) of the sample. Suicide was evident in six (6%) of the sample. Unauthorised leave was evident in six (6%) of the sample. The least common risk incident was substance abuse, which just three (3%) of the sample were identified for.

In total across the sample there were 566 incidents of violence, 446 incidents of self-harm, 11 incidents of suicide, ten incidents of unauthorised leave, three incidents of substance abuse, 23 incidents of self-neglect, and 52 incidents of victimisation.

4.3.5 Predictive Validity

The AUC values resulting from ROC analysis are presented in Table 4.2. Neither strength nor vulnerability scores predicted any of the seven risk outcomes. Using the adjusted *p* value, the only SRE that was predictive of its relative outcome was self-harm, AUC=.71, $p<.001$, demonstrating a large effect size.

4.3.6 START Scores and Number of Risk Incidents

The correlations between strength scores, vulnerability scores, and number of risk outcomes are presented in Table 4.3. As the data were not normally distributed (skewness and kurtosis >1.96) Spearman's rho correlations are reported (Field, 2013). Using the adjusted *p* value, neither strength nor vulnerability scores were significantly correlated with any of the risk incidents.

Table 4.2

Predictive Validity of START Scores and SREs

	Strength Score			Vulnerability Score			SRE		
Outcome	**AUC**	***p***	**95% CI**	**AUC**	***p***	**95% CI**	**AUC**	***p***	**95% CI**
Violence	.50	.97	[.38, .61]	.55	.39	[.44, .67]	.56	.30	[.45, .68]
Self-harm	.48	.73	[.37, .60]	.55	.37	[.44, .67]	.71	<.001*	[.61, .82]
Suicide	.37	.30	[.21, .54]	.47	.80	[.28, .66]	.65	.21	[.42, .89]
Unauthorised leave	.70	.10	[.46, .94]	.53	.82	[.29, .77]	.57	.59	[.29, .84]
Substance abuse	.61	.52	[.34, .88]	.84	.05	[.77, .91]	.56	.74	[.22, .89]
Self-neglect	.39	.13	[.25, .52]	.64	.06	[.49, .79]	.60	.21	[.44, .75]
Victimisation	.43	.31	[.30, .56]	.63	.06	[.50, .75]	.55	.42	[.42, .68]

Note. SRE, Specific Risk Estimate; AUC, Area Under the Curve; CI, Confidence Interval

*significant at the <.001 level

Table 4.3

Spearman's rho Correlations between START Scores and Number of Risk Incidents

	Strength Score		**Vulnerability Score**	
Outcome	r_s	***p***	r_s	***p***
Violence	-.08	.22	.19	.03
Self-harm	-0.5	.33	.12	.12
Suicide	-.11	.15	-.03	.40
Unauthorised leave	.16	.05	.03	.40
Substance abuse	.06	.26	.20	.02
Self-neglect	-.16	.06	.20	.03
Victimisation	-.08	.20	.18	.03

Note. *significant at the <.001 level (1-tailed)

4.3.7 SREs and Number of Risk Incidents

The results of the one-way ANOVAs are displayed in Table 4.4.

Table 4.4

One-way ANOVA for the Significance of SRE rating and number of Risk Incidents

Outcome	***F***	**(df)**	***p***
Violence*	3.24	(2, 42.54)	.05
Self-harm*	15.95	(2, 27.47)	<.001
Suicide*	0.98	(2, 14.05)	.40
Self-neglect*	1.18	(2, 24.39)	.33
Victimisation	0.25	(2, 98)	.78

Note. *Levene's statistic was significant (p<.05), therefore the Welch statistic is reported

Unauthorised leave and substance abuse were excluded from these analyses as there were not enough data to make comparisons (n=6 and n=3, respectively). Using the adjusted p value, significant between-group (high, medium, and low) effects were found for self-harm (p<.001).

Post-hoc tests revealed a significant difference between ratings of low risk of self-harm (M=0.34, SD=0.62) and ratings of medium risk of self-harm (M=1.84, SD=1.89) in relation to number of self-harm risk incidents (p<.001), suggesting those rated as "medium risk" engaged in significantly more self-harm incidents than those rated as "low risk". A significant difference was also found between ratings of low risk of self-harm and ratings of high

risk of self-harm (M=1.95, SD=2.07) in relation to number of self-harm incidents (p=.04), suggesting those rated as "high risk" engaged in significantly more self-harm risk incidents than those rated as "low risk". No statistically significant difference was found between rating of medium and high risk of self-harm and number of self-harm risk incidents (p=.64).

4.3.8 Incremental Validity

The results of the block-entry linear regression analyses are presented in Table 4.5. The only significant analysis with the adjusted p value was for self-harm, where the SRE had incremental validity over strength and vulnerability scores (p<.001).

Table 4.5

Block-entry Linear Regression Analyses

	R^2	R^2 Change	*F* Change	*p*
Violence				
Model 1				
Vulnerability Score	.05	.05	4.72	.03
Strength Score	.05	.00	.00	1.00
Model 2				
Strength Score	.01	.01	.61	.44
Vulnerability Score	.05	.04	4.04	.05
Model 3				
Strength & Vulnerability Score	.05	.05	2.33	.10
SRE	.10	.06	5.96	.02
Self-Harm				

Model 1				
Vulnerability Score	.02	.02	2.26	.14
Strength Score	.03	.01	.50	.48
Model 2				
Strength Score	.00	.00	.01	.92
Vulnerability Score	.03	.03	2.74	.10
Model 3				
Strength & Vulnerability Score	.03	.03	1.37	.26
SRE	.18	.15	17.91	<.001*
Suicide				
Model 1				
Vulnerability Score	.00	.00	.00	.95
Strength Score	.01	.01	.96	.33
Model 2				

Strength Score	.01	.01	.79	.38
Vulnerability Score	.01	.00	.18	.68
Model 3				
Strength & Vulnerability Score	.01	.01	.48	.62
SRE	.05	.04	4.45	.04
Unauthorised Leave				
Model 1				
Vulnerability Score	.00	.00	.25	.62
Strength Score	.04	.03	3.27	.07
Model 2				
Strength Score	.02	.02	2.23	.14
Vulnerability Score	.04	.01	1.29	.26
Model 3				
Strength & Vulnerability Score	.04	.04	1.76	.18

SRE	.08	.05	5.19	.03
Substance Abuse				
Model 1				
Vulnerability Score	.03	.03	2.84	.10
Strength Score	.04	.02	1.60	.21
Model 2				
Strength Score	.00	.00	.29	.59
Vulnerability Score	.04	.04	4.16	.04
Model 3				
Strength & Vulnerability Score	.04	.04	2.23	.11
SRE	.04	.00	.03	.86
Self-Neglect				
Model 1				
Vulnerability Score	.03	.03	2.85	.10

Strength Score	.04	.01	.82	.37
Model 2				
Strength Score	.02	.02	2.13	.15
Vulnerability Score	.04	.02	1.52	.22
Model 3				
Strength & Vulnerability Score	.04	.04	1.83	.17
SRE	.05	.02	1.56	.22
Victimisation				
Model 1				
Vulnerability Score	.02	.02	2.21	.14
Strength Score	.02	.00	.00	.95
Model 2				
Strength Score	.00	.00	.36	.55
Vulnerability Score	.02	.02	1.83	.18

Model 3				
Strength & Vulnerability Score	.02	.02	1.09	.34
SRE	.02	.00	.10	.76

Note. * significant at the <.001 level

4.4 Discussion

At initial viewing, the results of the current study question the lack of clinical utility of the START with this population. The range in both strength and vulnerability scores suggested that raters have the ability to differentiate between those who exhibit more or less strengths and risks. This is worth mentioning as all of the participants were residing in medium levels of security, so the fact that raters of the START could still distinguish between those more or less likely to engage in risk behaviours supports the clinical utility of the START assessment.

The results of the current study throw into question the clinical utility of the START strength scale, as total strength scores did not significantly predict any of the risk outcomes. Similarly, these scores were not associated with number of risk outcomes. This effect may be understandable given the nature of forensic risk assessments. Within forensic services clinicians are far more experienced in rating the risk factors of an individual, and spend much less time looking at the protective factors. A study by Collins, Desmarais, Nicholls, and Brink (cited in Webster et al., 2009) looked at the satisfaction of clinicians using the START in clinical practice. They found that while 88% of their sample agreed that, overall, the START was user-friendly, only 56% agreed that it was easier to make the finer distinctions in rating the independent scales (i.e. deciding between 0 vs. 1 vs. 2). It may be that strength scores did not significantly predict the risk outcomes in the present study

because raters were less confident in rating the items as a strength, and struggled to accurately distinguish between ratings.

The START vulnerability scores also did not significantly predict any of the seven risk outcomes. Similarly vulnerability scores were not associated with number of risk incidents. Whilst previous literature with females is minimal, O'Shea and Dickens (2015) investigated gender differences in the predictive ability of the START, identifying that the START was a stronger predictor of both aggression and self-harm in women than in men. These findings were not replicated in the current study. This may be a reflection on the quality of the START data within the present study. It was not possible to control raters of the START, and thus this inconsistency, whilst reflective of the practical use of the START within a clinical setting, may have impacted on the quality of the data and subsequent predictive validity. Conversely, the lack of predictive validity may also reflect the effective management of violence and self-harm risk within the medium secure setting, such that an individual may present as high risk of violence, however effective management strategies mean this is not reflected in number of risk incidents.

The SRE data revealed that self-harm was the only significant predictor of its respective outcome. Similarly, the self-harm SRE demonstrated significant between-group differences. This suggests that staff were able to effectively differentiate between risk levels for individuals. It may be that this was the only differences found due to the frequency of this behavioural event. Self-harm is one of

the most common risk incidents within secure services (Woods & Ashley, 2007). Therefore this is where the majority of the data, and corresponding literature, is focused. As both the frequency and variance amongst individuals engaging in self-harm are higher, staff will unsurprisingly be better at determining risk in relation to these incidents.

It is important to consider the impact of the type of sample utilized in the current study. It may not be unsurprising that self-harm was found to be the only significant predictor of risk. This is a common occurrence in women with a diagnosis of Borderline Personality Disorder, therefore the frequency of this behaviour is likely to be high within a medium-secure mental health setting. Subsequently, the use of this population may have inadvertently skewed the findings of the current study. It may be that as this is a frequent problem within this setting, staff are more attuned to it, and so are more accurately able to rate this behaviour in terms of risk prediction. It may also be the case that other risk behaviours, such as violence and aggression, may be artifacts of self-harm in an attempt to manage this behaviour. Therefore, the underwhelming results of the START may not be a reflection of the tool in terms of predictive accuracy, but a reflection of the primary focus of staff within this setting and the risk behaviours displayed by the population in question.

In terms of incremental validity, for the all of the seven risk domains, strength and vulnerability scores did not uniquely contribute to number of risk incidents. However the SRE

demonstrated incremental validity above and beyond the strength and vulnerability scores for self-harm. This finding would suggest that in terms of self-harming behaviours, the most clinically useful part of the START is the SRE information which operationalizes the behaviour. This is derived from a consideration of the strength and vulnerability scale, and in the clinical use of the START, the strength and vulnerability scales are not scored as they have been for research purposes. Therefore it is the SREs that will be most likely used in the formulation of risk management plans. The results of this study suggest that for incidents of self-harm, the SRE offers a unique contribution to the prediction of risk, therefore supporting its use in the development of risk management plans over more general indices from the START.

4.4.1 Limitations and Future Considerations

A criticism of the current study is the lack of control over START raters. Whilst the incident data were coded independently and checked for inter-rater reliability, START data was collated from pre-existing records. This means there was no way of checking inter-rater reliability or even knowing how many independent START raters were included. Whilst this appears to be a methodological flaw, this is the reality of the practical use of the START; each individual will have a variety of raters across time. Therefore the findings of the present study are reflective of the true clinical utility of the START within secure mental health services.

A similar criticism could also be the suitability of the people who were collecting the data. Whilst all staff members are required

to attend one START training when they join the service, staff were not specifically trained for this study. It may be that some raters of the START were very recently trained, and thus adhere to the coding guidelines more closely than those who received their training some time ago. It was hoped that the staff training would supersede the need for 'experienced clinicians', however the varying time between training and data collection may have impacted on the quality and consistency of the data gathered. It may be that future research selects specific members of staff to be trained on the START and immediately collect data in order to reduce the variety in "quality" of raters, as this could have impacted on the present findings.

Whilst the overall sample size for the current study was satisfactory in terms of power, it may be of benefit to recruit a larger sample for future research. This is because suicide, substance abuse, and unauthorised leave were vastly underrepresented in the current sample. Whilst this is good news clinically, statistically it means that any significant findings will show large confidence intervals, and thus should be taken as provisional findings only. This may explain the current gap in the literature regarding these outcome measures, as the security in place in forensic hospitals reduces the likelihood that these incidents will take place. Future research should consider recruiting on a larger scale in order to get some definitive findings regarding the predictive validity of the START in relation to these lesser-observed outcomes.

Similarly, an issue with the current study was the method of risk incident collection. The pseudo-prospective design of the study meant that as opposed to collecting incident data with the START domains in mind, pre-existing data was gathered through the standard incident recording system and re-coded to fit the START criteria. This may present some biases as to what data will be gathered. For example, self-neglect may not often be recorded as a risk incident if it is regarded as a daily occurrence such as lack of personal hygiene. This means the data collected may not have represented the true number of risk incidents for each individual. This would be overcome if future research adopted a true prospective design and coded observed behaviour as risk outcomes relating to the seven START risk domains.

Underreporting of incidents may also have contributed to the minimal findings in the present study. This issue has been discussed in the literature, with reasons for underreporting including time restraints, insufficient processes, a lack of knowledge or specific beliefs about risk, cultural norms, and a perceived lack of value in the process (Kingston, Evans, Smith, & Berry, 2004). Potential underreporting will have impacted on the accuracy of the results, as the true number of risk incidents may be underrepresented, subsequently impacting on the predictive validity of the START. Training staff on the importance and value of reporting risk incidents coupled with a prospective design would overcome this issue, and allow a more accurate representation of risk incidents.

The follow-up period for this study was identified as three months. This is in line with the START manual recommendations. However, previous literature has suggested that the START increases in predictive validity from six-month to 12-month follow-up periods (Viljoen et al., 2016). It may be that limiting the follow-up period to three-months could mean certain infrequent risk behaviours are unlikely to be seen, and the START items may not be sensitive enough to reflect this over such a short period of time, therefore creating false conclusions about the predictive accuracy of this tool.

From a statistical perspective, during the regression analyses the SRE data were treated as ordinal numerical data in the same respect a Likert scale would be treated. This comes with its own limitations, as this means that an assumption has been made that one unit change between each rating is equal. This assumption could impact on the findings of this analysis, and thus limit the conclusions that can be made.

Though missing START data was dealt with in accordance with the manual protocol, the existence of missing information and the reasons therefore reflect negatively on the staff's familiarity with their patients, and thus the validity of their clinical opinions. This has an impact on the confidence of the findings of the present study, and subsequent generalisability of conclusions made. Future research should ponder this reflection, and consider rather than pro-rating START scores as suggested in the manual, the data be excluded from the study.

4.4.2 Conclusion

Similarly to previous literature on the START, the findings of the current study are stronger for certain outcomes than others. Whilst the clinical utility of the START has been demonstrated within this female sample for self-harming behaviours, this is not the case for the other six risk outcomes, and it may be that more information needs to be gathered regarding these risk domains. The strongest aspect of the START in the current study appears to be the SRE.

Though initial viewing of the results of the current study may suggest that the START is not a clinically useful risk prediction tool for outcomes other than self-harm, this may be an inaccurate conclusion. Given the population used, it is important to note the frequency of self-harming behaviours, the primary focus of staff in terms of risk management, and the fact that other risk behaviours may be more accurately labelled as artefacts of self-harm within this population. So while the results appear to question the clinically utility of the START for females within levels of medium security, a larger and more varied sample could identify if this is a valid conclusion, or if the lack of predictive validity is specific to this population.

Rationale for Chapter Five

The research presented thus far in this thesis has pertained to the role of protective factors in the assessment of risk. However how does this assessment help to inform the management and treatment of risk behaviours? Chapter five attempts to answer this question by utilising a Coping Skills intervention in order to promote protective factors to subsequently reduce self-harming behaviours in a 20-year-old female with a diagnosis of Borderline Personality Disorder.

Chapter Five

Reducing Risk Behaviours by Promoting Protective Factors: A Case Study on the Effectiveness of a Coping Skills programme for a Female with Borderline Personality Disorder and Self-Harming Behaviours

5.0 Abstract

Purpose: The aim of this case study was to demonstrate the benefit of increasing protective factors to reduce an adverse outcome. In this case, promoting coping skills in order to reduce self-harming behaviours.

Background: Miss X is a 20 year-old female with a diagnosis of borderline personality disorder. Miss X's main risk areas and difficulties include physical and verbal aggression, self-harm, suicidal ideation, and difficulty managing emotions.

Method: The intervention comprised of a ten-session Coping Skills programme, combining both DBT and CBT techniques. This programme also included out of session work to be completed each week to support the skills learnt within the group. Pre and post assessments were collected, in addition to a six-month follow up assessment.

Results: Whilst significant differences were not found in the psychometric data from pre to post, levels of hopelessness did reduce at the six-month follow-up. Similarly, incidents of self-harm also reduced at the six-month follow-up.

Conclusions: The behavioural data from this case study yielded potentially promising results. By promoting protective factors in the form of adaptive and appropriate coping strategies, this intervention resulted in a reduction in self-harming behaviours in this individual, thus supporting the utility of promoting protective factors to reduce risk behaviours.

5.1 Introduction

Many individuals with personality disorder have maladaptive dysfunctional coping mechanisms. Research conducted by Bitjttebier and Vertommen (1999) has evidenced that psychiatric in-patients frequently fail to access social support and use excessive amounts of avoidant coping. Women in secure services, particularly those diagnosed with borderline personality disorder often have poor strategies to regulate their emotions, and these poor strategies include deliberate self-harm (Brown et al, 2002).

In mental health care settings the prevalence of borderline personality disorder is high, with some research reporting figures of 50% or greater (National Collaborating Centre for Mental Health, 2009). Whilst within the community the proportion of males and females with a diagnosis of borderline personality disorder is roughly equal, within services there is a strong preponderance of women (National Collaborating Centre for Mental Health, 2009).

The diagnostic criteria of borderline personality disorder as outlined by the Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-V; American Psychiatric Association, 2013) characterises these individuals as demonstrating a desperate avoidance of real or imagined abandonment, instability in interpersonal relationships, identity disturbances, marked impulsivity in "reckless" or self-damaging areas, suicidality and self-harm, difficulty controlling anger, and instability of mood. This symptomatology highlights the importance of promoting appropriate

coping mechanisms within this population, and an ability to understand and better regulate their emotions.

Self-harm behaviours are one of the most commonly associated symptoms of borderline personality disorder. Self-harming behaviour can serve a variety of functions for an individual with borderline personality disorder, including relief from acute distress or feelings in addition to reconnecting after a period of dissociation (National Collaborating Centre for Mental Health, 2009).

Research has suggested that individuals with borderline personality disorder use deliberate, non-suicidal self-harm as an experiential avoidance mechanism (Chapman, Specht, & Cellucci, 2005). This may reflect an inability to cope with negative or difficult emotions. In this respect self-harm acts as a maladaptive coping mechanism. Therefore, to reduce self-harm in an individual with borderline personality disorder, it is necessary to promote protective factors in the form of appropriate coping mechanisms. The aim of the current case study is to evaluate the effectiveness of a Coping Skills programme in promoting positive coping strategies and subsequently reducing incidents of self-harm.

5.1.1 Client Overview

Miss X is a 20 year-old female currently residing in a high dependency locked rehabilitation unit. Miss X holds a diagnosis of borderline personality disorder. Currently detained under Section 3 of the Mental Health Act (1983, 2007), Miss X's main risk areas and difficulties include physical and verbal aggression, self-harm, suicidal ideation, and difficulty managing emotions.

5.1.2 Assessment and Analysis

The following assessments were administered to inform Miss X's formulation and to assess change from pre to post-programme (see Appendix H for consent form):

(i) Short-Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Desmarais, 2009):

The START is a 20-item risk assessment tool that looks at seven risk domains. These include violence risk to others, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, and victimisation by others. Research has demonstrated the utility of using the START in secure services, with the START demonstrating excellent inter-rater reliability (*ICC* range from .85 to .95, $p<.001$; Desmarais, Nicholls, Wilson, & Brink, 2012). The START also demonstrates good construct validity across different security levels, making it an appropriate tool for use in the present case study (Desmarais, Nicholls and Brink as cited in Webster et al., 2009).

Miss X's START assessment was completed by her primary nurse, with an MDT review and input. This initial START assessment showed that Miss X's greatest vulnerabilities were her relationships with others (including peers and staff), occupational skills such as understanding the value of education and work opportunities for her, emotional state/resilience, substance use, impulse control, and coping.

Miss X also demonstrated some strengths, for example self-care (including personal hygiene, appropriate dress, and normal eating), in addition to a grasp of material resources competencies (including appropriate management of finances).

Miss X's specific risk estimates for the seven risk domains of the START are displayed in Table 5.1.

Table 5.1

START Specific Risk Estimates

Specific Risk Estimate	**Rating**
Violence risk to others	Medium
Self-harm	High
Suicide	Medium
Unauthorised leave	Medium
Substance abuse	Medium
Self-neglect	Medium
Victimisation	Medium

(ii) Historical, Clinical, Risk Assessment (HCR-20 (v3); Douglas, Hart, Webster, & Belfrage, 2013):

The HCR-20 (v3) is a structured professional judgement tool designed to assess violence risk. In turn this helps to inform risk management plans by identifying key areas of vulnerability for an individual (Douglas et al., 2013). There are 20 items assessed on a three point scale ('yes' indicating the item is present; 'partial' indicating the item is possibly present or present to a limited extent; and 'no' indicates the item is absent). There are ten

historical factors, five clinical factors, and five risk management factors incorporated into the HCR-20 (v3).

Miss X's HCR-20 was completed using a combination of file review, semi-structured interview, and MDT discussion.

The key Historical items for Miss X were violence and personality disorder. Miss X has demonstrated violence within the home, directed towards her sister. She has also demonstrated violence and aggression towards both staff and patients in response to an inability to manage her emotions. Miss X also holds a diagnosis of borderline personality disorder.

The key Clinical items of the HCR-20 (v3) were insight and instability. Miss X has little insight into her diagnosis and how this may impact on her behaviour. Miss X struggles to identify difficult emotions and will use self-harming as her primary coping mechanism. In reference to instability, Miss X engages in verbal aggression towards staff members if she feels are needs are not being adequately met, or she feels treated unfairly on the ward. This can also lead to self-harming behaviours.

In terms of the Risk items, Miss X's main areas for concern were pertaining to personal support and stress/coping. Miss X has no contact with her family except for her father, although this is intermittent and unreliable. Miss X has repeatedly demonstrated an inability to cope under stressful situations. Miss X struggles to identify what emotions she finds difficult to manage, and reverts to maladaptive coping mechanisms such as self-harm.

(iii) Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974):

The BHS consists of 20 true-false statements that are designed to assess negative expectancies about the future (pessimism). Scores are summed for a total score ranging from 0 to 20, where higher scores indicate a greater level of hopelessness. The BHS demonstrates excellent test-retest reliability, r=.69; p<.001, and internal consistency with Kuder-Richardson-20 reliabilities ranging from .87 to .93 (Beck & Steer, 1993). The BHS also displays good content validity, adhering closely to Stotland's (1969) understanding of hopelessness. Additionally, concurrent validity for the BHS has also been found to be good, with the BHS demonstrated good correlations with clinical ratings of hopelessness, r=.74; p<.005 (Beck & Steer, 1993). As a result the BHS was deemed an appropriate measure for this case study.

Miss X's responses pre-programme resulted in a score of 20. This indicates that she displayed a severe level of hopelessness, and therefore puts her at a very high risk of suicide. It is hypothesised that by increasing Miss X's protective factors through developing effective coping strategies, her hopelessness scores will decrease as she has more skills to apply in situations where she may have previously found it difficult to cope.

(iv) Coping Styles Questionnaire (CSQ; Rogers & Najarian, 1989):

The CSQ is a 48-item self-report measure to identify how the respondent copes with stressful situations. This measure identifies

both adaptive and maladaptive coping styles. The CSQ has four subscales: rational coping, emotional coping, avoidance coping, and detachment. "Rational coping" and "detachment" represent adaptive coping strategies. A high score on "rational coping" indicates that in a stressful situation, the individual will adopt a rational coping strategy. A high score on "detachment" indicates that in a stressful situation, the individual will feel independent of both the situation and the associated emotions. "Emotional coping" and "avoidance coping" represent maladaptive coping strategies. A high score on "emotional coping" indicates that in a stressful situation, the individual will not be able to cope emotionally. A high score on "avoidance coping" indicates that an individual will seek to avoid a stressful situation.

These four subscales demonstrate satisfactory internal consistency, with coefficients of .85, .74, .69, and .99 respectively (Roger, Jarvis, & Najarian, 1993). Substantial test-retest reliability has also been demonstrated across these four subscales, ranging from .70 to .80 (Roger, Jarvis, & Najarian, 1993).

Miss X's pre-intervention scores indicated higher levels of maladaptive coping mechanisms, with her highest rating being for emotional coping. This indicates that Miss X finds it difficult to cope emotionally when faced with stressful situations. It was hypothesised that by increasing her awareness of difficult emotions and introducing new adaptive coping strategies, Miss X's scores for maladaptive coping strategies would significantly decrease, and her scores for adaptive coping strategies would significantly increase.

Behavioural data was also gathered in the form of incident data. All incidents of self-harm are recorded by ward staff, providing a measure of frequency of self-harming incidents. The month prior to the programme began Miss X displayed a total eleven incidents of self-harming behaviours. This data was gathered again both post-programme and at a six-month follow-up to see if any change in psychometrics was also reflected in behavioural data.

5.1.3 Formulation

A formulation draws upon psychological theory to generate a working hypothesis of why an individual engages in behaviours or demonstrates difficulties in certain areas (Johnstone & Dallos, 2014). Information from relationships, social contexts, and personal history are incorporated to explain what lead to and maintains the "problem" behaviour. A formulation should be collaborative between the multi-disciplinary team and the client, and is used to inform appropriate treatments and interventions for the individual (Johnstone & Dallos, 2014). A formulation is open to interpretation, and should be reviewed regularly. With this in mind, Miss X was invited to engage in a joint formulation process, based on both the aforementioned assessments and semi-structured interview.

The CBT-based 5 P's approach was used to structure the formulation. This approach acknowledges the presenting issues, and predisposing, precipitating, perpetuating, and protective factors (Johnstone & Dallos, 2014). These are further outlined in Table 5.2.

Table 5.2

The Five P's of CBT Formulation

The Five P's	**Relationship to Therapy**
Presenting issues. Statement of the client's presenting problems in terms of emotions, thoughts, and behaviours.	This process begins to define the current problems the person faces. This introduces specificity and individualisation. This process helps to develop the therapeutic relationship, clarifies problems, and instils hope.
Predisposing factors. The distal external and internal factors that increased the person's vulnerability to their current problems.	Provides a longitudinal understanding of the problems and a focus for more in-depth interventions that aim to maintain change and prevent relapse.
Precipitating factors. The proximal external and internal factors hat triggered the current presenting issues.	Introduces the cognitive model and provides initial focus for interventions. Aims to build clients' confidence in themselves, therapy and therapist.

Perpetuating factors. The internal and external factors that maintain the current problems.	Provides a focus for intervention by breaking the maintenance cycle.
Protective factors. The person's resiliency and strengths that help maintain emotional health.	Provides a path of least resistance by suggesting interventions that build on existing resiliency and strengths.

Note. Adapted from Johnstone & Dallos (2014).

Miss X engaged fully in the formulation process and was open and honest throughout. The formulation produced is displayed in Figure 5.1.

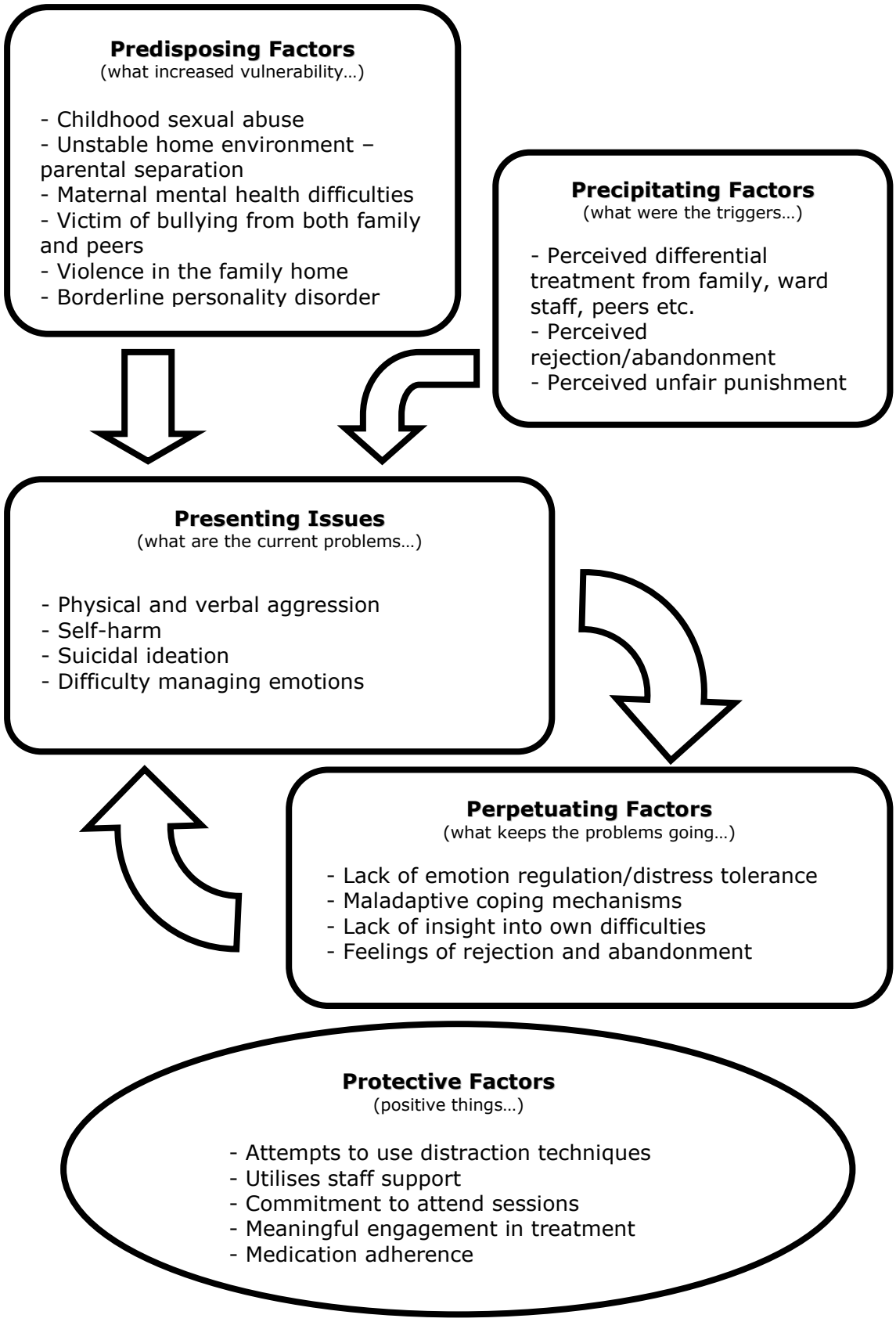


Figure 5.1 Case Formulation

5.1.4 Case Conceptualisation

Miss X's early childhood was characterised by instability. Within the family home there was domestic violence between parents which Miss X was witness to. Miss X also engaged in violence against her sister, who is two years her senior, with deliberate intent to cause harm. At the age of nine Miss X's parents separated due to an unstable home environment. This created a strain on Miss X's relationship with her mother, who blamed Miss X for her father leaving the family home. Miss X's mother suffered her own mental health difficulties with a diagnosis of borderline personality disorder, which also contributed to the strain on family life. One such example is that Miss X's mother would regularly drive Miss X and her sister to a bridge and explain that one day she would jump.

Miss X was the victim of bullying throughout her childhood from both family and peers. Much of the family bullying was as a result of her mother blaming Miss X for the parental separation. This reached its peak when Miss X was thirteen years old and her mother sent her to live with her uncle as she did not want Miss X to remain in the family home. Miss X internalised this as rejection from her mother. This also increased her feelings of unfair treatment as her sister remained in the family home. During her time with her uncle, Miss X became the victim of sexual abuse perpetrated by her uncle within the home. Miss X did not wish to disclose further information about this abuse, but reports suggest this lasted for approximately 18 months.

Miss X was known to CAMHS during her teenage years, and entered supported living at the age of 16. Miss X was struggling to cope in the community and had three inpatient admissions in quick succession as a result of attempted overdoses. At the time she was cutting herself using a razor blade, and was expressing suicidal ideation. Records suggest Miss X was admitted to an inpatient unit at the age of 18. Soon after admission she was found attempting hanging from the vents in the showers. Miss X presented with intermittent ligatures and intermittent thoughts of self-harm and suicide.

Miss X struggles to manage her emotions. As a result of her childhood and family dynamics, Miss X regularly perceives differential treatment on the ward from both staff and peers. This can be a trigger for aggressive outbursts followed by self-harming behaviours. Miss X finds it difficult to understand why these situations make her so distressed, and as a result will often put herself in situations where she is likely to receive differential treatment. For example Miss X will attend the clinic for medication with other patients when she is not scheduled to receive any. This means she interprets the situation as staff giving preferential treatment to other patients, which causes her distress.

Research has demonstrated that individuals who find it more difficult to understand and identify their emotions are more likely to engage in self-harming behaviours (McKay, Gavigan, & Mulchycky, 2004; Paivo & McCullock, 2004). This is the case with Miss X who finds it difficult to identify and express her emotions appropriately.

Similarly, those with childhood abuse and negative affect, combined with low emotional expressivity are also more likely to engage in self-harming behaviours (Gratz, 2006). Miss X's childhood was characterised by abuse and bullying which has put her at an increased risk of engaging in self-harming behaviours. Miss X struggles to express her emotions appropriately and this often results in aggression towards staff and peers.

By increasing Miss X's awareness of difficult emotions and alternative coping strategies, it was hypothesised that the number of self-harming incidents will reduce.

5.2 Intervention

The aims of the Coping Skills programme are detailed in Table 5.3. The chosen intervention worked to incorporate components of both Dialectical Behavioural Therapy (DBT) and Cognitive Behavioural Therapy (CBT). DBT was originally developed as a therapy to treat individuals with borderline personality disorder and teaches skills to manage overwhelming emotions (Linehan, 1993; McKay, Wood and Brantley, 2007). DBT consists of four modules including distress tolerance, emotion regulation, interpersonal effectiveness, and mindfulness (Linehan, 1993).

Table 5.3

Treatment Aims

Treatment Aim	Evidence
Identify and understand difficult emotions	- An ability to give examples of situations that may be difficult to deal with - Identify thoughts and bodily sensations associated with difficult emotions
Learning new coping skills	- Demonstrate a knowledge of different coping mechanisms - Completing a coping skills diary
Implementing these new skills	- Reduction in incidents on ward - Change in CSQ and BHS scores - Increase in protective factors on START, specifically coping

There is a strong evidence base for the use of mindfulness in women with borderline personality disorder (Shapiro & Carlson, 2009) in particular for those detained in secure services. Chilvers, Thomas and Stanbury (2011) found that introducing a mindfulness programme into a secure service significantly reduced the number of aggressive incidents observed over a six-month period. The rationale for including mindfulness into the Coping Skills programme was to encourage adaptive coping mechanisms. The strong literature for support of mindfulness with patients with a diagnosis of borderline personality disorder also supported its inclusion into the programme.

The programme also included aspects of CBT. CBT uses a wide range of techniques to modify behaviours associated with personality disorder, these include: cognitive re-structuring, behaviour modification, exposure, psychoeducation and specific skills training (Matusiewicz et al, 2010). There has been a wealth of research that has found CBT to be effective in reducing symptoms associated with personality disorder (Matusiewicz et al, 2010). Davidson et al (2006) conducted the Borderline Personality Disorder Study of Cognitive Therapy (BOSCOT) trial, which found that after CBT, individuals with borderline personality disorder reported lower levels of anxiety, less symptom distress, and less dysfunctional cognitions. There is also support for the use of CBT in secure settings. Harvey and Smedley (2010) describe the benefits of a CBT approach in terms of being able to tailor the therapy to the individuals needs, it can be problem-focused, and in-depth. By

incorporating aspects of CBT including psychoeducation and cognitive restructuring, it was hoped that the Coping Skills programme would reduce maladaptive coping mechanisms and decrease levels of hopelessness.

In addition to the within-session work, each week incorporated a homework assignment in order to encourage participants to practice the skills they had learnt in the session. These also provided qualitative evidence of progress in the group and progress toward the group aims.

Miss X was chosen for this programme as she demonstrates an inability to cope with difficult emotions. Miss X engages regularly in self-harming behaviours and struggles to regulate her emotions in response to stressful situations pertaining to feelings of perceived mistreatment or rejection. Miss X also displays an inability to recognise these difficult emotions and regularly puts herself in situations where she is likely to become distressed.

Miss X does not hold a diagnosis of learning disability, and does not struggle with communication, therefore no adaptations to the Coping Skills programme were necessary. The aim of Miss X's attendance in the programme was to teach her to gain insight into the situations and subsequent emotions she finds difficult to manage, and to develop more effective coping strategies.

5.3 Results

Miss X attended all ten session in this programme. These sessions are detailed in Table 5.4. Miss X's engagement and interaction during the sessions was good. Miss X demonstrated an

eagerness to learn and was seen to be taking notes throughout. Miss X volunteered to type up these notes for other members of the group and support them with their out-of-session homework. Both the in-session and out-of-session work was completed fully and with good effort and understanding demonstrated. Miss X regularly engaged in discussions during sessions, although at times overshadowed her peers and had difficulty turn-taking. However with guidance and support from facilitators this improved throughout the programme.

Table 5.4

Session Details

Session Date	Content
06.07.2015	Introduction to the group Aims: • For group members to understand what the group is about • To understand what the group will help them with • Develop group rules Learning point: • Learning how to function as a group and setting individual goals
13.07.2015	Difficult emotions Aims: • What are emotions and specifically what are difficult emotions • The effects of difficult emotions on the body & mind Learning point: • What are emotions and specifically what are difficult emotions. How difficult emotions effect our body and our mind
20.07.2015	Causes of difficult emotions Aims: • To understand why it is important to cope with

difficult emotions

- To understand generic causes of difficult emotions
- To understand personal causes of difficult emotions

Learning point:

- It is important to manage our emotions to be able to have a good quality of life. Knowing our triggers to difficult emotions can help us deal with them better by implementing coping strategies when triggers happen

27.07.2015 Mindfulness

Aims:

- To introduce the four principles of Mindfulness

Learning point:

- Mindfulness can help one to stay in the moment and manage emotions in a non-judgemental and accepting manner

03.08.2015 Relaxation

Aims:

- To discuss the factors that affect how we cope with difficult emotions
- To talk about how relaxation can help us cope with emotions

Learning point:

- Relaxation is a skill that is developed through practice much like mindfulness. With practice

relaxation can support us in managing difficult emotions

10.08.2015 Behavioural chain analysis

Aims:

- To introduce the Behaviour Chain
- Group to understand how a chain of thoughts, behaviours, emotions and events can effect the final behaviour

Learning point:

- Considering how we reacted to an event and thinking about different ways we could have reacted to have reached more positive consequences will help us create more positive outcomes

17.08.2015 Move/time out skill

Aims:

- Recap on the Behaviour Chain
- Introduce 'move skill'

Learning point:

- The move/ time out skill is a technique to help us get away from stressful situations. It can be used for some stressors but can't be used for others

24.08.2015 Distraction

Aims:

- Introduce concept of distraction
- Discuss implementation of this skill

Learning point:

- Distraction is not appropriate when you have a problem that needs resolving immediately. This includes stress such as money problems or problems with bills

14.09.2015 Asking for help and "Stop & Think"

Aims:

- To understand the value of asking for help
- Introduce stop & think

Learning point:

- Asking for help is different from demanding help. Stopping and thinking gives us time to cope with our emotions

21.09.2015 Recap, reflect, and goal setting

Aims:

- To recap on the skills to manage difficult emotions
- To reflect on the group

(i) Short-Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Desmarais, 2009):

Figures 5.2 and 5.3 display START strength and vulnerability scores respectively, both pre and post-intervention. The results show that Miss X's ratings did not differ from pre to post-intervention. However, her score for Coping increased from a "no" to a "partial" as a strength, and also decreased from a "yes" to a "partial" as a vulnerability. This was also maintained at the six-month follow-up. This demonstrates that as a result of completing the intervention on Coping Skills, Miss X has been able to demonstrate an ability to better cope with difficult situations, which has been noticed and documented by staff via the START assessment.

Miss X's specific risk estimates post-intervention did not change. However, at the six-month follow-up staff had reduced Miss X's self-harm risk from High to Medium.

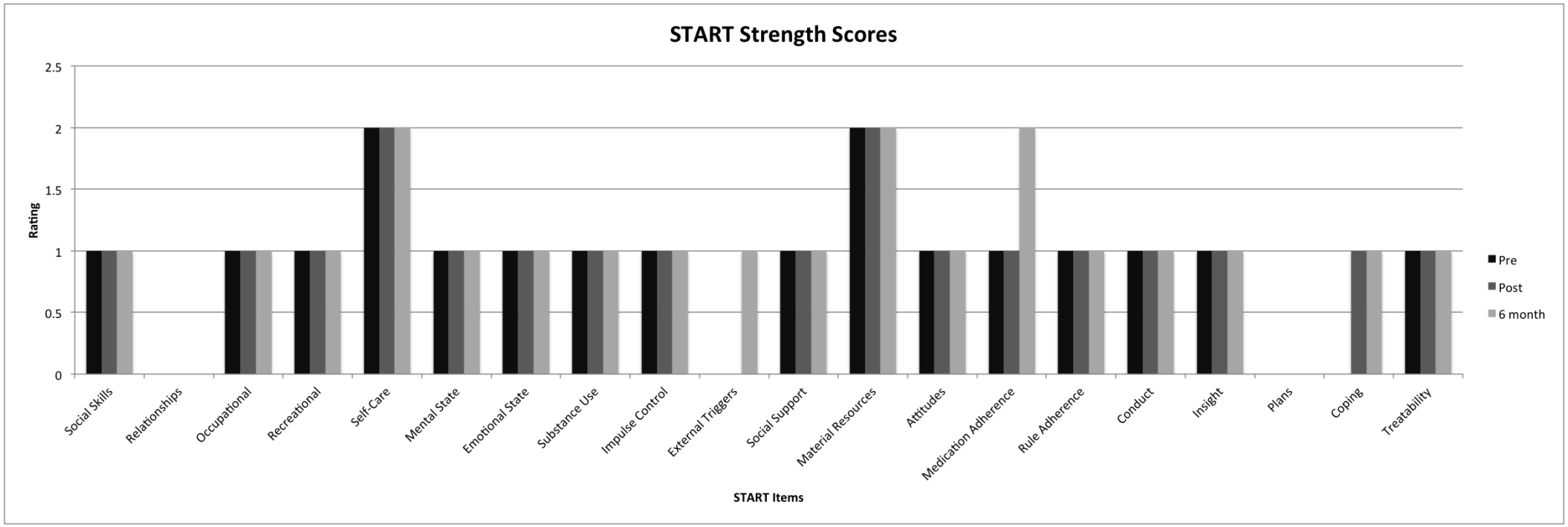


Figure 5.2 START Strength Scores

Note. 0=no, 1=partial, 2=yes

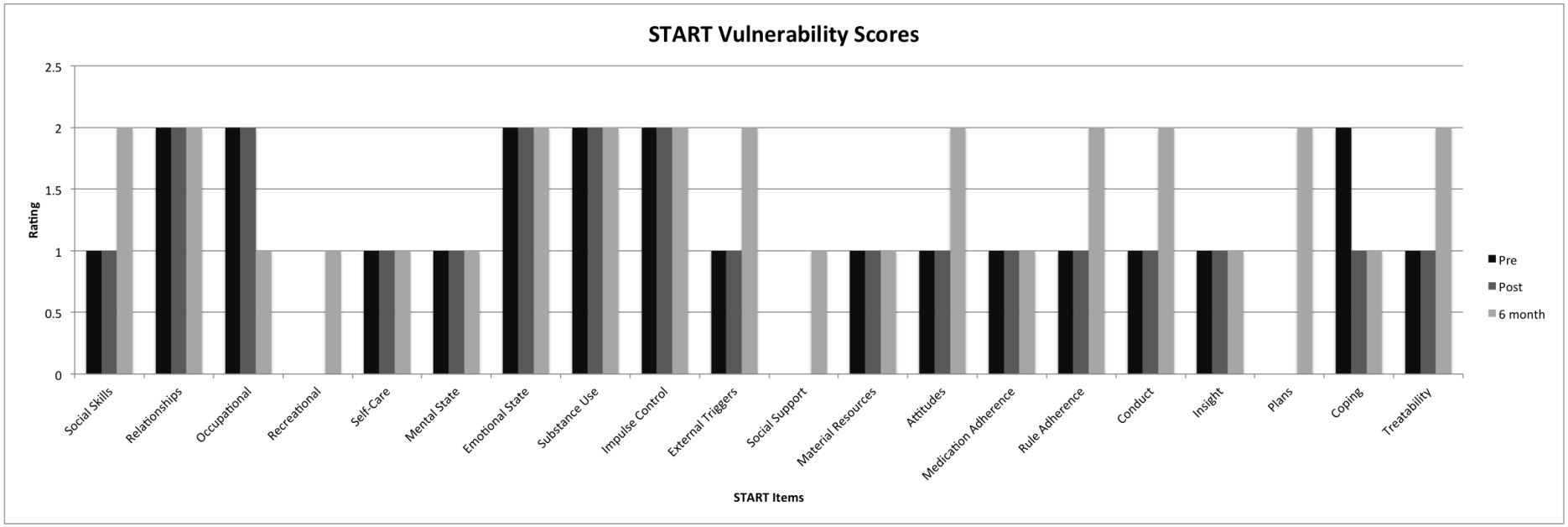


Figure 5.3 START Vulnerability Scores

Note. 0=no, 1=partial, 2=yes

(ii) Coping Styles Questionnaire (CSQ; Rogers & Najarian, 1989):

Table 5.5 displays the results for the CSQ (Roger & Najarian, 1989). Whilst Miss X's scores reflected a change pre to post-intervention, with a reduction in maladaptive coping styles and an increase in adaptive coping styles, it is important to determine if this can be considered clinically significant change. Jacobson and Truax (1991) define clinically significant change as two standard deviations from the mean. By these requirements, Miss X's changes in scores are not considered to be clinically significant. At the six-month follow-up, Miss X's scores for detachment had further increased, however her avoidance coping score had also increased (neither clinically significant). Miss X's scores for adaptive coping styles remained below the female norms, and scores for maladaptive coping styles increased above female norms.

Table 5.5

CSQ Scores

	Adaptive				**Maladaptive**			
	Rational Coping		Detachment		Emotional Coping		Avoidance Coping	
	Raw score	*z* score	Raw score	*z* score	Raw score	*z* score	Raw score	*z* score
Pre-intervention	11	-2.06	8	-1.64	27	+1.48	12	-0.69
Post-intervention	13	-1.75	9	-1.44	18	-0.04	11	-0.87
Six-month follow-up	12	-1.90	14	-0.41	19	+0.13	19	+0.62
Female norms:								
M	24.28		16.01		18.22		15.69	
SD	6.45		4.87		5.94		5.37	

(iii) Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974):

Miss X's pre-intervention score was 20 (z=1.82), which indicated a severe level of hopelessness. This measure was re-administered post-intervention and this score was also 20 (z=1.82). These results demonstrate no change in Miss X's level of hopelessness. Due to Miss X's history of suicidal ideation, her scores were compared to the "Suicide Attempters" norms in the BHS Manual (Beck et al., 1974). The mean score for this group was 8.86 (SD=6.11). Miss X's scored above the clinical norm, placing Miss X at an increased risk of suicide. This score and subsequent level of risk did not change as a result of completing the Coping Skills intervention at post-intervention assessment. At the six-month follow-up, Miss X's score on the BHS was 13 (z=0.68), demonstrating a reduction in hopelessness, although no reduction was clinically significant (Jacobson & Truax, 1991).

The post-intervention behavioural data, in respect to incident recordings, showed a reduction in incidents of self-harm from eleven in the month preceding the intervention, to eight in the month following the intervention. At the six-month follow-up, incidents over this month had reduced further to three incidents of self-harming behaviours.

5.4 Discussion

Whilst the present case study produced unremarkable results in terms of psychometrics, the behavioural findings are still noteworthy. Despite not demonstrating a clinically significant change, Miss X did show a reduction in maladaptive coping styles and an increase in adaptive coping. Whilst this may not constitute clinically significant change, the results demonstrate a change in the desired direction.

Miss X also demonstrated an increase in strength score on the START specifically relating to coping skills. This demonstrates recognition by her MDT that her ability to cope in stressful or difficult situations has increased. Miss X's score on the BHS did not change pre to post intervention. This is not necessarily a surprise, as whilst Miss X has learnt skills that will better enable her to cope, it may be that she is still experiencing the same levels of hopelessness but now has a better way of coping with these negative feelings. It may take more time for a significant change to be exhibited in the BHS score, as Miss X becomes more proficient in utilising her new skills.

There was little change in other areas of the START, however. This may not be a reflection of the lack of responsivity of the START, but rather the ineffectiveness of the intervention. One would expect an increase in other areas in addition to the rating of coping skills if an individual has adequately learnt and is effectively applying these new skills, as one is better able to manage other risks and vulnerabilities. However this is not evident in Miss X's

case, suggesting a lack of skills rather than poor responsivity of the START.

During any data collection, one must consider the potential bias of staff ratings, specifically regarding the increase in START strength score for coping skills. It could be argued that as staff were aware of Mss X's involvement in a group specifically designed to build coping skills, this would bias their opinion and create a more positive post-intervention rating, as to do otherwise would potentially criticise the value of the group or the skills of the facilitators.

The behavioural data gathered in the form of incident reports was insightful into the changes Miss X was making. Whilst the psychometric information did not yield a clinically significant change, the behavioural data demonstrated a reduction in self-harming behaviours recorded by ward staff from pre to post-intervention, and further still at the six-month follow-up. These findings suggest that Miss X was utilising the skills she learnt within the programme to better manage her emotions and using alternative strategies other than self-harm to cope in difficult situations.

However, it is important to recognise that there was very little behavioural data collected in the present study, and thus findings should be interpreted with caution. The reduction in risk behaviours may be due to chance, or may not be as a direct result of the intervention. In order to conclude that the behavioural change demonstrated by Miss X was as a direct result of the

intervention, more control would have been needed in relation to other treatments Miss X was receiving at the time, and the environment she was in in terms of reinforcing the skills learnt from the group. It would also be beneficial to have behavioural data recorded by an independent party who was not privy to or invested in the aims of outcome of the group in order to reduce any bias that may have been present. Without these adjustments, the behavioural data presented can only be considered partially supportive of the intervention's effectiveness.

Furthermore, there was little consideration given to the possibility that if Miss X gained insight into her problems and lack of coping skills, this may initially create a deterioration in presentation. Making an individual aware of their coping skills, or lack there of, may increase feelings of hopelessness, which may become an unintended by-product of engaging in coping skills work. Whilst the psychometric data did not suggest this, this may explain the minimal improvement Miss X made as a result of the intervention. A future consideration would be to acknowledge this potential short-term negative change and perhaps include parallel individual Psychology sessions to support with this aspect of the intervention.

Within secure settings, patients are not treated in a vacuum. Whilst Miss X demonstrated commitment to the Coping Skills programme, she was also attending a Mindfulness group, individual Psychology sessions, and other therapeutic activities such as Occupational Therapy and primary nurse sessions. It is possible that the mild change in scores on the assessments from pre to post

programme are reflective of the busy timetable Miss X had, and the lack of reinforcement from outside of the group. For instance, if Miss X's individual Psychology sessions had been used to compliment and strengthen the skills she was developing within the Coping Skills programme, this may have reinforced learning of these skills that may have been more evident in the post-intervention assessments.

On a similar note, it may have been beneficial to involve ward-based staff in the group. Whilst psychology was responsible for facilitating the group, having a member of ward staff present may have also been of benefit to the attendees. As psychology is not a ward-based discipline, involving this staff group may have helped to reinforce the skills learnt within the group, so that when patients request a primary nurse session, for example, staff can reinforce the coping skills they are learning within the group. The future implementation of this Coping Skills programme should incorporate this collaborative approach.

Miss X's busy timetable and lack of ward-based support is an important consideration, as if we know that someone will no get support outside of the group this may impact on change. It could be argued that a clinically informed decision should dictate to either alter these other conditions, or not put her through the group at this time, as there may have been a potential to set her up to fail. This oversight may have contributed to the minimal change from pre to post-intervention. Future research on the effectiveness of an intervention should ideally control the individual's environment as much as is clinically possible to ensure the individual has the best

chance for skill development, both within the intervention and outside of the timetabled therapeutic input.

The lack of change may also reflect the minimal length of the group. Ten sessions could be considered an introduction to coping skills, and it may have been optimistic to expect to see a significant change following such minimal psychological input. It may be that given a longer more substantiated intervention, perhaps including role play elements in order to rehearse the skills, Miss X would have displayed a significant change in psychometric data fro pre to post-intervention.

Delivering interventions via a group approach does offer benefits. The setting allows facilitation of ideas and a sharing of experiences between group members. Several psychological interventions recognise the benefit of delivering interventions via a group format. For example DBT is often delivered in this manner, and has demonstrated a significant reduction in risk behaviours across a variety of settings and client groups (Bloom, Woodward, Susmaras, & Pantalone, 2012; Feigenbaum et al., 2012). This supports the delivery of the Coping Skills programme as a group intervention, although as previously stated, results may be enhanced if this was combined with individual psychology sessions to support the skills learnt within the group.

In comparison to other members of the group, Miss X performed similarly. Only one member of the 6-person group demonstrated significant change from pre to post-intervention in relation to a reduction of psychometric scores on the BHS. This

suggests that Miss X's results are in line with the majority of the group, demonstrating that the intervention was largely ineffective in promoting coping skills. This is likely a reflection of the limitations listed previously, rather than Miss X's ability to learn the skills and apply them to herself effectively.

This case study has aimed to demonstrate the potential utility of promoting protective factors to decrease risk incidents. Whilst a change was not clinically significant in terms of the psychometric data collected, the behavioural data suggests the potential usefulness of this approach. By promoting protective factors in the form of adaptive and appropriate coping strategies, this intervention has resulted in a reduction in self-harming behaviours in this individual. Although more behavioural data is needed to fully support the hypothesis, the findings preliminarily suggest that focusing an intervention on promoting an individual's strengths can potentially reduce undesired risk behaviour.

Rationale for Chapter Six

The final chapter of this thesis will aim to draw together the findings of the research presented in the previous chapters. These findings will be contextualised within the current theoretical knowledge of forensic risk assessment and management, as well as discussing the clinical implications and directions for future research. The overall aims of the thesis will also be revisited and discussed in light of the current findings.

Chapter Six

Discussion

The findings from this thesis have not fully supported the clinical utility of including protective factors within forensic risk assessment. The findings of this thesis do not uphold the role of protective factors within risk prediction, but potentially there is some utility in risk management.

The overall aims of the current thesis will now be discussed in light of the current findings.

6.1 Aim 1: Identify current risk assessment tools that incorporate protective factors and their effectiveness at predicting risk behaviours within secure mental health services

The systematic review identified three risk assessment tools which incorporate protective factors, the SAPROF, SAVRY, and the START. The findings suggested that rather than the protective factors themselves being a good predictor of risk, it may be what they add to the process of risk assessment in allowing the rater to consider an individual's strengths to create a more balanced risk assessment. It is unlikely that protective factors will ever outperform risk factors for the prediction of risk incidents, but including a consideration of protective factors could help to increase the accuracy of a risk assessment tool.

6.2 Aim 2: To critically evaluate the structure of such a tool

A critique of the START questioned the clinical utility of protective factors. However a criticism of the current literature pertaining to the START is that the majority focuses on violent outcomes, ignoring the remaining six outcomes. Whilst the START

demonstrated excellent levels of inter-rater reliability, the critique questions the clinical utility of rating each individual item of the START as both a strength and a vulnerability. The strong negative correlations between rating of one item as a strength and a vulnerability suggests that raters view each item as a dichotomous variable. Webster et al., (2009) state that each item should be viewed separately as a strength and a vulnerability, such that an individual can score highly on one item in both aspects. For example, for the item 'Relationships', this could be a strength in terms of relationships with staff, but a vulnerability in terms of social or familial relationships, meaning this item is both a strength and a vulnerability. The current literature suggests that the START is not being rated in this way, and that perhaps for the purpose of risk prediction, this dichotomous answering style is more useful. In addition to this, the subjectivity of the Likert-type scale increases potential mis-rating of an individual's risk. Therefore, whilst at first glance the START may appear to be a balanced assessment of risk, closer examination does not overwhelmingly support the effectiveness of this tool.

6.3 Aim 3: To understand the specific role of a measure of protective strengths within a risk assessment tool

The START was investigated for predictive validity within a sample of female inpatients in levels of medium security. The SREs were the most predictive aspect of the START, which Webster et al., (2009) recommend as a product of considering both the strength and the vulnerability scale. These findings suggest that when both

strengths and risk factors are considered together, this creates a more accurate prediction of an individual's likelihood to engage in risk behaviours than either of these factors individually. However, if these items are being rated as opposites of each other, there is no justification to rating each item as both a strength and vulnerability, and perhaps having overall risk judgments would be more clinically useful.

Additionally, a consideration of the population is needed when interpreting the findings. As previously stated self-harm is likely to be the main risk within this population, giving staff a certain area of expertise in terms of predicting this type of risk. Therefore it may be that the inaccuracy in risk prediction is not a reflection of the START, but rather a reflection of staff and their ability or need to predict risk outside of self-harming behaviours.

6.4 Aim 4: To investigate if promoting said protective factors can reduce risk behaviours in an individual

Finally, it is important to consider the role of protective factors for the management of risk behaviours in addition to the prediction. A case study was used to identify the role of promoting protective factors in order to reduce a risk behaviour, in this case promoting coping skills in order to reduce self-harm behaviours. By promoting these skills across a ten-week programme, there was a reduction in the number of self-harm incidents displayed by the individual at both the immediate follow-up post-intervention, followed by a further reduction at the six-month follow-up. This case study demonstrates the potential effectiveness of promoting

protective factors in an individual in order to reduce a risk behaviours, however it is important to consider the potential bias of raters who are aware of the skills the individual is targeting through the intervention.

This approach to reducing risk behaviours follows the same principles as the Good Lives Model (Ward & Stuart, 2003). By focusing on an individual's strengths and providing them with the skills to lead more satisfying and rewarding lives, undesired outcomes can be reduce (Ward & Brown, 2004). This case study demonstrates how this model of rehabilitation could be effectively implemented within a secure environment to reduce risk behaviours, if steps are taken to ensure the clinical environment supports the development of these skills and reduces rater bias in pre to post-assessment.

6.5 Clinical Implications

This thesis has demonstrated that whilst the specific utility of protective factors in risk prediction is underwhelming. However there may be some utility in this approach in the subsequent risk management and treatment planning than the actual prediction of risk behaviours. By identifying protective factors in a risk assessment process, these could potentially be utilized in treatment planning to play to the individual's strengths, and use these to engage an individual in effective treatment and highlight areas for further focus. The case study offered some support for this approach, as by concentrating on promoting protective factors, the risk behaviour was reduced, although a change in pre to post-

intervention assessment scores was not evident. The use of strength-based approaches such as the Good Lives Model (Ward & Stuart, 2004) are increasing within mental health services, and the case study provided some support for this approach.

The START was the main focus of the current thesis, reflecting its prevalence within secure mental health settings. For the purpose of this research these scales were investigated individually, however in clinical practice, this is not the case. The findings of the current study offer very little clinical utility in separating out strength and vulnerability ratings, and suggest that the clinical judgment aspect of the SRE's provide the only accurate prediction of risk outcomes.

Though the current findings unsurprisingly make it difficult to argue that protective factors are better predictors of risk behaviours, perhaps the clinical utility of incorporating protective factors lies in the perception of patients. As Rogers (2000) argues, a risk-only assessment is inherently biased, and creates a more 'risky' view of the individual. If staff and patients are solely focused on risk factors, this may be demotivating for staff who see the positives in individuals, and patients who are expected to remain enthusiastic about their capabilities and potential when this aspect of risk behaviour is invalidated.

6.6 Future Directions

Whilst the role of protective factors is a very current and relevant field of research, there is still some way to go in determining the exact role these factors play, and how they can

best be utilised within clinical practice. As demonstrated during the systematic review of the current literature, the majority of the research has been conducted using the START. The SAPROF and the SAVRY received little attention in the literature, and thus more is needed to establish an understanding of how these tools utilise protective factors in their judgement of risk.

Similarly, most of the research is conducted on male samples. This reduces the generalizability of the findings, and although males are more represented in the forensic population, there is still a subset of females within secure services that the literature needs to acknowledge. Research into this has already started, with the HCR-20 offering an addition specifically for females. However, more work is needed to highlight and evidence the differences in risk assessment for men and women.

Future research should look at other areas where protective factors may be more clinically useful. For example, looking at the influence of protective factors on formulation and treatment planning, or how a measure of protective factors influences staff views of patients and subsequent interpersonal relationships. This could provide more of an insight into the clinical utility of protective factors that don't appear to lie in risk prediction.

6.7 Conclusion

It is evident from this thesis that more research is needed to establish the specific role of protective factors. The findings presented question the clinical utility of including such a measure in forensic risk assessment in relation to predictive accuracy.

Protective factors can often be overlooked within forensic settings, as risk factors are deemed to be more important. It could be argued that the acknowledgement of an individual's strengths could act to increase our understanding of the individual, potentially providing a more balanced view of the individual, ensuring appropriate management strategies and treatment goals are achieved. However, the current thesis did not examine this concept, and this could be a direction for future research into the role of protective factors within forensic care.

So, whilst one aims to always look on the bright side of life, perhaps in the case of forensic risk assessment one should turn to the dark side.

References

Abderhalden, C., Needham, I., Miserez, B., Almvik, R., Dassen, T., Haug, H., & Fischer, J. (2004). Predicting inpatient violence in acute psychiatric wards using the Broset-Violence-Checklist: A multicentre prospective cohort study. *Journal of Psychiatric and Mental Health Nursing, 11*(4), 422-427. doi: 10.1111/j.1365-2850.2004.00733.x

Abderhalden, C., Needham, I., Dassen, T., Halfens, R., Haug, H.-J., & Fischer, J. (2006). Predicting inpatient violence using an extended version of the Broset Violence-Checklist: Instrument development and clinical application. *BMC Psychiatry, 6(17), 6*-15. doi: 10.1186/1471-244X-6-17

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), 197-214. doi: 10.1186/1471-244X-13-197

Allnutt, S. H., Ogloff, J. R., Adams, J., O'Driscoll, C., Daffern, M., Carroll, A., Nanayakkara, V., & Chaplow, D. (2013). Managing aggression and violence: the clinician's role in contemporary mental health care. *Australian & New Zealand Journal of Psychiatry, 47*(8), 728-736. doi: 10.1177/0004867413484368

Almvik, R., Woods, P., & Rasmussen, K. (2007). Assessing risk for imminent violence in the elderly: the Broset Violence Checklist. *International Journal of Geriatric Psychiatry, 22*(9), 862-867. doi:10.1002/gps.1753

American Psychiatric Association. (2013). *Diagnostic and statistical*

manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Amore, M., Menchetti, M., Tonti, C., Scarlatti, F., Lundgren, E., Esposito, W., & Berardi, D. (2008). Predictors of violent behavior among acute psychiatric patients: Clinical study. *Psychiatry and Clinical Neurosciences, 62*(3), 247-255. doi: 10.1111/j.1440-1819.2008.01790.x

Anderson, T. R., Bell, C. C., Powell, T. E., Williamson, J. L., & Blount, M. A., Jr. (2004). Assessing Psychiatric Patients For Violence. *Community Mental Health Journal, 40*(4), 379-399. doi: 10.1023/B:COMH.0000035231.24472.bc

Andrews, D. A., Bonta, J., & Wormith, S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency, 52(7),* 7-27. doi: 10.1177/0011128705281756

Anonymous (2013). Violence in clinical psychiatry. *European Psychiatry. Conference: 21st European Congress of Psychiatry, EPA, 28*. Retrieved from http:// www.oudconsultancy.nl

Anonymous. (2013). Aggression in psychiatry: Impact of family history, substance use, psychiatric history, and dual diagnosis. *Indian Journal of Psychiatry, 55*, S82. Retrieved from http://www. connection.ebscohost.com

Arbach-Lucioni, K., Andres-Pueyo, A., Pomarol-Clotet, E., & Gomar-Sones, J. (2011). Predicting violence in psychiatric inpatients: A prospective study with the HCR-20 violence risk assessment scheme. *Journal of Forensic Psychiatry & Psychology, 22*(2), 203-222. doi:10.1080/14789949.2010.530290

Ash, P., & Alonso-Katzowitz, J. (2011). Commentary: Predicting aggression in youth. *Journal of the American Academy of Psychiatry and the Law, 39*(2), 180-188. Retrieved from http://www.jaapl.org

Baksheev, G. N., Sullivan, D., Fraser, R., Greenwood-Smith, C., McCarthy, J., Reid, D., Lemphers, A., & Purcell, R. (2015). Characteristics of Mentally Disordered Youth Referred to a Forensic Satellite Clinic for Violence Risk Assessment: A Case Control Study. *Psychiatry Psychology and Law, 22*(2), 235-246. doi:10.1080/13218719. 2014.941088

Ball, E. M., Young, D., Dotson, L. A., Brothers, L. T., & Robbins, D. (1994). Factors associated with dangerous behavior in forensic inpatients - results from a pilot-study. *Bulletin of the American Academy of Psychiatry and the Law, 22*(4), 605-620. Retrieved from http://www.jaapl.org

Barry-Walsh, J., & Daffern, M. (2010). The DASA, accuracy and risk assessment. *Australasian Psychiatry, 18*(6), 587. doi: 10.3109/10398562.2010.526215

Barry-Walsh, J., Daffern, M., Duncan, S., & Ogloff, J. (2009). The prediction of imminent aggression in patients with mental illness and/or intellectual disability using the Dynamic Appraisal of Situational Aggression instrument. *Australasian Psychiatry, 17*(6), 493-496. doi: 10.1080/ 10398560903289975

Barzman, D. H., Brackenbury, L., Sonnier, L., Schnell, B., Cassedy, A., Salisbury, S., Sorter, M., & Mossman, D. (2011). Brief Rating of Aggression by Children and Adolescents (BRACHA): Development of

a tool for assessing risk of inpatients' aggressive behavior. *Journal of the American Academy of Psychiatry and the Law, 39*(2), 170-179. Retrieved from http://www.jaapl.org

Beck, A. T., & Steer, R. A. (1993). *Beck hopelessness scale: Manual.* The Psychological Corporation: USA.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The hopelessness scale. *Journal of Consulting and Clinical Psychology, 42*(6), 861-865. doi: 10.1037/h0037562

Beck, J. C., McNiel, D. F., & Binder, R. L. (1996). Predicting inpatient violence. *American Journal of Psychiatry, 153*(6), 845-846. Retrieved from http://www.search.proquest.com

Becker, D. F., & Grilo, C. M. (2007). Prediction of suicidality and violence in hospitalized adolescents: Comparisons by sex. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 52*(9), 572-580. doi: 10.1177/070674370705200905

Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: integrating research and clinical practice. *Journal of Forensic Psychiatry, 9*(2), 328-338. doi:10.1080/09585189808402200

Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: A prospective study in two maximum-security correctional institutions. *Journal of Forensic Psychiatry, 11*(1), 167-175. doi:10.1080/095851800362445

Bijttebier, P., and Vertommen. H. (1999). Coping Strategies in

relation to personality disorders. *Personality and Individual Differences, 26*(5), 847-856. doi: 10.1016/S0191-8869(98)00187-1

Bjorkdahl, A., Olsson, D., & Palmstierna, T. (2006). Nurses' short-term prediction of violence in acute psychiatric intensive care. *Acta Psychiatrica Scandinavica, 113*(3), 224-229. doi:10.1111/j.1600-0447.2005.00679.x

Bjorkly, S. (1993). Scale for the prediction of aggression and dangerousness in psychotic-patients, an introduction. *Psychological Reports, 73*(3), 1363-1377. doi:10.2466/pr0.1993.73.3f.1363

Bjorkly, S. (1994). The scale for the prediction of aggression and dangerousness in psychotic-patients (pad) - a prospective pilot-study. *Criminal Justice and Behavior, 21*(3), 341-356. doi:10.1177/0093854894021003004

Bjorkly, S. (1995). Prediction of aggression in psychiatric-patients - a review of prospective prediction studies. *Clinical Psychology Review, 15*(6), 475-502. doi:10.1016/0272-7358(95)00016-i

Bjorkly, S. (1997). Clinical assessment of dangerousness in psychotic patients: Some risk indicators and pitfalls. *Aggression and Violent Behavior, 2*(2), 167-178. doi:10.1016/s1359-1789(96)00024-9

Bjorkly, S., Eidhammer, G., & Selmer, L. E. (2014). Concurrent Validity and Clinical Utility of the HCR-20(V3) Compared With the HCR-20 in Forensic Mental Health Nursing: Similar Tools but Improved Method. *Journal of Forensic Nursing, 10*(4), 234-242. doi:10.1097/jfn.0000000000000047

Bjorkly, S., Hartvig, P., Heggen, F., Brauer, H., & Moger, T. (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*(6), 388-394. doi: 10.1016/j.eurpsy.2009.07.004

Bjorkly, S., Havik, O. E., & Loberg, T. (1996). The interrater reliability of the Scale for the Prediction of Aggression and Dangerousness in Psychotic Patients (PAD). *Criminal Justice and Behavior, 23*(3), 440-454. doi:10.1177/0093854896023003002

Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation.* London: Sage Publications.

Bloom, H., Webster, C., Hucker, S., & De Freitas, K. (2005). The Canadian contribution to violence risk assessment: history and implications for current psychiatric practice. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie, 50*(1), 3-11. doi: 10.1177/070674370505000103

Bloom, J. M., Woodward, E. N., Susmaras, T., & Pantalone, D. W. (2012). Use of dialectical behavior therapy in inpatient treatment of borderline personality disorder: a systematic review. *Psychiatric Services, 63*(9), 881-888. doi: 10.1176/appi.ps.201100311

Boer, D., Hart, S., Kropp, P., & Webster, C. (1997). *Manual for the Sexual Violence Risk-20: Professional guidelines for assessing risk of sexual violence.* Burnaby, BC: Simon Fraser University, Mental Health, Law, and Policy Institute.

Borum, R. (1996). Improving the clinical practice of violence risk Assessment: Technology, guidelines, and training. *American*

Psychologist, 51(9), 945-956. doi:10.1037/0003-066x.51.9.945

Borum, R., Bartel, P., & Forth, A. (2002) *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY), Version 1, consultation edition.* Tampa: Department of Mental Health Law and Policy, University of South Florida

Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY).* Odessa, FL: Psychological Assessment Resources.

Boxer, P. (2007). Aggression in very high-risk youth: Examining developmental risk in an inpatient psychiatric population. *American Journal of Orthopsychiatry, 77*(4), 636-646. doi: 10.1037/0002-9432.77.4.636

Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *The International Journal of Forensic Mental Health, 9*(4), 271-281. doi:10.1080/14999013.21.534378

Brown, M., Comtois, Z., Katherine, A., and Linehan, M. (2002). Reasons for suicide attempts and nonsuicidal self-injury in women with borderline personality disorder. *Journal of Abnormal Psychology*, *111*(1), 198-202. doi: 10.1037/0021-843X.111.1.198

Buchanan, A. (1999). Risk and dangerousness. *Psychological Medicine, 29*(2), 465-473. doi:10.1017/s0033291798008101

Buchanan, A. (2008). Risk of violence by psychiatric patients: Beyond the "actuarial versus clinical" assessment debate. *Psychiatric Services,*

59(2), 184-190. doi:10.1176/appi.ps.59.2.184

Buchanan, A. (2013). Violence risk assessment in clinical settings: being sure about being sure. *Behavioral Sciences & the Law, 31*(1), 74-80. doi: 10.1002/bsl.2045

Buchanan, A., Binder, R., Norko, M., & Swartz, M. (2012). Psychiatric violence risk assessment. *American Journal of Psychiatry, 169*(3), 340. doi: 10.1176/appi.ajp.2012.169.3.340

Buchanan, A., Nich, C., Douglas, K. S., Babuscio, T., & Easton, C. J. (2013). Risk factors of violence during a 4-week period in a psychiatric outpatient population. *Journal of Nervous and Mental Disease, 201*(12), 1021-1026. doi: 10.1097/NMD.0000000000000061

Burke, T. (2010). Psychiatric disorder: Understanding Violence. In *Forensic mental health: Concepts, systems, and practice* (pp. 35-51). New York, NY: Oxford University Press

Burman, M., Armstrong, S., Batchelor, S., McNeill, F., & Nicholson, J. (2007). *Research and practice in risk assessment and risk management of children and young people engaging in offending behaviours: A Literature Review.* RMA Scotland

Calow, N., Lewis, A., Showen, S., & Hall, N. (2016). Literature synthesis: Patient aggression risk assessment tools in the emergency department. *Journal of Emergency Nursing, 42*(1), 19-24. doi:10.1016/j.jen.2015.01.023

Candini, V., Buizza, C., Ferrari, C., Boero, M. E., Giobbio, G. M., Goldschmidt, N., Greppo, S., Lozzino, L., Maggi, P., Melegari, A., Pasqualetti, P., Rossi, G., & de Girolamo, G. (2015). Violent

behavior of patients living in psychiatric residential facilities: A comparison of male patients with different violence histories. *International Journal of Law and Psychiatry, 39*(1), 46-51. doi:10.1016/j.ijlp.2015.01.020

Carroll, A. (2007). Are violence risk assessment tools clinically useful? *Australian and New Zealand Journal of Psychiatry, 41*(4), 301-307. doi:10.1080/00048670701213237

Castelletti, L., Rivellini, G., & Stratico, E. (2014). Predictive efficacy of violence risk assessment tools, implications for forensic and general psychiatry in Italy: A literature review. *Journal of Psychopathology, 20*(2), 153-162. Retrieved from http://www.jpsychopathol.it

Chagigiorgis, H., Michel, S. F., Seto, M. C., Laprade, K., & Ahmed, A. G. (2013). Assessing short-term, dynamic changes in risk: The predictive validity of the Brockville Risk Checklist. *The International Journal of Forensic Mental Health, 12*(4), 274-286. doi:10.1080/14999013.2013.857740

Chakhssi, F., de Ruiter, C., & Bernstein, D. (2010). Reliability and validity of the Dutch version of the Behavioural Status Index: A nurse-rated forensic assessment tool. *Assessment, 17*(1), 58-69. doi: 10.1177/1073191109338815

Chan, O., & Chow, K. K. (2014). Assessment and determinants of aggression in a forensic psychiatric institution in Hong Kong, China. *Psychiatry Research, 220*(1), 623-630. doi: 10.1016/ j.psychres.2014.08.008

Chapman, A. L., Specht, M. W., & Cellucci, T. (2005). Borderline

personality disorder and deliberate self-harm: Does experiential avoidance play a role? *Suicide and Life-Threatening Behavior*, *35(4),* 388-399. doi: 10.1521/suli.2005.35.4.388

Chen, S. C., & Hwu, H. G. (2009). Psychometric evaluation of Chinese version of Violence Scale for objective rating among inpatients with schizophrenia. *Journal of Clinical Nursing, 18*(13), 1889-1896. doi: 10.1111/j.1365-2702.2008.02669.x

Chen, S. C., Hwu, H. G., & Hu, F. C. (2014). Clinical prediction of violence among inpatients With schizophrenia using the Chinese modified version of Violence Scale: A prospective cohort study. *International Journal of Nursing Studies, 51*(2), 198-207. doi:10.1016/j.ijnurstu.2013.06.002

Chilvers, J., Thomas, C., & Stanbury, A. (2011). The impact of a ward-based mindfulness programme on recorded aggression in a medium secure facility for women with learning disabilities. *Journal of Learning Disabilities and Offending Behaviour, 2*(1), 27-41. doi:10.5042/jldob.2011.0026

Chu, C. M., Daffern, M., & Ogloff, J. R. (2013). Predicting aggression in acute inpatient psychiatric setting using BVC, DASA, and HCR-20 Clinical scale. *Journal of Forensic Psychiatry & Psychology, 24*(2), 269-285. doi:10.1080/14789949.2013.773456

Chu, C. M., Hoo, E., Daffern, M., & Tan, J. (2012). Assessing the risk of imminent aggression in institutionalized youth offenders using the dynamic appraisal of situational aggression. *Journal of Forensic Psychiatry & Psychology, 23*(2), 168-183. doi:10.1080/

14789949.2012.668207

Chu, C. M., Thomas, S. D., Daffern, M., & Ogloff, J. R. (2013). Should clinicians use average or peak scores on a dynamic risk-assessment measure to most accurately predict inpatient aggression? *International Journal of Mental Health Nursing, 22*(6), 493-499. doi:10.1111/j.1447-0349.2012.00846.x

Chu, C. M., Thomas, S. D., Ogloff, J. R., & Daffern, M. (2011). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) in a secure forensic hospital: Risk factors and strengths. *The International Journal of Forensic Mental Health, 10*(4), 337-345. doi: 10.1080/14999013.2011.629715

Chu, C. M., Thomas, S. D., Ogloff, J. R., & Daffern, M. (2013). The short- to medium-term predictive accuracy of static and dynamic risk assessment measures in a secure forensic hospital. *Assessment, 20*(2), 230-241. doi: 10.1177/1073191111418298

Claix, A., & Pham, T. (2004). Evaluation of the HCR-20 violence risk assessment scheme in a Belgian forensic population. *L'Encephale: Revue de psychiatrie clinique biologique et therapeutique, 30*(5), 447-453. doi:10.1016/S0013-7006(04)95459-0

Clark, L., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319. doi: 10.1037/1040-3590.7.3.309

Clarke, D. E., Brown, A. M., & Griffith, P. (2010). The Broset Violence Checklist: clinical utility in a secure psychiatric intensive care setting. *Journal of Psychiatric and Mental Health Nursing, 17*(7), 614-620. doi:10.1111/j.1365-2850.2010.01558.x

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi: 0.1037/0033-2909.112.1.155

Coid, J., Yang, M., Ullrich, S., Zhang, T. Q., Sizmur, S., Roberts, C., Farrington, D. P., & Rogers, R. D. (2009). Gender differences in structured risk assessment: Comparing the accuracy of five instruments. *Journal of Consulting and Clinical Psychology, 77*(2), 337-348. doi:10.1037/a0015155

Convit, A., Jaeger, J., Lin, S. P., Meisner, M., & Volavka, J. (1988). Predicting assaultiveness in psychiatric inpatients: A pilot study. *Hospital & Community Psychiatry, 39*(4), 429-434. doi: 10.1176/ps.39.4.429

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry wards: A systematic review. *Psychiatry Research, 189*(1), 10-20. doi:10.1016/j.psychres.2010.12.024

Coté, G., Crocker, A. G., Nicholls, T. L., & Seto, M. C. (2012). Risk assessment instruments in clinical practice. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 57*(4), 238-244. doi: 10.1177/070674371205700407

Crocker, J. H., Stargatt, R., & Denton, C. (2010). Prediction of aggression and restraint in child inpatient units. *Australian and New Zealand Journal of Psychiatry, 44*(5), 443-449. doi: 10.3109/00048670903489825

Cronin-Davis, J. (2000). The risk assessment of violent mentally disordered offenders in secure hospital environments: What are the factors considered by clinical staff? *British Journal of Occupational Therapy, 63*(6), 259-266. Retrieved from

http://www.researchgate.net

Dack, C., Ross, J., Papadopoulos, C., Stewart, D., & Bowers, L. (2013). A review and meta-analysis of the patient factors associated with psychiatric in-patient aggression. *Acta Psychiatrica Scandinavica, 127*(4), 255-268. doi:10.1111/acps.12053

Daffern, M. (2007). The predictive validity and practical utility of structured schemes used to assess risk for aggression in psychiatric inpatient settings. *Aggression and Violent Behavior, 12*(1), 116-130. doi: 10.1016/j.avb.2006.03.005

Daffern, M., & Howells, K. (2007). 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. *The International Journal of Forensic Mental Health, 6*(2), 137-143. doi:10.1080/1499913. 2007.10471258

Daffern, M., Howells, K., Ogloff, J., & Lee, J. (2005). Individual characteristics predisposing patients to aggression in a forensic psychiatric hospital. *Journal of Forensic Psychiatry & Psychology, 16*(4), 729-746. doi:10.1080/14789940500345595

Daffern, M., Ogloff, J. R., Ferguson, M., & Thomson, L. (2005). Assessing risk for aggression in a forensic psychiatric hospital using the Level of Service Inventory-Revised: Screening Version. *The International Journal of Forensic Mental Health, 4*(2), 201-206. doi:10.1080/14999013.2005.10471224

Daniel, M. R., Weir, S., & Tiffin, P. A. (2013). A novel CAMHS risk assessment system: Clinicians' views. *Journal of Forensic Practice,*

15(3), 182-191. doi: 10.1108/JFP-09-2012-0011

Davidson, K., Norrie, J., Tyrer, P., Gumley, A., Tata, P., Murray, H., & Palmer, S. (2006). The effectiveness of cognitive behavior therapy for borderline personality disorder: results from the borderline personality disorder study of cognitive therapy (BOSCOT) trial. *Journal of Personality Disorder, 20*(5), 450-465. doi: 10.1521/pedi.2006.20.5.450

de Borba Telles, L. E., Folino, J. O., & Taborda, J. G. V. (2012). Accuracy of the Historical, Clinical and Risk Management Scales (HCR-20) in predicting violence and other offenses in forensic psychiatric patients in Brazil. *International Journal of Law and Psychiatry, 35*(5), 427-431. doi: 10.1016/j.ijlp.2012.09.001

de Ruiter, C., & de Vogel, V. (2004). Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients. *Journal of Forensic Psychiatry and Psychology, 15*(1), 145-164. doi:10.1080/14789940410001655916

de Ruiter, C., & Nicholls, T. L. (2011). Protective factors in forensic mental health: A new frontier. *International Journal of Forensic Mental Health, 10*(3), 160-170. doi:10.1080/14999013.2011.600602

de Vogel, V., & de Ruiter, C. (2005). The HCR-20 in Personality Disordered Female Offenders: A Comparison with a Matched Sample of Males. *Clinical Psychology & Psychotherapy, 12*(3), 226-240. doi: 10.1002/cpp.452

de Vogel, V., & de Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into

the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*(3), 321-336. doi:10.1080/10683160600569029

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2009). *SAPROF: Guidelines for the assessment of protective factors for violence risk (English version).* Utrecht, The Netherlands: Forum Educatief

de Vogel, V., de Ruiter, C., Hildebrand, M., Bos, B., & van de Ven, P. (2004). Type of Discharge and Risk of Recidivism Measured by the HCR-20: A Retrospective Study in a Dutch Sample of Treated Forensic Psychiatric Patients. *The International Journal of Forensic Mental Health, 3*(2), 149-165. doi:10.1080/1499013.2014.10471204

de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law & Human Behavior, 28*(3), 235-251. doi: 10.1023/B:LAHU.0000029137.41974.eb

de Vogel, V., de Vries Robbé, M., de Ruiter, C., & Bouman, Y. (2011). Assessing protective factors in forensic psychiatric practice: Introducing the SAPROF. *International Journal of Forensic Mental Health, 10*(3), 171-177. doi: 10.1080/14999013.2011.600230

de Vogel, V., Stam, J., Bouman, Y. H. A., Ter Horst, P., & Lancel, M. (2016). Violent women: a multicentre study into gender differences in forensic psychiatric patients. *Journal of Forensic Psychiatry & Psychology, 27*(2), 145-168. doi:10.1080/14789949.2015.1102312

de Vries Robbé, M. D., de Vogel, V., & Douglas, K. S. (2013). Risk factors

and protective factors: A two-sided dynamic approach to violence risk assessment. *Journal of Forensic Psychiatry and Psychology, 24*(4), 440-457. doi: 10.1080/14789949.2013.818162

de Vries Robbé, M. D., de Vogel, V., Douglas, K. S., & Nijman, H. L. I. (2015). Changes in Dynamic Risk and Protective Factors for Violence During Inpatient Forensic Psychiatric Treatment: Predicting Reductions in Postdischarge Community Recidivism. *Law and Human Behavior, 39*(1), 53-61. doi:10.1037/lhb0000089

de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients: A retrospective validation study of the SAPROF. *The International Journal of Forensic Mental Health, 10*(3), 178-186. doi: 10.1080/14999013.2011.600232

de Vries Robbé, M., de Vogel, V., & Stam, J. (2012). Protective factors for violence risk: The value for clinical practice. *Psychology*, *3*(12), 1259. doi: 10.4236/psych.2012.312A187

de Vries Robbe, M., de Vogel, V., Douglas, K. S., & Nijman, H. L. (2015). Changes in dynamic risk and protective factors for violence during inpatient forensic psychiatric treatment: Predicting reductions in postdischarge community recidivism. *Law and Human Behavior, 39*(1), 53-61. doi: 10.1037/lhb0000089

Department of Health. (2009). *Best Practice in Managing Risk. Principles and evidence for best practice in the assessment and management of risk to self and others in mental health services.* Retrieved from https://www.gov.uk/government/

uploads/system/uploads/attachment_data/file/478595/best-practice-managing-risk-cover-webtagged.pdf

Dernevik, M., Grann, M., & Johansson, S. (2002). Violent behaviour in forensic psychiatric patients: Risk assessment and different risk-management levels using the HCR-20. *Psychology, Crime & Law, 8*(1), 93-111. doi:10.1080/10683160208401811

Desmarais, S. L., Nicholls, T. L., Read, J. D., & Brink, J. (2010). Confidence and accuracy in assessments of short-term risks presented by forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 1-22. doi: 10.1080/14789940903183932

Desmarais, S. L., Nicholls, T. L., Wilson, C. M., & Brink, J. (2012). Using dynamic risk and protective factors to predict inpatient aggression: Reliability and validity of START assessments. *Psychological assessment*, *24*(3), 685-700. doi: 10.1037/a0026668

Dickens, G. L., & O'Shea, L. E. (2015). How short should short-term risk assessment be? Determining the optimum interval for START reassessment in a secure mental health service. *Journal of Psychiatric and Mental Health Nursing, 22*(6), 397-406. doi: 10.1111/jpm.12232

Dietl, J., Hollweg, M., Kleindienst, N., Nedopil, N., Reich, U., & Stadtland, C. (2005). Risk assessment and prediction of violent and sexual recidivism in sex offenders: long-term predictive validity of four risk assessment instruments. *Journal of Forensic Psychiatry and Psychology, 16*(1), 92-108. doi: 10.1080/1478994042000270247

Dignam, P., Brock, P., Hein, S., & McEvoy, P. (2010). Aggression and restraint in child inpatient units. *Australian and New Zealand Journal of Psychiatry, 44*(9), 866. doi: 10.1023/A:1005575113507

Dolan, M., & Doyle, M. (2000). Violence risk prediction - Clinical and actuarial measures and the role of the Psychopathy Checklist. *British Journal of Psychiatry, 177*(4), 303-311. doi:10.1192/ bjp.177.4.303

Dolan, M., & Doyle, M. (2000). Violence risk prediction. *The British Journal of Psychiatry, 177*(4), 303-311. doi: 10.1192/ bjp.177.4.303

Dolan, M., & Fullam, R. (2007). The validity of the Violence Risk Scale second edition (VRS-2) in a British forensic inpatient sample. *Journal of Forensic Psychiatry & Psychology, 18*(3), 381-393. doi: 10.1080/14789940701489390

Dolan, M., Fullam, R., Logan, C., & Davies, G. (2008). The Violence Risk Scale Second Edition (VRS-2) as a predictor of institutional violence in a British forensic inpatient sample. *Psychiatry Research, 158*(1), 55-65. doi: 10.1016/j.psychres.2006.08.014

Douglas, K. S., & Ogloff, J. R. (2003). The Impact of Confidence on the Accuracy of Structured Professional and Actuarial Violence Risk Judgments in a Sample of Forensic Psychiatric Patients. *Law and Human Behavior, 27*(6), 573-587. doi: 10.1023/B:LAHU. 0000004887.50905.f7

Douglas, K. S., & Ogloff, J. R. (2003). Violence by Psychiatric Patients: The Impact of Archival Measurement Source on Violence Base Rates and Risk Assessment Accuracy. *Canadian Journal of Psychiatry-Revue*

Canadienne De Psychiatrie, 48(11), 734-740. doi: 10.1177/070674370304801105

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk of violence: User guide.* Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., Nicholls, T. L., & Ogloff, J. R. P. (2004). Assessing risk for violence among male and female civil psychiatric patients: the HCR-20, PCL: SV, and VSC: research report. *Behavioral Sciences and the Law, 22*(1), 127-158. doi: 10.1002/bsl.579

Douglas, K. S., Ogloff, J. R., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology, 67*(6), 917-930. doi: 10.1037/0022-006X.67.6.917

Dowsett, J. (2005). Measurement of risk by a community forensic mental health team. *Psychiatric Bulletin, 29*(1), 9-12. doi: 10.1192/pb.29.1.9

Doyle, M., & Dolan, M. (2002). Violence risk assessment: combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing*, *9*(6), 649-657. doi: 10.1046/j.1365-2850.2002.00535.x

Doyle, M., & Logan, C. (2012). Operationalizing the assessment and management of violence risk in the short-term. *Behavioral Sciences & the Law, 30*(4), 406-419. doi:10.1002/bsl.2017

Doyle, M., Dolan, M., & McGovern, J. (2002). The validity of North

American risk assessment tools in predicting in-patient violent behaviour in England. *Legal and Criminological Psychology, 7*(2), 141-154. doi: 10.1348/135532502760274756

Doyle, M., Shaw, J., Carter, S., & Dolan, M. (2010). Investigating the validity of the classification of violence risk in a UK sample. *The International Journal of Forensic Mental Health, 9*(4), 316-323. doi: 10.1080/1499013.2010.527428

Drieschner, K. H., Marrozos, I., & Regenboog, M. (2013). Prevalence and risk factors of inpatient aggression by adults with intellectual disabilities and severe challenging behaviour: A long-term prospective study in two Dutch treatment facilities. *Research in Developmental Disabilities, 34*(8), 2407-2418. doi: 10.1016/ j.ridd.2013.04.008

Duffy, D., Doyle, M., & Ryan, T. (2004). Risk assessment and management in acute mental health care. In M. Harrison, D. Howard & D. Mitchell (Eds.) *Acute mental health nursing: From acute concerns to the capable practitioner* (pp. 144-161). Thousand Oaks, CA: Sage Publications, Inc.

Dumais, A., Larue, C., Michaud, C., & Goulet, M. H. (2012). Predictive validity and psychiatric nursing staff's perception of the clinical usefulness of the French version of the Dynamic Appraisal of Situational Aggression. *Issues in Mental Health Nursing, 33*(10), 670-675. doi: 10.3019/01612840.2012.697254

Eaton, S., Ghannam, M., & Hunt, N. (2000). Prediction of violence on a psychiatric intensive care unit. *Medicine Science and the Law, 40*(2), 143-146. doi: 10.1177/002580240004000210

Edens, J. F., Skeem, J. L., & Douglas, K. S. (2006). Incremental Validity Analyses of the Violence Risk Appraisal Guide and the Psychopathy Checklist: Screening Version in a Civil Psychiatric Sample. *Assessment, 13*(3), 368-374. doi: 10.1177/1073191105284001

Eriksen, B. M. S., Bjorkly, S., Faerden, 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), 186-197. doi: 10.1080/14999013.2016.1170740

Fabian, J. M. (2008). Current standards and practices in violence risk assessment at a maximum security forensic hospital following a high profile sexual homicide. *Aggression and Violent Behavior, 13*(5), 337-345. doi: 10.1016/j.avb.2008.05.002

Fehon, D. C., Grilo, C. M., & Lipschitz, D. S. (2005). A comparison of adolescent inpatients with and without a history of violence perpetration - Impulsivity, PTSD, and violence risk. *Journal of Nervous and Mental Disease, 193*(6), 405-411. doi: 10.1097/01.nmd.0000165294.41091.fc

Feigenbaum, J. D., Fonagy, P., Pilling, S., Jones, A., Wildgoose, A., & Bebbington, P. E. (2012). A real- world study of the effectiveness of DBT in the UK National Health Service. *British Journal of Clinical Psychology*, *51*(2), 121-141. doi: 10.1111/j.2044-8260.2011.02017.x

Ferguson, C. J., Averill, P. M., Rhoades, H., Rocha, D., Gruber, N. P., & Gummattira, P. (2005). Social isolation, impulsivity and depression as predictors of aggression in a psychiatric inpatient population.

Psychiatric Quarterly, 76(2), 123-137. doi: 10.1007/s11089-005-2335-1

Field, A. (2013). *Discovering statistics using IBM SPSS satistics.* Sage.

Ferris, L. E., Sandercock, J., Hoffman, B., Silverman, M., Barkun, H., Carlisle, J., & Katz, C. (1997). Risk assessments for acute violence to third parties: A review of the literature. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 42*(10), 1051-1060. doi: 10.1177/070674379704201006

Fluttert, F. A. J., Van Meijel, B., Bjorkly, S., Van Leeuwen, M., & Grypdonck, M. (2013). The investigation of early warning signs of aggression in forensic patients by means of the "Forensic Early Signs of Aggression Inventory'. *Journal of Clinical Nursing, 22*(11-12), 1550-1558. doi:10.1111/j.1365-2702.2012.04318.x

Fluttert, F. A. J., Van Meijel, B., Van Leeuwen, M., Bjorkly, S., Nijman, H., & Grypdonck, M. (2011). The Development of the Forensic Early Warning Signs of Aggression Inventory: Preliminary findings: Toward a Better Management of Inpatient Aggression. *Archives of Psychiatric Nursing, 25*(2), 129-137. doi:10.1016/j.apnu.2010.07.001

Fujii, D. E., Tokioka, A. B., Lichton, A. I., & Hishinuma, E. (2005). Ethnic Differences in Prediction of Violence Risk With the HCR-20 Among Psychiatric Inpatients. *Psychiatric Services, 56*(6), 711-716. doi: 10.1176/appi.ps.56.6.711

Fuller, J., & Cowan, J. (1999). Risk assessment in a multi-disciplinary forensic setting: Clinical judgement revisited. *Journal of Forensic Psychiatry, 10*(2), 276-289. doi: 10.1080/09585189908403681

Gammelgård, M., Koivisto, A-M., Eronen, M., & Kaltiala-Heino, R. (2008). The predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) among institutionalised adolescents. *The Journal of Forensic Psychiatry & Psychology, 19*(3), 352-370. doi:10.1080/14789940802114475

Gammelgård, M., Koivisto, A.-M., Eronen, M., & Kaltiala-Heino, R. (2015). Predictive validity of the structured assessment of violence risk in youth: A 4-year follow-up. *Criminal Behaviour and Mental Health: CBMH, 25*(3), 192-206. doi: 10.1002/cbm.1921

Gratz, K. L. (2006). Risk factors for deliberate self-harm among female college students: The role and interaction of childhood maltreatment, emotional inexpressivity, and affect intensity/reactivity. *American Journal of Orthopsychiatry*, *76*(2), 238-250. doi: 10.1037/0002-9432.76.2.238

Gray, N. S., Benson, R., Craig, R., Davies, H., Fitzgerald, S., Huckle, P., Maggs, R., Taylor, J., Trueman, M., Williams, T., & Snowden, R. J. (2011). The Short-Term Assessment of Risk and Treatability (START): A prospective study of inpatient behavior. *The International Journal of Forensic Mental Health, 10*(4), 305-313. doi: 10.1080/14999013.2011.631692

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *The British Journal of Psychiatry, 192*(5), 384-387. doi: 10.1192/bjp.bp.107.044065

Green, D., Schneider, M., Griswold, H., Belfi, B., Herrera, M., & DeBlasi, A. (2016). A Comparison of the HCR-20(V3) Among Male and Female Insanity Acquittees: A Retrospective File Study. *International*

Journal of Forensic Mental Health, 15(1), 48-64. doi:10.1080/14999013.2015.1134726

Grevatt, M., Thomas-Peter, B., & Hughes, G. (2004). Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient violence? *Journal of Forensic Psychiatry & Psychology, 15*(2), 278-292. doi: 10.1080/1478894032000199095

Griffith, J. J., Daffern, M., & Godber, T. (2013). Examination of the predictive validity of the Dynamic Appraisal of Situational Aggression in two mental health units. *International Journal of Mental Health Nursing, 22*(6), 485-492. doi: 10.1111/inm.12011

Grossi, L. M., Green, D., Belfi, B., McGrath, R. E., Griswold, H., & Schreiber, J. (2015). Identifying Aggression in Forensic Inpatients Using the MMPI-2-RF: An Examination of MMPI-2-RF Scale Scores and Estimated Psychopathy Indices. *International Journal of Forensic Mental Health, 14*(4), 231-244. doi:10.1080/14999013.2015.1108943

Gunenc, C., O'Shea, L. E., & Dickens, G. L. (2015). Prevalence and predictors of verbal aggression in a secure mental health service: Use of the HCR-20. *International Journal of Mental Health Nursing, 24*(4), 314-323. doi:10.1111/inm.12130

Gunn, J. (1996). Let's get serious about dangerousness. *Criminal Behaviour and Mental Health, 6*(1), 51-64. doi: 10.1002/cbm.136

Guy, L. S., Packer, I. K., & Warnken, W. (2012). Assessing Risk of Violence Using Structured Professional Judgment Guidelines. *Journal of Forensic Psychology Practice, 12*(3), 270-283. doi:10.1080/

15228932.2012.674471

Hájek, A., & Hall, N. (2002). Induction and probability. In P. Machamer & M. Silberstein (Eds.), The *blackwell guide to the philosophy of science* (pp. 149-172). Oxford, UK: Blackwell Publishers

Hamilton, J. R., & Bullard, H. (1990). Dangerousness: Which patients should we worry about? In *Dilemmas and difficulties in the management of psychiatric patients* (pp. 117-129). New York, NY: Oxford University Press; US.

Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving Actuarial Risk Assessments for Sex Offenders*. Ottawa, Canada: Ministry of the Solicitor General

Hare, R. D. (1999). Psychopathy as a risk factor for violence. *Psychiatric Quarterly, 70*(3), 181-197. doi: 10.1023/A:1022094925150

Harris, A., & Lurigio, A. J. (2007). Mental illness and violence: A brief review of research and assessment strategies. *Aggression and Violent Behavior, 12*(5), 542-551. doi:10.1016/j.avb.2007.02.008

Harris, G. T., & Rice, M. E. (1994). *The Violent Patient: Handbook of Prescriptive Treatments for Adults* (pp. 463-486). New York, NY: Plenum Press; US.

Harris, G. T., & Rice, M. E. (1997). Risk appraisal and management of violent behavior. *Psychiatric Services, 48*(9), 1168-1176. Retrieved from http://www.link.springer.com

Hart, S. D., Hare, R. D., & Cox, D. (1995). *The Hare Psychopathy Checklist: Screening Version (PCL:SV).* Toronto, Ontario, Canada: Multi-Health Systems

Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk

assessment instruments: Evaluating the 'margins of error' of group v. individual predictions of violence. *British Journal of Psychiatry, 190*(49), 60-65. doi: 10.1192/bjp.190.5s60

Hart, S. D., Webster, C. D., & Menzies, R. J. (1993). A note on portraying the accuracy of violence predictions. *Law and Human Behavior, 17*(6), 695-700. doi:10.1007/bf01044690

Hartvig, P., Roaldset, J., Moger, T., Ostberg, B., & Bjorkly, S. (2011). The first step in the validation of a new screen for violence risk in acute psychiatry: The inpatient context. *European Psychiatry, 26*(2), 92-99. doi: 10.1016/j.eurpsy.2010.01.003

Harvey, J., and Smedley, K. (2010). *Psychological Therapy in Prisons and Other Secure Settings. New York:* Willian Publishing

Hawley, C. J., Littlechild, B., Sivakumaran, T., Sender, H., Gale, T. M., & Wilson, K. J. (2006). Structure and content of risk assessment proformas in mental healthcare. *Journal of Mental Health, 15*(4), 437-448. doi: 10.1080/09638230600801462

Heilbrun, K., Hart, S. D., Hare, R. D., Gustafson, D., Nunez, C., & White, A. J. (1998). Inpatient and postdischarge aggression in mentally disordered offenders - The role of psychopathy. *Journal of Interpersonal Violence, 13*(4), 514-527. doi: 10.1177/088626098013004007

Hildebrand, M., de Ruiter, C., & Nijman, H. (2004). PCL-R psychopathy predicts disruptive behavior among male offenders in a Dutch forensic psychiatric hospital. *Journal of Interpersonal Violence, 19*(1), 13-29. doi: 10.1177/0886260503259047

Hilton, N. Z., Harris, G. T., & Rice, M. E. (2010). *Risk assessment for criminal justice, offender intervention, and victim services*. Washington, DC: American Psychological Association; US.

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. doi: 10.1037/lhb0000179

Holloway, F. (2004). Risk: More questions than answers. *Advances in Psychiatric Treatment*, *10*(4), 273-274. doi: 10.1192/apt.10.4.273

Hoptman, M. J., Yates, K. F., Patalinjug, M. B., Wack, R. C., & Convit, A. (1999). Clinical prediction of assaultive behavior among male psychiatric patients at a maximum-security forensic facility. *Psychiatric Services, 50(11),* 1461-1466. doi: 10.1176/ps.50.11.1461

Hornsveld, R. H. J., Kraaimaat, F. W., Bouwmeester, S., Polak, M. A., & Zwets, A. J. (2014). Behavior on the ward of personality-disordered inpatients and chronically psychotic inpatients during a three-year stay in a Dutch forensic psychiatric hospital. *Journal of Forensic Psychiatry & Psychology, 25*(6), 684-704. doi:10.1080/14789949.2014.947308

Hornsveld, R. H. J., Van Dam-Baggen, C. M. J., Lammers, S. M. M., Nijman, H. L. I., & Kraaimaat, F. W. (2004). Violent forensic patients: Personality traits and behavior. *Tijdschrift voor Psychiatrie, 46*(3), 133-143. Retrieved from http://www.researchgate.net

Howe, J., Rosenfeld, B., Foellmi, M., Stern, S., & Rotter, M. (2016). Application of the HCR-20 version 3 in civil psychiatric patients. *Criminal Justice and Behavior, 43*(3), 398-412. doi: 10.1177/0093854815605527

Huchzermeier, C., Brus, E., Geiger, F., Kernbichler, A., & Aldenhoff, J. (2008). Predictive validity of the psychopathy checklist: Screening version for intramural behaviour in violent offenders - A prospective study at a secure psychiatric hospital in Germany. *The Canadian Journal of Psychiatry / La Revue Canadienne de Psychiatrie, 53*(6), 384-391. doi: 10.1177/070674370805300608

Ideker, K., Todicheeney-Mannes, D., & Kim, S. C. (2011). A confirmatory study of violence risk assessment tool (M55) and demographic predictors of patient violence. *Journal of Advanced Nursing, 67*(11), 2455-2462. doi: 10.1111/j.1365-2648.2011.05667.x

Inett, A., Wright, G., Roberts, L., & Sheeran, A. (2014). Predictive validity of the START with intellectually disabled offenders. *Journal of Forensic Practice, 16*(1), 78-88. doi: 10.1108/JFP-12-2012-0029

Iozzino, L., Ferrari, C., Large, M., Nielssen, O., & de Girolamo, G. (2015). Prevalence and Risk Factors of Violence by Psychiatric Acute Inpatients: A Systematic Review and Meta-Analysis. *PLoS ONE, 10*(6). doi:10.1371/journal.pone.0128536

Ivgi, D., Bauer, A., Khawaled, R., Rosca, P., Weiss, J. M., & Ponizovsky, A. M. (2015). Validation of the hcr-20 scale for assessing risk of violent behavior in Israeli psychiatric inpatients. *Israel Journal of Psychiatry and Related Sciences, 52*(2), 121-128. Retrieved from http://www.europepmc.org

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12-19. doi: 10.1037/0022-006X.59.1.12

Johnson, M. C., & Morris, R. G. (2008). The moderating effects of religiosity on the relationship between stressful life events and delinquent behavior. *Journal of Criminal Justice*, *36*(6), 486-493. doi: 10.1016/j.jcrimjus.2008.09.001

Johnstone, L., & Dallos, R. (2014). *Formulation in psychology and psychotherapy: Making sense of people's problems.* East Sussex, UK: Routledge

Jung, S., Ledi, D., & Daniels, M. K. (2013). Evaluating the concurrent validity of the HCR-20 scales. *Journal of Risk Research, 16*(6), 697-711. doi: 10.1080/13669877.2012.726247

Kasinathan, J., & Marsland, C. (2014). Imminent aggression risk in adolescents: Is it different from adults? *Australian and New Zealand Journal of Psychiatry, 48*(1), 26. Retrieved from http://www.psycnet.apa.org

Kasinathan, J., Marsland, C., Batterham, P., Gaskin, C., Adams, J., & Daffern, M. (2015). Assessing the risk of imminent aggression in mentally ill young offenders. *Australasian Psychiatry, 23*(1), 44-48. doi: 10.1177/1039856214563845

Kay, S. R., Wolkenfeld, F., & Murrill, L. M. (1988). Profiles of aggression among psychiatric patients: II. Covariates and predictors. *Journal of Nervous and Mental Disease, 176*(9), 547-557. doi: 10.1097/00005053-198809000-00008

Ketelsen, R., Zechert, C., Driessen, M., & Schulz, M. (2007). Characteristics of aggression in a German psychiatric hospital and predictors of patients at risk. *Journal of Psychiatric and Mental Health Nursing, 14*(1), 92-99. doi: 10.1111/j.1365-2850.2007.01049.x

Khiroya, R., Weaver, T., & Maden, T. (2009). Use and perceived utility of structured violence risk assessments in English medium secure forensic unit. *Psychiatric Bulletin, 33*(4), 129-132. doi: 10.1192/pb.bp.108.019810

Kim, S. C., Ideker, K., & Todicheeney-Mannes, D. (2012). Usefulness of aggressive behaviour risk assessment tool for prospectively identifying violent patients in medical and surgical units. *Journal of Advanced Nursing, 68*(2), 349-357. doi: 10.1111/j.1365-2648.2011.05744.x

Kingston, M. J., Evans, S. M., Smith, B. J., & Berry, J. G. (2004). Attitudes of doctors and nurses towards incident reporting: a qualitative analysis. *Medical Journal of Australia*, *181(*1), 36-39. Retrieved from http://www.mja.com.au

Klassen, D., & O'Connor, W. A. (1988). A prospective study of predictors of violence in adult male mental health admissions. *Law and Human Behavior, 12*(2), 143-158. doi: 10.1007/BF01073122

Kling, R., Corbiere, M., Milord, R., Morrison, J. G., Craib, K., Yassi, A., Sidebottom, C., Kidd, C., Long, V., & Saunders, S. (2006). Use of a violence risk assessment tool in an acute care hospital: effectiveness in identifying violent patients. *AAOHN Journal, 54*(11), 481-487. doi: 10.1177/216507990605401102

Kobes, M., Nijman, H., & Bulten, E. B. H. (2012). Assessing Aggressive Behavior in Forensic Psychiatric Patients: Validity and Clinical Utility of Combining Two Instruments. *Archives of Psychiatric Nursing, 26*(6), 487-494. doi:10.1016/j.apnu.2012.04.004

Krakowski, M., & Czobor, P. (2004). Gender differences in violent behaviors: Relationship to clinical symptoms and psychosocial factors. *American Journal of Psychiatry, 161*(3), 459-465. doi:10.1176/appi.ajp.161.3.459

Kruger, C., & Rosema, D. (2010). Risk factors for violence among long-term psychiatric in-patients: A comparison between violent and non-violent patients. *African Journal of Psychiatry, 13*(5), 366-375. doi: 10.4314/ajpsy.v13i5.63102

Lamberg, L. (2007). New tools aid violence risk assessment. *Journal of the American Medical Association, 298*(5), 499-501. doi: 10.1001/jama.298.5.499

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159-174. doi: 10.2307/2529310

Langan, J. (2010). Challenging assumptions about risk factors and the role of screening for violence risk in the field of mental health. *Health, Risk and Society, 12*(2), 85-100. doi: 10.1080/13698571003632429

Langton, C. M., Barbaree, H. E., Hansen, K. T., Harkins, L., & Peacock, E. J. (2007). Reliability and Validity of the Static-2002 Among Adult Sexual Offenders With Reference to Treatment Status. *Criminal Justice and Behavior, 34*(5), 616-640. doi:10.1177/0093854806296851

Langton, C. M., Hogue, T. E., Daffern, M., Mannion, A., & Howells, K. (2009). Prediction of institutional aggression among personality disordered forensic patients using actuarial and structured clinical risk assessment tools: prospective evaluation of the HCR-20, VRS, Static-99, and Risk Matrix 2000. *Psychology Crime & Law, 15*(7), 635-659. doi:10.1080/10683160802516224

Langton, C. M., Hogue, T. E., Daffern, M., Mannion, A., & Howells, K. (2011). Personality Traits as Predictors of Inpatient Aggression in a High-Security Forensic Psychiatric Setting: Prospective Evaluation of the PCL-R and IPDE Dimension Ratings. *International Journal of Offender Therapy and Comparative Criminology, 55*(3), 392-415. doi:10.1177/0306624x10370828

Lantta, T., Kontio, R., Daffern, M., Adams, C. E., & Valimaki, M. (2016). Using the dynamic appraisal of situational aggression with mental health inpatients: A feasibility study. *Patient Preference and Adherence, 10*, 691-701. doi: 10.2147/PPA.S103840

Large, M. (2010). Accuracy and risk assessment. *Australasian Psychiatry, 18*(3), 267. Retrieved from http://www.journals.sagepub.com

Laurell, J., Belfrage, H., & Hellstrom, A. (2010). Facets on the psychopathy checklist screening version and instrumental violence in forensic psychiatric patients. *Criminal Behaviour and Mental Health, 20*(4), 285-294. doi: 10.1002/cbm.779

Lewis, A. H. O., & Webster, C. D. (2004). General instruments for risk assessment. *Current Opinion in Psychiatry, 17*(5), 401-405. Retrieved from http://www.journals.lww.com

Lincoln, T., Hodgins, S., Jockel, D., Freese, R., Born, R., Eucker, S.,

Schmidt, P., Gretenkord, L., & Muller-Isberner, R. (2006). Patients in forensic and general psychiatry: Do risk assessment measures detect differing risks of future violence? *Nervenarzt, 77*(5), 576-586. doi:10.1007/s00115-005-1918-5

Lindsay, W. R., Hogue, T. E., Taylor, J. L., Steptoe, L., Mooney, P., O'Brien, G., Johnston, S., & Smith, A. H. W. (2008). Risk Assessment in Offenders With Intellectual Disability: A comatison across three security levels. *International Journal of Offender Therapy and Comparative Criminology, 52*(1), 90-111. doi:10.1177/0306624X07308111

Linehan, M. (1993). *Manual for treating borderline personality disorder*. New York: Guilford Press

Littlechild, B., & Hawley, C. (2010). Risk assessments for mental health service users ethical, valid and reliable? *Journal of Social Work, 10*(2), 211-229. doi:10.1177/1468017309342191

Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, T. A. H. (2010). The impact of protective factors in desistance from violent reoffending: A study in three samples of adolescent violent offenders. *Journal of Interpersonal Violence, 25*(3), 568–587. doi: 10.1177/0886260509334403

Lussier, P., Verdun-Jones, S., Deslauriers-Varin, N., Nicholls, T., & Brink, J. (2010). Chronic violent patients in an inpatient psychiatric hospital: Prevalence, description, and identification. *Criminal Justice and Behavior, 37*(1), 5-28. Retrieved from http://www.journals.sagepub.com

Macpherson, G. J., & Kevan, I.-M. (2004). Predictive validity of the hcr-

20 violence risk assessment scheme within a maximum security special hospital. *Issues in Forensic Psychology, 5*, 62-80. Retrieved from http://www.journals.sagepub.com

Mao, R., Hamed, A. B. A., Goh, A. S., & Wen, D. W. (2012). Clinical application of the broset violence checklist to predict inpatient violence in a tertiary psychiatric hospital in singapore: A preliminary study. *Annals of the Academy of Medicine Singapore, 1*(10), 17-21. Retrieved from http://www.tandfonline.com

Margari, F., Matarazzo, R., Casacchia, M., Roncone, R., Dieci, M., Safran, S., Fiori, G., & Simoni, L. (2005). Italian validation of MOAS and NOSIE: a useful package for psychiatric assessment and monitoring of aggressive behaviours. *International Journal of Methods in Psychiatric Research, 14*(2), 109-118. doi: 10.1002/mpr.22

Martinaki, S., Tsopelas, C., Ploumpidis, D., Douzenis, A., Tzavara, H., Skapinakis, P., & Mavreas, V. (2013). Evaluation of dangerousness of Greek mental patients. *Psychiatriki, 24*(3), 185-196. Retrieved from http://www.europepmc.org

Martini, A., Fantini, S., D'Ovidio, M. C., Ceracchi, A., & De Santis, A. (2012). Risk assessment of aggression toward emergency health care workers. *Occupational Medicine-Oxford, 62*(3), 223-225. doi:10.1093/occmed/kqr199

Matusiewicz, A.K., Hopwood, C.J., Banducci, A.N., and Lejuez, C.W. (2010). The effectiveness of cognitive behavioural therapy for personality disorders. *Psychiatric Clinic of North America, 33*(3), 657-685. doi: 10.1016/j.psc.2010.04.007

McDermott, B. E., & Holoyda, B. J. (2014). Assessment of aggression in

inpatient settings. *CNS Spectrums, 19*(5), 425-431. doi:10.1017/s1092852914000224

McDermott, B. E., Edens, J. F., Quanbeck, C. D., Busse, D., & Scott, C. L. (2008). Examining the role of static and dynamic risk factors in the prediction of inpatient violence: Variable- and person-focused analyses. *Law and Human Behavior, 32*(4), 325-338. doi: 10.1007/s10979-007-9094-8

McDermott, B. E., Quanbeck, C. D., Busset, D., Yastro, K., & Scott, C. L. (2008). The accuracy of risk assessment instruments in the prediction of impulsive versus predatory aggression. *Behavioral Sciences & the Law, 26*(6), 759-777. doi:10.1002/bsl.842

McKay, D., Gavigan, C. A., & Kulchycky, S. (2004). Social skills and sex-role functioning in borderline personality disorder: relationship to self- mutilating behavior. *Cognitive Behaviour Therapy*, *33*(1), 27-35. doi: 10.1080/16506070310002199

McKay, M., Wood, J.C., and Brantley, J. (2007) *The dialectical behaviour therapy skill workbook: Practical exercises for learning mindfulness, interpersonal effectiveness, emotion regulation and distress tolerance.* Oakland C.A: New Harbour Publications Inc.

McKelvey, J. S. (2013). Impulsivity, cognitive insight and prediction of aggression in psychiatric inpatients. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74*(6-B)

McKenzie, B., & Curr, H. (2005). Predicting violence in a medium secure setting: A study using the historical and clinical scales of the HCR-20. *The British Journal of Forensic Practice, 7*(3), 22-28. doi:

10.1108/14636646200500018

McMillan, D., Hastings, R. P., & Coldwell, J. (2004). Clinical and actuarial prediction of physical violence in a forensic intellectual disability hospital: A longitudinal study. *Journal of Applied Research in Intellectual Disabilities, 17*(4), 255-265. doi: 10.1111/j.1468-3148.2004.00213.x

McNiel, D. E., & Binder, R. L. (1991). Clinical assessment of the risk of violence among psychiatric inpatients. *American Journal of Psychiatry, 148*(Oct 91), 1317-1321. Retrieved from http://www.search.proquest.com

McNiel, D. E., & Binder, R. L. (1994). Screening for risk of inpatient violence: Validation of an actuarial tool. *Law and Human Behavior, 18*(5), 579-586. doi: 10.1007/BF01499176

McNiel, D. E., & Binder, R. L. (1995). Correlates of accuracy in the assessment of psychiatric inpatients' risk of violence. *The American Journal of Psychiatry, 152*(6), 901-906. Retrieved from http://www.search.proquest.com

McNiel, D. E., & Binder, R. L. (1996). "Predicting inpatient violence": Reply. *The American Journal of Psychiatry, 153*(6), 845-846. Retrieved from http://www.ncbi.nlm.nih.gov

McNiel, D. E., Gregory, A. L., Lam, J. N., Binder, R. L., & Sullivan, G. R. (2003). Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology, 71*(5), 945-953. doi: 10.1037/0022-006X.71.5.945

McNiel, D. E., Lam, J. N., & Binder, R. L. (2000). Relevance of interrater agreement to violence risk assessment. *Journal of Consulting and*

Clinical Psychology, 68(6), 1111-1115. doi: 10.1037/0022-006X.68.6.1111

McNiel, D. E., Sandberg, D. A., & Binder, R. L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human Behavior, 22*(6), 655-669. doi: 10.1023/A:1025754706716

Meaden, A., Hacker, D., & Spencer, K. (2013). Acute aggression risk: An early warning signs methodology. *Journal of Forensic Practice, 15*(1), 21-31. doi: 10.1108/14636641311299059

Menzies, R., & Webster, C. D. (1995). Construction and validation of risk assessments in a 6-year follow-up of forensic patients - a tridimensional analysis. *Journal of Consulting and Clinical Psychology, 63*(5), 766-778. doi:10.1037/0022-006x.63.5.766

Miller, H. A. (2006). A dynamic assessment of offender risk, needs, and strengths in a sample of general offenders. *Behavioral Sciences and the Law, 24*(6), 767-782. doi: 10.1002.bsl.728

Miller, H. A. (2006). *Manual of the Inventory of Offender Risk, Needs, and Strengths (IORNS).* Odessa, FL: Psychological Assessment Resources

Miller, M. C. (2000). A model for the assessment of violence. *Harvard Review of Psychiatry, 7*(5), 299-304. doi: 10.3109/hrp.7.5.299

Monahan, J. (1997). Actuarial support for the clinical assessment of violence risk. *International Review of Psychiatry, 9*(2), 167-169. doi: 10.1080/09540269775358

Monahan, J., Steadman, H. J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Silver, E., Appelbaum, P. S., Grisso, T., Mulvey, E. P., & Roth,

L. H. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *The British Journal of Psychiatry, 176*(4), 312-319. doi: 10.1192/bjp.176.4.312

Morrissey, C., Hogue, T., Mooney, P., Allen, C., Johnston, S., Hollin, C., Lindsay, W. R., & Taylor, J. L. (2007). Predictive validity of the PCL-R in offenders with intellectual disability in a high secure hospital setting: Institutional aggression. *Journal of Forensic Psychiatry and Psychology, 18*(1), 1-15. doi: 10.1080/08990220601116345

Morrissey, C., Mooney, P., Hogue, T. E., Lindsay, W. R., & Taylor, J. L. (2007). Predictive validity of the PCL-R for offenders with intellectual disability in a high security hospital: Treatment progress. *Journal of Intellectual and Developmental Disability, 32*(2), 125-133. doi: 10.1080/0898446231570018

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consluting and Clincal Psychology, 62*(4), 783-792. doi: 10.1037/0022-006X.62.4.783

Motamedi, B. (2011). The frequency and types of aggressive and violent behaviors in acute psychiatric inpatient settings. *European Child and Adolescent Psychiatry, 20*, 135-142. doi: 10.1176/ps.49.11.1452

Motamedi, B. (2012). Violence in psychiatric wards. *European Psychiatry. Conference: 20th European Congress of Psychiatry, EPA, 27*

Mudde, N., Nijman, H., van der Hulst, W., & van den Bout, J. (2011). Predicting aggression during the treatment of forensic psychiatric patients by means of the HCR-20. *Tijdschrift voor Psychiatrie, 53*(10), 705-713. Retrieved from http://www.europepmc.org

Mulvey, E. P., & Lidz, C. W. (1995). Conditional prediction: A model for research on dangerousness to others in a new era. *International Journal of Law and Psychiatry, 18*(2), 129-143. doi: 10.1016/0160-2527(95)00002-Y

Nagi, C., Ostapiuk, E. B., Craig, L. A., Hacker, D., & Beech, A. R. (2009). Using the revised Problem Identification Checklist to predict inpatient and community violence: A pilot study. *The British Journal of Forensic Practice, 11*(3), 4-13. doi: 10.1108/14636646200900017

Nanayakkara, V., O'Driscoll, C., & Allnutt, S. (2012). HCR-20 Risk Profiles of Patients Referred to Forensic Mental Health Services in New South Wales. *Psychiatry Psychology and Law, 19*(5), 711-719. doi:10.1080/13218719.2011.615819

National Collaborating Centre for Mental Health. (2009). *Borderline Personality Disorder: Treatment and Management*. London, England: The British Psychological Society & The Royal College of Psychiatrists.

Neal, T. M. S., Miler, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior, 42*(9), 952-968. doi:10.1177/0093854815572252

Neufeld, E., Perlman, C. M., & Hirdes, J. P. (2012). Predicting inpatient aggression using the InterRAI risk of harm to others clinical assessment protocol: a tool for risk assessment and care planning. *Journal of Behavioral Health Services & Research, 39*(4), 472-480. doi:http://dx.doi.org/10.1007/s11414-011-9271-x

Nicholls, T. L. (2004). Violence risk assessments with female NCRMD

acquittees: Validity of the HCR-20 and PCL-SV. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*(8), 4055. Retrieved from http://www.psycnet.apa.org

Nicholls, T. L., Brink, J., Desmarais, S. L., Webster, C. D., & Martin, M. L. (2006). The Short-Term Assessment of Risk and Treatability (START): A prospective validation study in a forensic psychiatric sample. *Assessment*, *13*(3), 313-327. doi: 10.1177/1073191106290559

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing risk for violence among male and female civil psychiatric patients: The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law, 22*(1), 127-158. doi: 10.1002/bsl.579

Nijman, H., Merckelbach, H., Evers, C., Palmstierna, T., & Campo, J. A. (2002). Prediction of aggression on a locked psychiatric admissions ward. *Acta Psychiatrica Scandinavica, 105*(5), 390-395. doi:10.1034/j.1600-0447.2002.0o426.x

Nonstad, K., Nesset, M. B., Kroppan, E., Pedersen, T. W., Nottestad, J. A., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norwegian high secure hospital. *The International Journal of Forensic Mental Health, 9*(4), 294-299. doi: 10.1080/14999013.2010.534958

Nonstad, K., Nesset, M. B., Kroppan, E., Pederson, T. W., Nøttestad, J. A., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norweigian

high secure hospital. *International Journal of Forensic Mental Health, 9*(4), 294-299. doi: 10.1080/14999013.2010.534958

O'Shea, L. E., & Dickens, G. L. (2014). Short-Term Assessment of Risk and Treatability (START): Systematic Review and Meta-Analysis. *Psychological Assessment, 26*(3), 990-1002. doi:10.1037/a0036794

O'Shea, L. E., & Dickens, G. L. (2015). Predictive validity of the Short-Term Assessment of Risk and Treatability (START) for aggression and self-harm in a secure mental health service: Gender differences. *The International Journal of Forensic Mental Health, 14*(2), 132-146. doi: 10.1080/14999013.2015.1033112

O'Shea, L. E., & Dickens, G. L. (2015). Predictive validity of the START for unauthorised leave and substance abuse in a secure mental health setting: A pseudo-prospective cohort study. *International Journal of Nursing Studies, 52*(5), 970-979. doi: 10.1016/ j.ijnurstu.2015.02.007

O'Shea, L. E., & Dickens, G. L. (2015). The HCR-20 as a measure of reliable and clinically significant change in violence risk among secure psychiatric inpatients. *Comprehensive Psychiatry, 62*, 132-140. doi: 10.1016/j.comppsych.2015.07.009

O'Shea, L. E., & Dickens, G. L. (2016). Role of assessment components and recent adverse outcomes in risk estimation and prediction: Use of the Short Term Assessment of Risk and Treatability (START) in an adult secure inpatient mental health service. *Psychiatry Research, 240*, 398-405. doi: 10.1016/j.psychres.2016.04.068

O'Shea, L. E., Mitchell, A. E., Picchioni, M. M., & Dickens, G. L. (2013). Moderators of the predictive efficacy of the Historical, Clinical and

Risk Management-20 for aggression in psychiatric facilities: Systematic review and meta-analysis. *Aggression and Violent Behavior, 18*(2), 255-270. doi: 10.1016/j.avb.2012.11.016

O'Shea, L. E., Picchioni, M. M., & Dickens, G. L. (2016). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) for multiple adverse outcomes in a secure psychiatric inpatient setting. *Assessment, 23*(2), 150-162. doi: 10.1177/1073191115573301

O'Shea, L. E., Picchioni, M. M., Mason, F. L., Sugarman, P. A., & Dickens, G. L. (2014). Differential predictive validity of the Historical, Clinical and Risk Management Scales (HCR-20) for inpatient aggression. *Psychiatry Research, 220*(1-2), 669-678. doi: 10.1016/j.psychres.2014.07.080

O'Shea, L. E., Picchioni, M. M., Mason, F. L., Sugarrnan, P. A., & Dickeris, G. L. (2014). Predictive validity of the HCR-20 for inpatient self-harm. *Comprehensive Psychiatry, 55*(8), 1937-1949. doi:10.1016/j.comppsych.2014.07.010

O'Shea, L. E., Picchioni, M. M., McCarthy, J., Mason, F. L., & Dickens, G. L. (2015). Predictive validity of the HCR-20 for inpatient aggression: The effect of intellectual disability on accuracy. *Journal of Intellectual Disability Research, 59*(11), 1042-1054. doi: 10.1111/jir.12184

O'Shea, L., & Dickens, G. (2015). Contribution of protective factors assessment to risk prediction: Systematic review and meta-analysis. *European Psychiatry, 30*, 211. doi: 10.1016/S0924-9338(15)30171-1

O'Shea, L., Picchioni, M., McCarthy, J., Mason, F., & Dickens, G. (2015). Predictive validity of the HCR-20 for inpatient aggression: The effect of intellectual disability on accuracy. *Journal of Intellectual Disability Research*, *59*(11), 1042-1054. doi: 10.1111/jir.12184

O'Shea, L. E., Picchioni, M. M., & Dickens, G. L. (2015). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) for multiple adverse outcomes in a secure psychiatric inpatient setting. *Assessment*, *23*(2), 150-162. doi: 10.1177/1073191115573301

Ogloff, J. R., & Daffern, M. (2006). The dynamic appraisal of situational aggression: An instrument to assess risk for imminent aggression in psychiatric inpatients. *Behavioral Sciences & the Law, 24*(6), 799-813. doi: 10.1002/bsl.741

Otto, R. K., & Douglas, K. S. (Eds.). (2011). *Handbook of violence risk assessment*. New York: Routledge.

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports, 10*(3), 799-812. doi: 10.2466/pr0.1962.10.3.799

Owen, C., Tarantello, C., Jones, M., & Tennant, C. (1998). Violence and aggression in psychiatric units. *Psychiatric Services, 49*(11), 1452-1457. doi: 10.1176/ps.49.11.1452

Paivio, S. C., & McCulloch, C. R. (2004). Alexithymia as a mediator between childhood trauma and self-injurious behaviors. *Child Abuse & Neglect*, *28*(3), 339-354. doi: 10.1016/j.chiabu.2003.11.018

Palmstierna, T., & Olsson, D. (2007). Violence from young women

involuntarily admitted for severe drug abuse. *Acta Psychiatrica Scandinavica, 115*(1), 66-72. doi: 10.1111/j.1600-0447.2006.00853.x

Patel, K. (2012). Talk 2 - Violence risk assessment in forensic psychiatry: Guesswork or science? *Indian Journal of Psychiatry, 54*, S10. doi: 10.1590/S0101-60832004000600002

Penney, S. R., McMaster, R., & Wilkie, T. (2014). Multirater reliability of the historical, clinical, and risk management-20. *Assessment, 21*(1), 15-27. doi:10.1177/1073191113514107

Penterman, B., & Nijman, H. (2011). Assessing aggression risks in patients of the ambulatory mental health crisis team. *Community Mental Health Journal, 47*(4), 463-471. doi: 10.1007/s10597-010-9348-7

Penterman, E., & Nijman, H. (2009). Assessing aggression in patients of the mental health crisis service. *Tijdschrift voor Psychiatrie, 51*(6), 355-364. Retrieved from http://www.europepmc.org

Penterman, E., Nijman, H., Saalmink, K., Rasing, S., & van der Staak, C. (2013). Assessing aggressive behaviour at the psychiatric emergency service with a checklist: A replication study. *Tijdschrift voor Psychiatrie, 55*(2), 93-100. Retrieved from http://www.europepmc.org

Pepe, M. S., Janes, H., Longton, G., Leisenring, W., & Newcomb, P. (2004). Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology*, *159*(9), 882-890. doi: 10.1093/aje/kwh101

Peters, M. R. (1998). Anger as a predictor of aggressive behavior in an adolescent inpatient setting. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 59*(2-B), 883.

Pham, T., Remy, S., Dailliet, A., & Lienard, L. (1998). Psychopathy and assessment of violent behaviors in a security hospital. *L'Encephale: Revue de psychiatrie clinique biologique et therapeutique, 24*(3), 173-179.

Phillips, N. L., Stargatt, R., & Brown, A. (2012). Risk assessment of self- and other-directed aggression in adolescent psychiatric inpatient units. *Australian and New Zealand Journal of Psychiatry, 46*(1), 40-46. doi: 10.1177/0004867411430876

Phillips, N. L., Stargatt, R., & Fisher, L. (2011). Risk assessment: Predicting physical aggression in child psychiatric inpatients units. *Australian and New Zealand Journal of Psychiatry, 45*(8), 638-645. Retrieved from http://www.tandfonline.com

Plutchik, R., Van Praag, H. M., Conte, H. R., & Picard, S. (1989). Correlates of suicide and violence risk: I. The suicide risk measure. *Comprehensive Psychiatry, 30*(4), 296-302. doi: 10.1016/0010-440X(89)90053-9

Pomp, L., Spreen, M., Bogaerts, S., & Volker, B. (2010). The role of personal social networks in risk assessment and management of forensic psychiatric patients. *Journal of Forensic Psychology Practice, 10*(4), 267-284. doi: 10.1080/15228932.2010.481232

Quinn, R., Miles, H., & Kinane, C. (2013). The validity of the Short-Term Assessment of Risk and Treatability (START) in a UK medium secure forensic mental health service. *The International Journal of Forensic*

Mental Health, 12(3), 215-224. doi: 10.1080/14999013.213.832714

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and Managing Risk.* Washington, DC: American Psychological Association

Rabinowitz, J., & Garelik-Wyler, R. (1999). Accuracy and confidence in clinical assessment of psychiatric inpatients risk of violence. *International Journal of Law and Psychiatry, 22*(1), 99-106. doi: 10.1016/S0160-2527(98)00032-6

Rechenmacher, J., Muller, G., Abderhalden, C., & Schulc, E. (2014). The diagnostic efficiency of the extended German Broset Violence Checklist to assess the risk of violence. *Journal of Nursing Measurement, 22*(2), 201-212. doi: 10.1891/1061-3749.22.2.201

Reimann, B. J. (2009). Predicting institutional aggression and seclusion within a medium secure forensic hospital using the Psychopathy Checklist-revised (PCL-R) and the Violence Risk Appraisal Guide (VRAG). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 69*(9-B), 5790. Retrieved from http://www.search.proquest.com

Reimann, B. J., & Nussbaum, D. (2011). Predicting seclusion in a medium secure forensic inpatient setting. *The International Journal of Forensic Mental Health, 10*(2), 150-158. Retrieved from http://www.tandfonline.com

Reynolds, G. P., McKelvey, J. S., Reinharth, J., Payne, E. B., Tropper, A., Selig, P., Malhotra, A., Russ, M., & Serper, M. R. (2013). Predictors of persistent aggression on the psychiatric inpatient service. *Comprehensive Psychiatry, 54 (8)*, e34. Retrieved from

http://www.search.proquest.com

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. doi: 10.1007/s10979-005-6832-7

Rizzo, M., & Smith, S. (2012). Risk assessment and risk management among severely mentally ill patients requiring hospitalisation. *Journal of Forensic Psychiatry & Psychology, 23*(4), 497-509. doi:10.1080/14789949.2012.704638

Roaldset, J. O. (2015). Risk assessment and clinical decision-making. *Australian and New Zealand Journal of Psychiatry, 49*(1), 90. Retrieved from http://www.journals.sagepub.com

Roaldset, J. O., & Bjorkly, 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*(1), 153-159. doi: 10.1016/j.psychres.2010.04.012

Roaldset, J. O., Bjorkly, S., & Gotestam, K. G. (2009). Patients' self-report as predictor of self-harm and violence in an acute psychiatric unit. A prospective naturalistic study. *European Psychiatry, 24*(2), 153-159. doi: 10.1016/S0924-9338(09)71199-X

Roaldset, J. O., Hartvig, P., & Bjorkly, S. (2011). V-RISK-10: Validation of a screen for risk of violence after discharge from acute psychiatry. *European Psychiatry, 26*(2), 85-91. doi: 10.1016/j.europsy. 2010.04.002

Roaldset, J. O., Hartvig, P., Linaker, O. M., & Bjorkly, S. (2012). A multifaceted model for risk assessment of violent behaviour in

acutely admitted psychiatric patients. *Psychiatry Research, 200*(2-3), 773-778. doi: 10.1016/j.psychres.2012.04.038

Rogers, D., & Najarian, B. (1989). The construction and validation of a new scale for measuring emotion control. *Personality and Individual Differences, 10*(8), 845-853. doi: 10.1016/0191-8869(89)90020-2

Rogers, D., Jarvis, G., & Najarian, B. (1993). Detachment and coping: The construction and validation of a new scale for measuring coping strategies. *Personality and Individual Differences, 15*(6), 619-626. doi: 10.1016/0191-8869(93)90003-L

Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior, 24*(5), 595-605. doi: 10.1023/A:1005575113507

Rosenfeld, B., & Penrod, S. D. (2011). *Research Methods In Forensic Psychology.* Hoboken, New Jersey: John Wiley & Sons.

Schaap, G., Lammers, S., & de Vogel, V. (2009). Risk assessment in female forensic psychiatric patients: A quasi-prospective study into the validity of the HCR-20 and PCL-R. *Journal of Forensic Psychiatry & Psychology, 20*(3), 354-365. doi: 10.1080/14789940802542873

Scott, C. L., & Resnick, P. J. (2006). Violence Risk Assessment in Persons with Mental Illness. *Aggression and Violent Behavior, 11*(6), 598-611. doi: 10.1016/j.avb.2005.12.003

Serin, R. C., Chadwick, N., & Lloyd, C. D. (2016). Dynamic risk and

protective factors. *Psychology, Crime, & Law, 22*(1-2), 151-170. doi: 10.1080/1068316X.2015.1112013

Serper, M. R., Goldberg, B. R., Herman, K. G., Richarme, D., Chou, J., Dill, C. A., & Cancro, R. (2005). Predictors of aggression on the psychiatric inpatient service. *Comprehensive Psychiatry, 46*(2), 121-127. doi: 10.1016/j.comppsych.2004.07.031

Serper, M., Beech, D. R., Harvey, P. D., & Dill, C. (2008). Neuropsychological and symptom predictors of aggression on the psychiatric inpatient service. *Journal of Clinical and Experimental Neuropsychology, 30*(6), 700-709. doi: 10.1080/13803390701684554

Shapiro, S.L., and Carlson, L.E. (2009). *The Art and Science of Mindfulness: Integrating Mindfulness Into Psychology and the Helping Professions.* Washington D.C, American Psychological Society.

Shuman, D. W., & Darwin-Looney, B. (2008). Legal issues of prediction, protection, and expertise *Textbook of violence assessment and management* (pp. 429-440). Arlington, VA: American Psychiatric Publishing, Inc ; US.

Sinani, L., Kola, V., Cenko, F., Elezi, F., Balaj, K., Saraci, J., Dervishi, P., & Gjolena, G. (2013). Assessment and treatment of patients with violent behaviors in emergency psychiatry. The HCR-20 violence risk assessment scheme. *European Psychiatry, 28*(1), 247. doi: 10.1016/S0924-9338(13)77013-5

Singh, J. P., & Petrila, J. (2013). Measuring and interpreting the predictive validity of violence risk assessments: an overview of the special

issue. *Behavioral Sciences & the Law, 31*(1), 1-7. doi: 10.1002/bsl.2054

Singh, J. P., Fazel, S., Gueorguieva, R., & Buchanan, A. (2014). Rates of violence in patients classified as high risk by structured risk assessment instruments. *British Journal of Psychiatry, 204*(3), 180-187. doi:10.1192/bjp.bp.113.131938

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513.

Singh, J. P., Grann, M., Lichtenstein, P., Langstrom, N., & Fazel, S. (2012). A novel approach to determining violence risk in schizophrenia: Developing a stepped strategy in 13,806 discharged patients. *PLoS ONE, 7*(2). doi:10.1371/journal.pone.0031727

Singh, J. P., Serper, M., Reinharth, J., & Fazel, S. (2011). Structured Assessment of Violence Risk in Schizophrenia and Other Psychiatric Disorders: A Systematic Review of the Validity, Reliability, and Item Content of 10 Available Instruments. *Schizophrenia Bulletin, 37*(5), 899-912. doi:10.1093/schbul/sbr093

Sjostrom, N., Eder, D. N., Malm, U., & Beskow, J. (2001). Violence and its prediction at a psychiatric hospital. *European Psychiatry, 16*(8), 459-465. doi: 10.1016/S0924-9338(01)00607-1

Skeem, J. L., & Mulvey, E. P. (2001). Psychopathy and community violence among civil psychiatric patients: Results from the MacArthur violence risk assessment study. *Journal of Consulting and Clinical Psychology, 69*(3), 358-374. doi: 10.1037/0022-

006X.69.3.358

Skeem, J. L., Manchak, S. M., Lidz, C. W., & Mulvey, E. P. (2013). The utility of patients' self-perceptions of violence risk: Consider asking the person who may know best. *Psychiatric Services, 64*(5), 410-415. doi:10.1176/appi.ps.001312012

Skeem, J. L., Mulvey, E. P., & Lidz, C. W. (2000). Building mental health professionals' decisional models into tests of predictive validity: The accuracy of contextualized predictions of violence. *Law and Human Behavior, 24*(6), 607-628. doi: 10.1023/A:1005513818748

Skeem, J., Schubert, C., Stowman, S., Beeson, S., Mulvey, E., Gardner, W., & Lidz, C. (2005). Gender and risk assessment accuracy: Underestimating women's violence potential. *Law and Human Behavior, 29*(2), 173-186. doi: 10.1007/s10979-005-3401-z

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior, 24*(3), 271-296. doi:10.1023/a:1005595519944

Smith, S. T., Edens, J. F., & McDermott, B. E. (2013). Fearless Dominance and Self-Centered Impulsivity Interact to Predict Predatory Aggression among Forensic Psychiatric Inpatients. *International Journal of Forensic Mental Health, 12*(1), 33-41. doi:10.1080/14999013.2012.760186

Smith, T. E., & Munich, R. L. (1992). Suicide, violence, and elopement: Prediction, understanding, and management. *American Psychiatric Press Review of Psychiatry, 11*, 535-554. Retrieved from

http://www.search.proquest.com

Snowden, R. J., Gray, N. S., Taylor, J., & Fitzgerald, S. (2009). Assessing risk of future violence among forensic psychiatric inpatients with the Classification of Violence Risk (COVR). *Psychiatric Services, 60*(11), 1522-1526. doi: 10.1176/ps.2009.60.11.1522

Snowden, R., Gray, N. S., & Taylor, J. (2010). Risk assessment for future violence in individuals from an ethnic minority group. *The International Journal of Forensic Mental Health, 9*(2), 118-123. doi: 10.1080/14999013.2010.501845

Soliman, A. E. D., & Reza, H. (2001). Risk factors and correlates of violence among acutely ill adult psychiatric inpatients. *Psychiatric Services, 52*(1), 75-80. doi:10.1176/appi.ps.52.1.75

Soyka, M., Graz, C., Bottlender, R., Dirschedl, P., & Schoech, H. (2007). Clinical correlates of later violence and criminal offences in schizophrenia. *Schizophrenia Research, 94*(1-3), 89-98. doi:10.1016/j.schres.2007.03.027

Spiessl, H., Krischker, S., & Cording, C. (1998). Aggressive behaviour in a psychiatric hospital: A 6-year study of 17943 admissions based upon the psychiatric basic documentation. *Psychiatrische Praxis, 25*(5), 227-230. Retrieved from http://www.europepmc.org

Stadtland, C., Hollweg, M., Kleindienst, N., Dietl, J., Reich, U., & Nedopil, N. (2005). Risk assessment and prediction of violent and sexual recidivism in sex offenders: Long-term predictive validity of four risk assessment instruments. *Journal of Forensic Psychiatry and Psychology, 16*(1), 92-108. doi: 10.1080/1478994042000270247

Stafford, E., & Cornell, D. G. (2003). Psychopathy scores predict

adolescent inpatient aggression. *Assessment, 10*(1), 102-112. doi:10.1177/1073191102250341

Starzomski, A., & Wilson, K. (2015). Development of a measure to predict short-term violence in psychiatric populations: The Imminent Risk Rating Scale. *Psychological Services, 12*(1), 1-8. doi: 10.1037/a0037281

Steadman, H. J. (1992). "Prediction of dangerous behavior: A review and analysis of 'second-generation' research" and "Expert testimony on violence and dangerousness: Roles for mental health professionals": Comment. *Forensic Reports, 5*(1), 155-158. Retrieved from http://www.researchgate.net

Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Clark Robbins, P., Mulvey, E. P., Grisso, T., Loren, H., & Banks, S. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior, 24*(1), 83-100. doi: 10.0147-7307/00/0200-0083

Steinert, T. (2002). Prediction of inpatient violence. *Acta Psychiatrica Scandinavica, Supplementum*, *4*(12), 133-141. doi: 10.1034/j.1600-0447.106.s412.29.x

Steinert, T. (2002). Violent behaviour of patients in institutions predictors and limitations. *Psychiatrische Praxis, 29*(2), 61-67. doi:10.1055/s-2002-22038

Steptoe, L. R., Lindsay, W. R., Murphy, L., & Young, S. J. (2008). Construct validity, reliability and predictive validity of the dynamic risk assessment and management system (DRAMS) in offenders with intellectual disability. *Legal and Criminological Psychology,*

13(2), 309-321. doi: 10.1348/135532507X218251

Stokman, C. L. (1984). Dangerousness and violence in hospitalized mentally ill offenders. *Psychiatric Quarterly, 56*(2), 138-143. doi: 10.1007/BF01064949

Strand, S., & Belfrage, H. (2001). Comparison of HCR-20 scores in violent mentally disordered men and women: Gender differences and similarities. *Psychology, Crime & Law, 7*(1), 71-79. doi: 10.1080/10683160108401784

Strub, D. S., Douglas, K. S., & Nicholls, T. L. (2014). The validity of Version 3 of the HCR-20 violence risk assessment scheme amongst offenders and civil psychiatric patients. *The International Journal of Forensic Mental Health, 13*(2), 148-159. doi: 10.1080/14999013.2014.911785

Strub, D. S., Douglas, K. S., & Nicholls, T. L. (2016). Violence risk assessment of civil psychiatric patients with the HCR-20: Does gender matter? *The International Journal of Forensic Mental Health, 15*(1), 81-96. doi: 10.1080/14999013.2016.1141438

Sturup, J., Kristiansson, M., & Lindqvist, P. (2011). Violent behaviour by general psychiatric patients in Sweden-validation of Classification of Violence Risk (COVR) software. *Psychiatry Research, 188*(1), 161-165. doi: 10.1016/j.psychres.2010.12.021

Sturup, J., Monahan, J., & Kristiansson, M. (2013). Violent behavior and gender of Swedish psychiatric patients: A prospective clinical study. *Psychiatric Services, 64*(7), 688-693. doi: 10.1176/appi.ps.201200266

Swanson, J. W. (2008). Preventing the unpredicted: managing violence

risk in mental health care. *Psychiatric Services, 59*(2), 191-193. doi:10.1176/appi.ps.59.2.191

Szmukler, G. (2001). The mathematics of risk assessment for serious violence. *Psychiatric Bulletin, 25*(9), 359. doi: 10.1192/pb.25.9.359

Szmukler, G. (2003). Risk assessment: 'Numbers' and 'values'. *Psychiatric Bulletin, 27*(6), 205-207. doi: 10.1192/pb/27.6.205

Szmukler, G. (2012). Risk assessment for suicide and violence is of extremely limited value in general psychiatric practice. *Australian and New Zealand Journal of Psychiatry, 46*(2), 173-174. Retrieved from http://www.journals.sagepub.com

Szmukler, G., & Rose, N. (2013). Risk assessment in mental health care: values and costs. *Behavioral Sciences & the Law, 31*(1), 125-140. doi: 10.1002/bsl.2046

Taylor, J. L., DuQueno, L., & Novaco, R. W. (2004). Piloting a ward anger rating scale for older adults with mental health problems. *Behavioural and Cognitive Psychotherapy, 32*(4), 467-479. doi:10.1017/s135246580400164x

Taylor, L., & Large, M. (2013). Dynamic appraisal of situational aggression lacks utility. *International Journal of Mental Health Nursing, 22*(6), 579. doi: 10.111/inm.12019

Telles, L. E. d. B., Folino, J. O., & Taborda, J. G. V. (2012). Accuracy of the Historical, Clinical and Risk Management Scales (HCR-20) in predicting violence and other offenses in forensic psychiatric patients in Brazil. *International Journal of Law and Psychiatry, 35*(5-6), 427-431. doi: 10. 1016/j.ijlp.2012.09.001

Telles, L. E. D., Day, V. P., Folino, J. O., & Taborda, J. G. V. (2009).

Reliability of the Brazilian version of HCR-20 Assessing Risk for Violence. *Revista Brasileira de Psiquiatria, 31*(3), 253-256. doi:10.1590/s1516-44462009005000001

Tengstrom, A. (2001). Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry, 55*(4), 243-249. doi:10.1080/080394801681019093

Tengstrom, A., Grann, M., Langstrom, N., & Kullgren, G. (2000). Psychopathy (PCL-R) as a predictor of violent recidivism among criminal offenders with schizophrenia. *Law and Human Behavior, 24*(1), 45-58. doi:10.1023/a:1005474719516

Tenneij, N. H., Didden, R., Stolker, J. J., & Koot, H. M. (2009). Markers for aggression in inpatient treatment facilities for adults with mild to borderline intellectual disability. *Research in Developmental Disabilities, 30*(6), 1248-1257. doi:10.1016/j.ridd.2009.04.006

Teo, A. R., Holley, S. R., Leary, M., & McNiel, D. E. (2012). The Relationship Between Level of Training and Accuracy of Violence Risk Assessment. *Psychiatric Services, 63*(11), 1089-1094. doi:10.1176/appi.ps.201200019

Thomas, S. D., Daffern, M., Martin, T., Ogloff, J. R., Thomson, L. D., & Ferguson, M. (2009). Factors associated with seclusion in a statewide forensic psychiatric service in Australia over a 2-year period. *International Journal of Mental Health Nursing, 18*(1), 2-9. doi:10.1111/j.1447-0349.2008.00567.x

Thomson, L. D. G., & Marshall, L. A. (2013). The predictive validity of the HCR-20 following clinical implementation: Does it work in practice?

Journal of Forensic Psychiatry and Psychology, 24(3), 371-385. doi: 10.1080/14789949.2013.800894

Thomson, L., Davidson, M., Brett, C., Steele, J., & Darjee, R. (2008). Risk assessment in forensic patients with schizophrenia: The predictive validity of actuarial scales and symptom severity for offending and violence over 8 - 10 years. *The International Journal of Forensic Mental Health, 7*(2), 173-189. doi: 10.1080/14999013.2008.9914413

Troisi, A., Kustermann, S., Di Genio, M., & Siracusano, A. (2003). Hostility during admission interview as a short-term predictor of aggression in acute psychiatric male inpatients. *Journal of Clinical Psychiatry, 64*(12), 1460-1464. doi: 10.4088/JCP.v64n1240

Troquete, N. A., van den Brink, R. H., Beintema, H., Mulder, T., van Os, T. W., Schoevers, R. A., & Wiersma, D. (2015). Predictive validity of the short-term assessment of risk and treatability for violent behavior in outpatient forensic psychiatric patients. *Psychological Assessment, 27*(2), 377-391. doi: 10.1037/a0038270

Vaaler, A. E., Iversen, V. C., Morken, G., Flovig, J. C., Palmstierna, T., & Linaker, O. M. (2011). Short-term prediction of threatening and violent behaviour in an Acute Psychiatric Intensive Care Unit based on patient and environment characteristics. *BMC Psychiatry, 11*. doi:10.1186/1471-244x-11-44

van de Sande, R., Noorthoorn, E., Wierdsma, A., Hellendoorn, E., van der Staak, C., Mulder, C. L., & Nijman, H. (2013). Association between short-term structured risk assessment outcomes and seclusion. *International Journal of Mental Health Nursing, 22*(6), 475-484.

doi:10.1111/inm.12033

Vance, J. E., Bowen, N. K., Fernandez, G., & Thompson, S. (2002). Risk and protective factors as predictors of outcome in adolescents with psychiatric disorder and aggression. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(1), 36-43. doi: 10.1097/00004583-200201000-00009

Vernham, Z., Tapp, J., & Moore, E. (2016). Observer ratings of interpersonal behavior as predictors of aggression and self-harm in a high-security sample of male forensic inpatients. *Journal of Interpersonal Violence, 31*(9), 1597-1617. Retrieved from http://www.journals.sagepub.com

Viljoen, S., Nicholls, T. L., Roesch, R., Gagnon, N., Douglas, K., & Brink, J. (2016). Exploring gender differences in the utility of strength-based risk assessment measures. *International Journal of Forensic Mental Health, 15*(2), 149-163. doi:10.1080/14999013.2016. 1170739

Villeneuve, D. B., & Quinsey, V. L. (1995). Predictors of general and violent recidivism among mentally disordered inmates. *Criminal Justice and Behavior, 22*(4), 397-410. doi:10.1177/ 0093854895022004004

Vitacco, M. J., Gonsalves, V., Tomony, J., Smith, B. E., & Lishner, D. A. (2012). Can standardized measures of risk predict inpatient violence? Combining static and dynamic variables to improve accuracy. *Criminal Justice and Behavior, 39*(5), 589-606. Retrieved from http://www.journals.sagepub.com

Vitacco, M. J., Van Rybroek, G. J., Rogstad, J. E., Yahr, L. E., Tomony, J.

D., & Saewert, E. (2009). Predicting short-term institutional aggression in forensic patients: A multi-trait method for understanding subtypes of aggression. *Law and Human Behavior, 33*(4), 308-319. doi: 10.1007/s10979-008-9155-7

Vojt, G., Marshall, L. A., & Thomson, L. D. G. (2010). The assessment of imminent inpatient aggression: A validation study of the DASA-IV in Scotland. *Journal of Forensic Psychiatry and Psychology, 21*(5), 789-800. doi: 10.1080/14789949.2010.489952

Vojt, G., Thomson, L. D., & Marshall, L. A. (2013). The predictive validity of the HCR-20 following clinical implementation: Does it work in practice? *Journal of Forensic Psychiatry & Psychology, 24*(3), 371-385. doi: 10.1080/14789949.2013.800894

Waldheter, E. J., Jones, N. T., Johnson, E. R., & Penn, D. L. (2005). Utility of social cognition and insight in the prediction of inpatient violence among individuals with a severe mental illness. *Journal of Nervous and Mental Disease, 193*(9), 609-618. doi:10.1097/01.nmd.0000177788.253573.de

Ward, T., & Brown, M. (2004). The good lives model and conceptual issues in offender rehabilitation. *Psychology, Crime & Law*, *10*(3), 243-257. doi: 10.1080/10683160410001662744

Ward, T., Mann, R. E., & Gannon, T. A. (2007). The good lives model of offender rehabilitation: Clinical implications. *Aggression and Violent Behavior*, *12*(1), 87-107. doi: 10.1016/j.avb.2006.03.004

Waschgler, K., Ruiz-Hernandez, J. A., Llor-Esteban, B., & Garcia-Izquierdo, M. (2013). Patients' aggressive behaviours towards nurses:

development and psychometric properties of the hospital aggressive behaviour scale- users. *Journal of Advanced Nursing, 69*(6), 1418-1427. doi:10.1111/jan.12016

Watt, A., Topping-Morris, B., Rogers, P., Doyle, M., & Mason, T. (2003). Pre-admission nursing assessment in a Welsh medium secure unit (1991-2000): Part 2-comparison of traditional nursing assessment with the HCR-20 risk assessment tool. *International Journal of Nursing Studies, 40*(6), 657-662. doi: 10.1016/S0020-7489(03)00017-8

Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, D. L. (2009). Manual for the Short-Term Assessment of Risk and Treatability (START) (Version 1.1). Port Coquitlam, BC, Canada: British Columbia Mental Health and Addiction Services

Weizmann-Henelius, G., Virkkunen, M., Gammelgard, M., Eronen, M., & Putkonen, H. (2015). The PCL-R and violent recidivism in a prospective follow-up of a nationwide sample of female offenders. *Journal of Forensic Psychiatry & Psychology, 26*(5), 667-685. doi:10.1080/14789949.2015.1049192

Whittington, R., Hockenhull, J. C., McGuire, J., Leitner, M., Barr, W., Cherry, M. G., Fletnje, B., Quinn, Y., & Dickson, R. (2013). A systematic review of risk assessment strategies for populations at high risk of engaging in violent behaviour: update 2002-8. *Health Technology Assessment (Winchester, England), 17*(50), i-xiv, 1-128. doi:http://dx.doi.org/10.3310/hta17500

Wilson, C. M., Desmarais, S. L., Nicholls, T. L., & Brink, J. (2010).

The role of client strengths in assessments of violence risk using the Short-Term Assessment of Risk and Treatability (START). *International Journal of Forensic Mental Health, 9*(4), 282-293. doi: 10.1080/14999013.2010.534694

Wilson, C. M., Desmarais, S. L., Nicholls, T. L., Hart, S. D., & Brink, J. (2013). Predictive validity of dynamic factors: Assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior, 37*(6), 377-388. doi: 10.1037/lhb0000025

Witt, K., Lichtenstein, P., & Fazel, S. (2015). Improving risk assessment in schizophrenia: epidemiological investigation of criminal history factors. *The British Journal of Psychiatry, 206*(5), 424-430. doi: 10.1192/bjp.bp.114.144485

Woods, P., & Almvik, R. (2002). The Broset violence checklist (BVC). *Acta Psychiatrica Scandinavica, 106*, 103-105. doi:10.1034/j.1600-0447.106.s412.22.x

Woods, P., & Ashley, C. (2007). Violence and aggression: a literature review. *Journal of Psychiatric and Mental Health Nursing, 14*(7), 652-660. doi: 10.1111/j.1365-2850.2007.01149.x

Woods, P., Olver, M., & Muller, M. (2015). Implementing violence and incident reporting measures on a forensic mental health unit. *Psychology, Crime & Law, 21*(10), 973-986. doi: 10.1080/1068316X.2015.1077247

Yao, X., Li, Z., Arthur, D., Hu, L., & Cheng, G. (2012). The application of a violence risk assessment tool among Chinese psychiatric service users: A preliminary study. *Journal of Psychiatric and Mental Health Nursing, 19*(5), 438-445. doi: 10.1111/j.1365-2850.2011.01821.x

Yao, X., Li, Z., Arthur, D., Hu, L., & Cheng, G. (2014). Validation of the Violence Risk Screening-10 instrument among clients discharged from a psychiatric hospital in Beijing. *International Journal of Mental Health Nursing, 23*(1), 79-87. doi: 10.1111/j.1447-0349.2012.00890.x

Young, M. H., Justice, J., & Erdberg, P. (1999). Risk factors for violent behavior among incarcerated male psychiatric patients: A multimethod approach. *Assessment, 6*(3), 243-258. doi:10.1177/107319119900600305

Zhou, J., Witt, K., Xiang, Y., Zhu, X., Wang, X., & Fazel, S. (2016). Violence risk assessment in psychiatric patients in China: A systematic review. *Australian and New Zealand Journal of Psychiatry, 50*(1), 33-45. Retrieved from http://www.journals.sagepub.com

Appendices

Appendix A

PsycINFO

1. Mental health services/ (29785)
2. Hospitals/ or psychiatric hospitals/ (19375)
3. Hospitalized patients/ (11372)
4. exp psychiatric patients/ (27735)
5. Inpatient.mp (25243)
6. Violence/ or patient violence/ (26630)
7. Violen*.mp (76493)
8. Aggress*.mp (76612)
9. Assault.mp (9466)
10. Physical abuse/ or verbal abuse/ or violence/ (30609)
11. Sexual abuse/ (17950)
12. Hostility/ (4729)
13. Aggressive behavio*.mp (34766)
14. Anger/ (7777)
15. Threat/ (7891)
16. Risk.mp (283496)
17. Risk assess*.mp (15319)
18. Risk assessment/ (11217)
19. IORNS.mp (3)
20. Inventory of offender risks needs and strengths.mp (0)
21. SAVRY.mp (67)
22. Structured assessment of violence risk in youth.mp (136)
23. SAPROF.mp (13)
24. Structured assessment of protective factors.mp (23)
25. (Short-term assessment of risk and treatability).mp [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (73)
26. Protective factors/ or protective factors.mp (8594)
27. Prediction/ (17800)
28. Statistical validity/ (15066)
29. Predictive validity.mp or predictive validity/ (20366)
30. Test validity/ or test sensitivity/ or test specificity/ (62175)
31. Area under curve.mp or area under curve/ (193)
32. AUC.mp (1834)
33. Test reliability/ (44037)
34. Accuracy.mp (58996)
35. 1 or 2 or 3 or 4 or 5 (98959)
36. 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 (175912)
37. 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 (285807)
38. 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 (160802)
39. 35 and 36 and 37 and 38 (203)

MEDLINE

1. Mental health services/ (28197)
2. Hospitals/ (68497)
3. Hospitals, Psychiatric/ (23896)
4. Inpatients/ (15714)
5. Violence/ (26799)

6. Assault.mp (7417)
7. Physical abuse/ (69)
8. Physical abuse.mp (3647)
9. Sexual abuse.mp or sexual abuse/ (15065)
10. Hostility/ (4827)
11. Aggression/ (29247)
12. Aggress*.mp (156611)
13. Violen*.mp (50488)
14. Anger/ (6572)
15. Threat.mp (35748)
16. Risk assessment/ (204255)
17. Risk assess*.mp (223297)
18. Risk/ (107646)
19. IORNS.mp (3)
20. Inventory of offender risks needs and strengths.mp (0)
21. SAVRY.mp (21)
22. Structured assessment of violence risk in youth.mp (19)
23. SAPROF.mp (8)
24. Structured assessment of protective factors.mp (4)
25. (Short-term assessment of risk and treatability).mp [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] (13)
26. Protective factors/ (886)
27. Prediction.mp (43814)
28. Predictive validity.mp (4362)
29. Statistical validity.mp (243)
30. Test validity.mp (311)
31. "Sensitivity and Specificity"/ (307782)
32. Area under curve.mp or area under curve/ (33045)
33. AUC.mp (38675)
34. Accuracy.mp (237795)
35. Test reliability.mp (720)
36. "Reproducibility of results"/ (320573)
37. 1 or 2 or 3 or 4 (130604)
38. 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 (258268)
39. 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 (324932)
40. 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 (857049)
41. 37 and 38 and 39 and 40 (79)

Embase

1. Mental health service/ (47867)
2. Hospital/ (359462)
3. Mental hospital/ (24589)
4. Hospital patient/ (111564)
5. Inpatient.mp (84021)
6. Violence/ or assault/ or battering/ or homicide/ or physical violence/ or sexual violence/ or verbal hostility/ (56431)

7. Violen*.mp (69832)
8. Aggress*.mp (233404)
9. Aggressive behavio*.mp (15280)
10. Aggression/ or aggressiveness/ or anger/ or threat/ (66283)
11. Risk.mp (2700518)
12. Risk assess*.mp (410836)
13. Risk assessment/ (392747)
14. IORNS.mp (7)
15. Inventory of offender risks needs and strengths.mp (0)
16. SAVRY.mp (34)
17. Structured assessment of violence risk in youth.mp (33)
18. SAPROF.mp (11)
19. Structured assessment of protective factors.mp (6)
20. (Short-term assessment of risk and treatability).mp [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] (29)
21. Protective factors.mp (8107)
22. Prediction/ (261836)
23. Predictive validity/ or predictive validity.mp (10032)
24. Statistical validity.mp (328)
25. Validity/ or predictive validity/ (41954)
26. "Sensitivity and Specificity"/ (250190)
27. Area under curve.mp or area under curve/ (76845)
28. AUC.mp (67802)
29. Accuracy/ or accuracy.mp (589194)
30. Reliability/ or interrater reliability/ (114127)
31. 1 or 2 or 3 or 4 or 5 (548787)
32. 6 or 7 or 8 or 9 or 10 (320540)
33. 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 (2702476)
34. 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 (1159053)
35. 31 and 32 and 33 and 34 (362)

Web of Science

1. TS=("mental health service" OR hospital OR "psychiatric hospital" OR inpatient OR "hospitalized patient" OR "psychiatric patient") *Databases:* SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH. *Timespan:* all years. (679983)
2. TS=(violen* OR aggress* OR assault OR hostility OR threat OR "aggressive behavio*" OR "physical violence" OR "sexual violence" OR anger) *Databases:* SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH. *Timespan:* all years. (424790)
3. TS=(risk OR "risk assess*" OR "risk assessment" OR IORNS OR "Inventory of offender risks needs and strengths" OR SAVRY OR "structured assessment of violence risk in youth" OR SAPROF OR "structured assessment of protective factors" OR "short-term assessment of risk and treatability" OR "protective factors") *Databases:* SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH. *Timespan:* all years. (2043508)
4. TS=(prediction OR "predictive validity" OR "statistical validity" OR validity OR reliability OR sensitivity OR specificity

OR "area under curve" OR AUC OR accuracy) *Databases:* SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH. *Timespan:* all years. (3144248)

5. #1 AND #2 AND #3 AND #4 *Databases:* SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH. *Timespan:* all years. (732)

ASSIA

all ("mental health service" OR hospital OR "psychiatric hospital" OR inpatient OR "hospitalized patient" OR "psychiatric patient")
AND
all (violen* OR aggress* OR assault OR hostility OR threat OR "aggressive behavio*" OR "physical violence" OR "sexual violence" OR anger)
AND
all (risk OR "risk assess*" OR "risk assessment" OR IORNS OR "Inventory of offender risks needs and strengths" OR SAVRY OR "structured assessment of violence risk in youth" OR SAPROF OR "structured assessment of protective factors" OR "short-term assessment of risk and treatability" OR "protective factors")
AND
all (prediction OR "predictive validity" OR "statistical validity" OR validity OR reliability OR sensitivity OR specificity OR "area under curve" OR AUC OR accuracy) (85)

Cochrane Library

("mental health service" OR hospital OR "psychiatric hospital" OR inpatient OR "hospitalized patient" OR "psychiatric patient") in Title, Abstract or Keywords
AND
(violen* OR aggress* OR assault OR hostility OR threat OR "aggressive behavio*" OR "physical violence" OR "sexual violence" OR anger) in Title, Abstract or Keywords
AND
(risk OR "risk assess*" OR "risk assessment" OR IORNS OR "Inventory of offender risks needs and strengths" OR SAVRY OR "structured assessment of violence risk in youth" OR SAPROF OR "structured assessment of protective factors" OR "short-term assessment of risk and treatability" OR "protective factors") in Title, Abstract or Keywords
AND
(prediction OR "predictive validity" OR "statistical validity" OR validity OR reliability OR sensitivity OR specificity OR "area under curve" OR AUC OR accuracy) in Title, Abstract or Keywords
(27)

Appendix B

Inclusion Form

Source database:…….

Full Reference:
……
……

Inclusion criteria	**Yes Unclear Discuss No**	**Comments**
Study design: Is the study a cohort study, meta-analysis, case series or case report, or systematic review? *Population:* Does the population consist of inpatients? Is the population in a mental health setting? *Exposure:* Is a violence risk assessment tool used? Does this tool include protective factors? *Outcomes:* Have violent behaviours been measured?		
Not an opinion paper		
If all questions answered with 'yes', include study	**Include** **Do not include**	

Reason for exclusion:……

……

……

Appendix C

Quality Assessment Form

Source database:…….

Full Reference:
……
…….

Question	Score Y(2) P(1) N(0) U	Comments
Were the study objectives clear? *Selection bias* Was the sample recruited in an acceptable way? Was the sample representative of the population? *Measurement bias* Was the risk assessment tool described? Was the tool applied using enough information? Was inter-rater reliability measured? Was inter-rater reliability 0.8 or above? Were violent behaviours clearly recorded? Were violent behaviours gathered uniformly? Was missing information dealt with appropriately? *Attrition* Was non-completion rate recorded? Was non-completion stage detailed? Was non-completion reason recorded? *Results* Were the results recorded? Is validity discussed? (AUC values?) Are the results reliable? Do the results fit with other related evidence? Are the results generalizable? Were confounding variables discussed/taken into account?		
Quality Score Number Unclear		

Quality assessment form adapted from the Critical Appraisal Skills Programme (CASP, 2004).

Scoring:

If the condition is met = 2
If the condition is partially met = 1
If the condition is not met = 0

Appendix D

Data Extraction Tool

General information

Date of data extraction:……

Author:……

Article title:……
……

Type of risk assessment studied*:*……..……

Source (e.g. journal, conference) year/volume/pages/country of origin:…….....
……

Notes:……
……

Specific Information

Study characteristics

1. Study type:

Cohort study Meta-analysis Case series Case report

Systematic Review

2. Location of study:……

3. Characteristics of participants:
   Age:……..

   Gender: Male Female

   Other information:……

4. Number of participants:…….

5. Mental health setting:……..

Exposure

1. Violent risk assessment used:…….

Outcomes and outcome measures

1. Type of violent behaviours:……
   ……

2. Drop-out rates (plus proportion of those who did not agree to participate, if possible) and reasons for drop-out:……
   ……
   ……
3. Follow-up period:……...

Analysis

1. Stats used (AUC values):……
   Inter-rater reliability (%agreement or kappa):……

2. Notes:

Appendix E

Study	Reason for Exclusion
Anonymous	Study design
Abderhalden et al., (2006)	Exposure
Abderhalden et al., (2004)	Exposure
Allnut et al., (2013)	Exposure
Almvik, Woods, & Rasmussen (2007)	Exposure
Amore et al., (2008)	Exposure
Anderson et al., (2004)	Exposure
Anonymous (2013)	Exposure
Arbach-Lucioni, Andres-Pueyo, Pomarol-Clotet, & Gomar-Sones (2011)	Exposure
Ash & Alonso-Katzowitz (2011)	Commentary/opinion
Baksheev et al., (2015)	Exposure
Ball et al., (1994)	Exposure
Barry-Walsh & Daffern (2010)	Exposure
Barry-Walsh, Daffern, Duncan, & Ogloff (2009)	Exposure
Barzman et al., (2011)	Exposure
Beck, McNeil, & Binder (1996)	Commentary/opinion
Becker & Grilo (2007)	Exposure
Belfrage (1998)	Exposure
Belfrage, Fransson, & Strand (2000)	Exposure
Bjorkdahl, Olsson, & Palmstierna (2006)	Exposure
Bjorkly (1993)	Exposure

Bjorkly (1994)	Exposure
Bjorkly (1995)	Exposure
Bjorkly (1997)	Exposure
Bjorkly, Eidhammer, & Selmer (2014)	Exposure
Bjorkly et al., (2009)	Exposure
Bjorkly, Havik, & Loberg (1996)	Exposure
Bloom, Webster, Hucker, & De Freitas (2005)	Study design
Borum (1996)	Study design
Boxer (2007)	Exposure
Buchanan (1999)	Study design
Buchanan (2008)	Study design
Buchanan (2013)	Study design
Buchanan, Binder, Norko, & Swartz (2012)	Study design
Buchanan et al., (2013)	Exposure
Burke (2010)	Study design
Calow, Lewis, Showen, & Hall (2016)	Population
Candini et al., (2015)	Exposure
Carroll (2007)	Study design
Castelletti, Rivellni, & Stratico (2014)	Exposure
Chagigiorgis, et al., (2013)	Exposure
Chakhssi, de Ruiter, & Bernstein	Exposure
Chan & Chow (2014)	Exposure
Chen & Hwu (2014)	Exposure
Chen, Hwu, & Hu (2014)	Exposure
Chu, Daffern, & Ogloff (2013)	Exposure

Chu, Hoo, Daffern, & Tan (2012)	Exposure
Chu, Thomas, Daffern, & Ogloff (2013)	Exposure
Chu, Thomas, Ogloff, & Daffern (2011)	Exposure
Claix & Pham (2004)	Exposure
Clarke, Brown, & Griffith (2010)	Exposure
Coid et al., (2009)	Exposure
Convit et al., (1988)	Exposure
Cornaggia, Beghi, Pavone, & Barale (2011)	Exposure
Cote, Crocker, Nicholls, & Seto (2012)	Exposure
Crocker, Stargatt, & Denton (2010)	Exposure
Cronin-Davis (2000)	Exposure
Dack et al., (2013)	Exposure
Daffern (2006)	Study design
Daffern & Howells (2007)	Exposure
Daffern, Howells, Ogloff, & Lee (2005)	Exposure
Daffern, Ogloff, Ferguson, & Thomson (2005)	Exposure
Daniel, Weir, & Tiffin (2013)	Study design
De Borba Telles, Folino, & Raborda (2012)	Exposure
De Ruiter & De Vogel (2004)	Exposure
De Vogel & De Ruiter (2005)	Exposure
De Vogel & De Ruiter (2006)	Exposure
De Vogel et al., (2004)	Exposure
De Vogel, De Ruiter, Van Beek, & Mead (2004)	Exposure
De Vogel et al., (2016)	Exposure

De Vries Robbé, De Vogel, Douglas, & Nijman (2015)	Population
De Vries Robbé, De Vogel, & Spa (2011)	Population
Dernevik, Grann, & Johansson (2002)	Exposure
Desmarais, Nicholls, Read, & Brink (2010)	Outcome
Dietl et al., (2005)	Exposure
Dignam, Brock, Hein, & McEvoy (2010)	Exposure
Dolan & Doyle (2000)	Exposure
Dolan & Fullam (2007)	Exposure
Dolan, Fullam, Logan, & Davies (2008)	Exposure
Douglas, Nicholls, & Ogloff (2004)	Exposure
Douglas & Ogloff (2003a)	Exposure
Douglas & Ogloff (2003b)	Exposure
Douglas, Ogloff, Nicholls, & Grant (1999)	Exposure
Dowsett (2005)	Exposure
Doyle, Dolan, & McGovern (2002)	Exposure
Doyle & Logan (2012)	Study design
Doyle, Shaw, Carter, & Dolan (2010)	Exposure
Drieschner, Marrozos, & Regenboog (2013)	Exposure
Dumais, Larue, Michard, & Goulet (2012)	Exposure
Eaton, Ghannam, & Hunt (2000)	Exposure
Edens, Skeem, & Douglas (2006)	Exposure
Eriksen et al., (2016)	Exposure
Fabian (2008)	Commentary/opinion
Fehon Grilo, & Lipschitz (2005)	Exposure

Ferguson et al., (2005)	Exposure
Ferris et al., (1997)	Exposure
Fluttert et al., (2013)	Exposure
Fluttert et al., (2011)	Exposure
Fujii, Tokioka, Lichton, & Hishinuma (2005)	Exposure
Fuller & Cowan (1999)	Exposure
Gammelgård, Koivisto, Eronen, & Kaltiala-Heino (2008)	Outcome Population
Gammelgård et al., (2015)	
Gray, Taylor, & Snowden (2008)	Exposure
Green et al., (2016)	Exposure
Grevatt, Thomas-Peter, & Hughes (2004)	Exposure
Griffith, Daffern, & Godber (2013)	Exposure
Grossi et al., (2015)	Exposure
Gunenc, O'Shea, & Dickens (2015)	Exposure
Gunn (1996)	Study design
Guy, Packer, & Warnken (2012)	Exposure
Hamilton & Bullard (1990)	Study design
Hare (1999)	Exposure
Harris & Lurigio (2007)	Study design
Harris & Rice (1994)	Study design
Harris & Rice (1997)	Exposure
Hart, Webster, & Menzies (1993)	Commentary/opinion
Hartvig et al., (2011)	Exposure
Hawley et al., (2006)	Exposure

Heilbrun et al., (1998)	Exposure
Hildebrand, De Ruiter, & Nijman (2004)	Exposure
Hilton, Harris, & Rice (2010)	Exposure
Holloway (2004)	Commentary/opinion
Hoptman et al., (1999)	Exposure
Hornsveld et al., (2014)	Exposure
Hornsveld et al., (2004)	Exposure
Howe et al., (2016)	Exposure
Huchzermeier et al., (2008)	Exposure
Ideker, Todicheeney-Mannes, & Kim (2011)	Exposure
Iozzino et al., (2015)	Exposure
Ivgi et al., (2015)	Exposure
Jung, Ledi, & Daniels (2013)	Exposure
Kasinathan & Marsland (2014)	Study design
Kasinathan et al., (2015)	Exposure
Kay, Wolkenfeld, & Murrill (1988)	Exposure
Ketelsen, Sechert, Driessen, & Schulz (2007)	Exposure
Kim, Ideker, Todicheeney-Mannes (2012)	Population
Klassen & O'Connor (1988)	Exposure
Kling et al., (2006)	Exposure
Kobes, Nijman, & Bulten (2012)	Exposure
Krakowski & Czobar (2004)	Exposure
Kruger & Rosema (2010)	Exposure
Lamberg (2007)	Study design

Langan (2010)	Study design
Langton et al., (2007)	Exposure
Langton et al., (2009)	Exposure
Langton et al., (2011)	Exposure
Lantta et a., (2016)	Exposure
Large (2010)	Commentary/opinion
Laurell, Belfrage, & Hellstron (2010)	Exposure
Lewis & Webster (2004)	Commentary/opinion
Lincoln et al., (2006)	Exposure
Lindsay et al., (2008)	Exposure
Littlechild & Hawley (2010)	Commentary/opinion
Lussier et al., (2010)	Exposure
Macpherson & Kevan (2004)	Exposure
Mao, Hamed, Gho, & Wen (2012)	Exposure
Margari et al., (2005)	Exposure
Martinaki et al., (2013)	Exposure
Martini et al., (2012)	Exposure
McDermott et al., (2008a)	Exposure
McDermott & Holoyda (2014)	Study design
McDermott et al., (2008b)	Exposure
McKelvey (2013)	Study design
McKenzie & Curr (2005)	Exposure
McMilan, Hastings, & Coldwell (2004)	Exposure
Mc Niel & Binder (1991)	Exposure
McNiel & Binder (1994)	Exposure
McNiel & Binder (1995)	Exposure

McNiel & Binder (1996)	Commentary/opinion
McNiel et al., (2003)	Exposure
McNiel, Lam, & Binder (2000)	Exposure
Meaden, Hacker, & Spencer (2013)	Exposure
Menzies & Webster (1995)	Exposure
Miller (2000)	Study design
Monahan (1997)	Study design
Monahan et al., (2000)	Exposure
Morrissey et al., (2007a)	Exposure
Morrissey et al., (2007b)	Exposure
Motamedi (2011)	Exposure
Motamedi (2012)	Study design
Mudde, Nijman, Van der Hulst, & Van den Bout (2011)	Exposure
Nagi et al., (2009)	Exposure
Nanayakkara, O'Driscoll, & Allnutt (2012)	Exposure
Neal, Miler, Shealy (2015)	Exposure
Neufeld, Perlman, & Hirdes (2012)	Exposure
Nicholls (2004)	Exposure
Nicholls, Ogloff, & Douglas (2004)	Exposure
Nijman et al., (2002)	Exposure
O'Shea et al., (2015)	Exposure
O'Shea & Dickens (2015a)	Unobtainable
O'Shea & Dickens (2015b)	Outcome
O'Shea & Dickens (2015c)	Exposure
O'Shea & Dickens (2016)	Outcome

O'Shea, Mitchell, Picchioni, & Dickens (2013)	Exposure
O'shea et al., (2014a)	Exposure
O'Shea et al., (2014b)	Exposure
O'Shea et al., (2015)	Exposure
Ogloff & Daffern (2006)	Exposure
Owen, Tarantello, Jones, & Tennant (1998)	Exposure
Palmstierna & Olsson (2007)	Exposure
Patel (2012)	Study design
Penney, McMaster, & Wilkie (2014)	Exposure
Penterman & Nijman (2011)	Population
Penterman & Nijman (2009)	Population
Penterman et al., (2013)	Population
Peters (1998)	Exposure
Pham, Remy, Dailliet, & Lienard (1998)	Exposure
Phillips, Stargatt, & Brown (2012)	Exposure
Phillips, Stargatt, & Fisher (2011)	Exopsure
Plutchik, Van Praag, Conte, & Picard (1989)	Outcome
Pomp, Spreen, Bogaerts, & Volker (2010)	Exposure
Rabinowitz & Garelik-Wyler (1999)	Exposure
Rechenmacher, Muller, Abdergalden, & Schulc (2014)	Exposure
Reimann (2009)	Exposure
Reimann & Nussbaum (2011)	Exposure

Reynolds et al., (2013)	Exposure
Rizzo & Smith (2012)	Study design
Roaldset (2015)	Commentary/opinion
Roaldset & Bjorkly (2010)	Exposure
Roaldset, Bjorkly, & Gotestam (2009)	Exposure
Roaldset, Hartvig, & Bjorkly (2011)	Exposure
Roaldset, Hartvig, Linaker, & Bjorkly (2012)	Exposure
Robbé, De Vogel, Douglas, & Nijman (2015)	Population
Schapp, Lammers, & De Vogel (2009)	Exposure
Scott & Resnick (2006)	Study design
Serper, Beeck, Harvey, & Dill (2008)	Exposure
Serper et al., (2005)	Exposure
Shuman & Darwin-Looney (2008)	Study design
Sinani et al., (2013)	Exposure
Singh, Fazel, Gueurguieva, Buchanan (2014)	Exposure Population
Singh, Serper, Reinharth, & Fazel (2011)	
Singh et al., (2012)	Exposure
Singh & Petrila (2013)	Commentary/opinion
Sjostrom, Eder, Malm, & Beskow (2001)	Exposure
Skeem et al., (2005)	Exposure
Skeem et al., (2013)	Exposure
Skeem & Mulvey (2001)	Exposure
Skeem, Mulvey, & Lidz (2000)	Exposure

Slovic, Monahan, & MacGregor (2000)	Exposure
Smith, Edens, & McDermott (2013)	Exposure
Smith & Munich (1992)	Commentary/opinion
Snowden, Gray, & Taylor (2010)	Exposure
Snowden, Gray, Taylor, & Fitzgerald (2009)	Exposure
Soliman & Reza (2001)	Exposure
Soyka et al., (2007)	Exposure
Spiessl, Krischker, & Cording (1998)	Commentary/opinion
Stadtland et al., (2005)	Exposure
Stafford & Cornell (2003)	Exposure
Starzomski & Wilson (2015)	Exposure
Steadman (1992)	Commentary/opinion
Steadman et al., (2000)	Exposure
Steinert (2002a)	Exposure
Steinert (2000b)	Exposure
Steptoe, Lindsay, Murphy, & Young (2008)	Exposure
Stokman (1984)	Commentary/opinion
Strand & Belfrage (2001)	Exposure
Strub, Douglas, & Nicholls (2014)	Exposure
Strub, Douglas, & Nicholls (2016)	Exposure
Sturup, Kristiansson, & Lindqvist (2011)	Exposure
Sturup, Monahan, & Kristiansson (2013)	Exposure
Swanson (2008)	Commentary/opinion
Szmukler (2001)	Commentary/opinion
Szmukler (2003)	Commentary/opinion

Szmukler (2012)	Commentary/opinion
Szmukler & Rose (2013)	Commentary/opinion
Taylor, DuQueno, & Novaco (2004)	Exposure
Taylor & Large (2013)	Exposure
Telles, Day, Folino, & Taborda (2009)	Exposure
Telles, Folino, & Taborda (2012)	Exposure
Tengstrom (2001)	Exposure
Tengstromm Grann, Langstrom, & Kullgren (2000)	Exposure
Tenneij, Didden, Stolker, & Koot (2009)	Exposure
Teo, Holley, Leary, & McNiel (2012)	Exposure
Thomas et al., (2009)	Exposure
Thomson et al., (2008)	Exposure
Thomson & Marshall (2013)	Exposure
Troisi, Kustermann, Di Genio, & Siracusano (2003)	Exposure
Troquette et al., (2015)	Population
Vaaler et al., (2011)	Exposure
Van De Sande et al., (2013)	Exposure
Vance, Bowen, & Fernandez (2002)	Outcome
Vernham, Tapp, & Moore (2016)	Exposure
Vileneuve & Quinsey (1995)	Exposure
Vitacco et al., (2012)	Exposure
Vitacco et al., (2009)	Exposure
Vojt, Marshall, & Thomson (2010)	Exposure
Vojt, Thomson & Marshall (2013)	Exposure

Waldheter, Jones, Johnson & Penn (2005)	Exposure
Waschgler, Ruiz-Hernandez, Llor-Esteban, & Garcia-Izquierdo (2013)	Study design
Watt et al., (2003)	Exposure
Weizmann-Henelius et al., (2015)	Exposure
Whittington et al., (2013)	Exposure
Witt, Lichtenstein, & Fazel (2015)	Exposure
Woods & Almvik (2002)	Exposure
Woods & Ashley (2007)	Exposure
Woods, Olver, & Muller (2015)	Exposure
Yao, Arthur, & Cheng (2012)	Exposure
Yao, Arthur, Hu, & Cheng (2014)	Exposure
Young, Justice, & Erdberg (1999)	Exposure
Zhou et al., (2016)	Outcome

Appendix F

The University of Nottingham

Faculty of Medicine and Health Sciences

Research Ethics Committee
School of Medicine Education Centre
B Floor, Medical School
Queen's Medical Centre Campus
Nottingham University Hospitals
Nottingham
NG7 2UH

Direct line/e-mail
+44 (0) 115 8232561
Louise.Sabir@nottingham.ac.uk

24th September 2015

Krista Jansen
Doctorate in Forensic Psychology
c/o Dr Vince Egan
Associate Professor of Forensic Psychology Practice
Centre for Family and Forensic Psychology
Psychiatry and Applied Psychology
School of Medicine
Room B23, YANG Fujia Building
Jubilee Campus
Nottingham
NG8 1BB

Dear Ms Jansen

Ethics Reference No: K17092015 SoM DFP PAP – please always quote
Study Title: Institutional Aggression in a Secure Forensic Mental Health Setting: Predictive Validity of the START.
Chief Researcher/Academic Supervisor: Dr Vince Egan, Associate Professor of Forensic Psychology Practice, Centre for Family and Forensic Psychology Psychiatry and Applied Psychology, School of Medicine.
Lead researchers/student: Ms Krista Jansen, Doctorate in Forensic Psychology Student,
Duration of Study: 01/10/2015-29/02/2016 3mths
No of Subjects: n/a Data Set Study: 107 female data sets using the Partnership in Care databases across 5 sites.

Thank you for submitting the above application which has been considered by the Committee at its sub-committee meeting on the 17th September 2015 and the following documents were received:

Predictive Validity of the START:
- FMHS Research Ethics Committee application Form dated 09/09/2015
- Research Proposal Final Version 1.0: 23/07/2015

These have been reviewed and are satisfactory and the study is approved.

Approval is given on the understanding that the Conditions of Approval set out below are followed.

1. Please can you submit copies of letters of permission from the 5 Partnership in Care sites when these are available for our records.

2. You must follow the protocol agreed and inform the Committee of any changes using a notification of amendment form (please request a form).

3. You must notify the Chair of any serious or unexpected event.

The University of Nottingham

4. This study is approved for the period of active recruitment requested. The Committee also provides a further 5 year approval for any necessary work to be performed on the study which may arise in the process of publication and peer review.

5. An End of Project Progress Report is completed and returned when the study has finished (Please request a form).

Yours sincerely

pp

Professor Ravi Mahajan
Chair, Faculty of Medicine & Health Sciences Research Ethics Committee

Appendix G

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Invitation to Take Part

Good morning,

I am a Forensic Psychologist in Training currently based at Calverton Hill. As part of my doctorate I am required to complete a research project. I have decided to look at the predictive validity of the START assessment in relation to inpatient aggression across medium secure female services. I intend to use existing data that is routinely collected by PiC to look at START scores and number/type of incidents from the IRAS system. This means that patients will not need to be approached in order to collect the data, as it is all already on the system.

I am writing to gain written permission for the use of your units in my research. I have attached my full research proposal and a letter of approval from the University of Nottingham Ethics Committee. The committee require a letter from each unit granting permission for the use of the data. Please complete the attached template if you are happy for me to use data from your respective units.

If you would like any more information, please do not hesitate to contact me.

Many thanks,
Krista Jansen
Forensic Psychologist in Training

[Unit address]

27th October 2015

To whom it may concern,

I give my permission for **Krista Jansen** to use data from **[Unit Name]** in her research project entitled "Institutional Aggression in a Secure Forensic Mental Health Setting: Predictive Validity of the START".

Ethics reference no: K17092015 SoM DFP PAP

[Signature]

[Your name here]

[Title]



RESEARCH PROPOSAL

(Final Version 1.0: 23/07/15)

Full Title of Study: Institutional Aggression in a Secure Forensic Mental Health Setting: Predictive Validity of the START
Short Title of Study: Predictive Validity of the START

Name of Researcher: Krista Jansen

Researcher's e-mail address: lwxkaj@nottingham.ac.uk

Supervisors' names: Vince Egan
Paul Mooney

Supervisors' e-mail addresses: Vince.Egan@nottingham.ac.uk

Paul.Mooney@partnershipsincare.co.uk

Introduction

Being able to accurately predict inpatient aggression is at the forefront of mental health services. With inpatient aggression producing a high emotional and financial cost for staff, patients, and institutions (Duffy, Doyle, & Ryan, 2004), a valid and reliable method of predicting aggression is highly sought after.

Current research highlights the difficulties in accurately predicting institutional aggression. Individuals who are aggressive within an institution are a heterogeneous group, differing in both symptoms and type and severity of previous aggression (Gadon, Johnstone, & Cooke, 2006; McDermott et al., 2008). This makes prediction more difficult for clinicians as the "typical" aggressive inpatient is not easy to identify.

Current Predictive Methods

HCR-20 This violence risk assessment scale has been the focus of much research attention in recent years. Despite its popularity among mental health services, recent findings show mixed results for the HCR-20 and short-term prediction of inpatient violence. For example, over a three-month period following initial admission, Gray et al., (2003) found that the HCR-20 was predictive of both verbal and physical aggression, as well as violence to property. However, Grevatt, Thomas-Peter, and Hughes (2004) found that HCR-20 scores did not predict physical aggression over a

6-month period. Conflicting evidence such as these suggests that the predictability of the HCR-20, at least in the short-term, is still under question.

Differences in the predictive validity of the HCR-20 between genders has also been noted. Nicholls, Ogloff, and Douglas (2004) found that the HCR-20 was not predictive of violence in men, but was predictive in women. These findings suggest that there is a gender difference in predicting inpatient aggression, and thus it cannot be assumed that what works in predicting institutional aggression in men will be equally effective for women.

PCL-R As a primary measure of psychopathy, the PCL-R is regularly utilised as a form of violence risk assessment. Several studies have looked at the effectiveness of the PCL-R and violent recidivism. Findings generalise across many offender-types, and have shown that the PCL-R predicts violent recidivism (Corrado et al., 2004; Hemphill, Hare, & Wong, S., 1998). However, these studies tend to focus on a prison population, and few studies have looked at the predictive value of the PCL-R in forensic psychiatric inpatients.

When looking at an inpatient population, the research shows that the PCL-R is only predictive of verbal aggression, and shows little or no correlation with acts of physical institutional aggression (Buffington-Vollum, Edens, Johnson and Johnson, 2002; Hildebrand, De Ruiter and Nijman, 2004).

Similarly to the HCR-20, the PCL-R has mainly focused on predicting violence in men. Salekin, Rogers, Ustad, and Sewell (1998) looked at psychopathy and recidivism among female inmates. They found distinct differences between male and female offenders, concluding that psychopathy appears to be less pronounced and subsequently less predictive of violent recidivism in women than in men. This again shows the importance of validating measures of risk assessment on women and not just men in order to accurately predict their risk of future violence or aggression.

START The aforementioned measures mixed findings in relations to their ability to predict inpatient aggression. This may be due to that fact that these assessments do not fully acknowledge the role of protective factors that may reduce the likelihood of an individual engaging in institutional aggression. Protective factors are an important consideration when looking at violent risk assessment. Combining a risk assessment scale such as the HCR-20 and a measure of protective factors such as the Structured Assessment of Protective Factors (SAPROF) has shown to outperform the predictive validity of each these assessments individually (de Vries Robbé, de Vogel, & de Spa, 2011).

The Short-Term Assessment of Risk and Treatability (START) adopts this method, by rating each of the ten facets as both a vulnerability and a strength (Webster et al., 2009). In this way, the START allows us to look at both the risk factors associated with aggressive behaviour within an institution, and also the strengths that could reduce the likelihood of aggressive behaviour.

Desmarais, Nicholls, Wilson, and Brink (2012) looked at the predictive validity of the START for inpatient aggression. This study

used a male sample, and found that the START was more predictive that HCR-20 subscale scores and PCL-R total scores. This study demonstrated that the use of both risk and protective scores allowed for a more accurate prediction of inpatient aggression (Desmarais et al., 2012). However, with only a male sample, these results cannot be generalised to a female population.

Rationale, Aims & Hypotheses

Previous research in the area of predicting institutional aggression has primarily looked at male populations. The current study will aim to investigate the predictive validity of the START over a three-month period with female forensic patients.

The research questions for this study are:

1) Is the START predictive of number of incidents over a three-month period?
    a. Which items are the most predictive?
2) Is the START predictive of type of incidents over a three-month period?
    a. Which items are the most predictive?

Method

Study design

This is a quantitative regression study, in which START ratings will be compared to number and type of incidents', as well as impulsivity scores, over a three-month period.

In total, a minimum of 107 participants are required for this study (effect size=.15, alpha=.05, power=.95).

Participants

Participants will be females currently detained in five medium-secure facilities within Partnerships in Care (see Table 3).

Inclusion Criteria

- Female inpatients currently residing in medium secure services within Partnerships in Care
- Over the age of 18 years old
- Have completed the START

Exclusion Criteria

- Females under the age of 18 years old
- Male inpatients
- Females in low secure or locked rehab
- Females who have not completed the START

Materials

The START is a 20-item risk assessment based on structured clinical judgement (Webster et al., 2009). Each item is scored on two scales: vulnerabilities and strengths. These items are rated on a three-point scale (0: *Minimally Present,* 1: *Moderately Present,* and 2: *Maximally Present*). Total scores from the vulnerability and strength scales will be obtained by summing the scores for individual scale items. Additionally, risk rating on seven domains (violence risk to others, self-harm, suicide, unauthorised leave,

substance abuse, self-neglect, and victimisation by others) could be rated as *Low, Moderate,* or *High*. However this risk is not derived from a total score, but rather a consideration of the 20-items. For the purpose of this study, only the actuarial ratings of vulnerabilities and strengths will be considered.

Procedure

Historical data from Partnerships in Care will be accessed to look at START data, impulsivity data, and incident data. Information on unit type and patient diagnoses will also be gathered. Incidents are recorded on an electronic system and are categorised into five different events (see Table 1). Incidents are also rated in terms of severity, ranging from Level 1 – No Harm, to Level 5 – Severe (see Table 2).

Information will be gathered for each patient on each item of the START, rated as both a vulnerability and a strength. Incident data will be retrieved from a three-month period following the date of the completion of the START.

Regression analyses will be used to identify a relationship between number and type of incidents and START information. ROC curve analyses will also be used to look at the predictive validity of the START.

Limitations

Types of unit may be different in that despite all units being medium secure, some may be high dependency, acute, or long-term care. This difference may produce a difference in incident reports and thus will be noted in discussion of the findings.

Additionally, although these units are all part of Partnerships in Care, the individual management style in each unit will be different. This may again create a difference in the classifications of incidents as a result of that units unique experiences and patients.

Finally, the patients involved in the research will have different diagnoses which will be noteworthy. Some units accommodate personality disorder, mental illness, or learning disability. This will be noted for each patient and considered in discussion of the findings.

Ethical considerations

The data used in this study is routinely collected by Partnerships in Care and will be historical data. In order to protect the participants' identity and to ensure all disclosures are confidential, each participant will be given a unique code to protect their anonymity. This proposal has been created in concordance with the Partnerships in Care Research Policy.

References

Buffington-Vollum, J., Edens, J. F., Johnson, D. W., & Johnson, J. K. (2004). Psychopathy as a predictor of institutional misbehaviour among sex offenders: A prospective replication. *Criminal Justice and Behavior, 29(5),* 497-511. doi: 10.1177/009385402236730

Corrado, R. R., Vincent, G. M., Hart, S. D., & Cohen, I. M. (2004). Predictive validity of the psychopathy checklist: Youth version for general and violent recidivism. *Behavioral Sciences and the Law, 22(1)*, 5-22. doi: 10.1002/bsl.574

Demarais, S. L., Nicholls, T. L., Wilson, C. M., & Bring, J. (2012). Using dynamic risk and protective factors to predict inpatient aggression: Reliability and validity of START assessments. *Psychological Assessment, 24(3),* 685–700. doi:10.1037/a0026668

Duffy, D., Doyle, M., & Ryan, T. (2004). Risk assessment and management in acute mental health care. In M. Harrison, D. Howard & D. Mitchell (Eds.) *Acute mental health nursing: From acute concerns to the capable practitioner* (pp. 144-161). Thousand Oaks, CA: Sage Publications, Inc.

Gadon, L., Johnstone, L., & Cooke, D. (2006). Situational variables and institutional violence: A systematic review of the literature. *Clinical Psychology Review, 26(5),* 515-534. doi: 10.1016/j.cpr.2006.02.002

Gray, N. S., Hill, C., McGleish, A., Timmons, D., MacCulloch, M. J., & Snowden, R. J. (2003). Prediction of violence and self-harm in mentally disordered offenders: A prospective study of the efficacy of HCR-20, PCL-R, and psychiatric symptomatology. *Journal of Consulting and Clinical Psychology, 71*(3), 185-218. doi: 10.1037/0022-006X.71.3.443

Grevatt, M., Thomas-Peter, B., & Hughes, G. (2004). Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient aggression. *Journal of Forensic Psychiatry & Psychology, 15*(2), 278-292. doi: 10.1080/1478994032000199095

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3(1),* 139-170. doi: 10.1111/j.2044-8333.1998.tb00355.x

Hildebrand, M., De Ruiter, C., & Nijman, H. (2004). PCL-R psychopathy predicts disruptive behaviour among male offenders in a Dutch forensic psychiatric hospital. *Journal of Interpersonal Violence, 19(1),* 13-29. doi: 10.1177/0886260503259047

McDermott, B. E., Edens, J. F., Quanbeck, C. D., Busse, D., & Scott, C. L. (2008). Examining the role of static and dynamic risk factors in the prediction of institutional violence: Variable and person-focused analyses. *Law and Human Behavior, 32(4)*, 325-338. doi: 10.1007/s10979-007-9094-8

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing risk for violence among male and female civil psychiatric patients:

The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law, 22(1)*, 127-158. doi: 10.1002/bsl.579

Salekin, R. T., Rogers, R., Ustad, K. L., & Sewell, K. W. (1998). Psychopathy and recidivism among female inmates. *Law and Human Behavior, 22(1),* 109-128. doi: 10.1023/A:1025780806538

de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients: A retrospective validation study of the SAPROF. *International Journal of Forensic Mental Health, 10(3),* 178-186. doi: 10.1080/14999013.2011.600232

Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, D. L. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START)* (Version 1.1). Port Coquitlam, BC, Canada: British Columbia Mental Heath and Addiction Services

Appendix H

Consent Form for Assessments

- I have been given information about the risk, well-being and "Coping Skills" group assessments and I have been given the chance to ask questions.



- I know that I can ask more questions or ask for more information at any time.



- I can stop the assessment at any time (but I can choose to start again later on).



- I know that if I tell Krista Jansen about any risks to myself or others, or information on my offence, this has to be shared with people who need to know this information. This could include my team or the police.



- I understand that my therapist is a Forensic Psychologist in Training, so my progress in individual sessions will be talked about with her/his Supervisor.



- I understand that notes may be taken in each session and computer notes are also kept for each session. Only authorised members of staff can see this information on a need to know basis. This information may be used to write psychology reports for the Patient Care Team and CPA meetings.



- I agree for my information to be used in a Case Study report which will be looked at by the University of Nottingham. I understand that the report will not include any information which could identify me.



Name: [redacted] Name: Krista Jansen

Signed: [redacted] Signed:
(Forensic Psychologist in Training)

Date: 29/6/15 Date: 29th June 2015

Appendix I

Partnerships in Care

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Risk Incidents in Females in Medium Secure Forensic Mental Health Services: Predictive Validity of the Short-Term Assessment of Risk and Treatability (START)

Krista Jansen – University of Nottingham, The Centre of Family and Forensic Psychology
Paul Mooney – Partnerships in Care; University of Nottingham, The Centre of Family and Forensic Psychology
Dr Vincent Egan – University of Nottingham, The Centre of Family and Forensic Psychology

Background

- Risk assessment has traditionally focused on vulnerabilities, and neglected an individual's strengths or protective factors
- The START looks at seven separate risk domains including risk of violence, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, and victimisation
- A unique feature of the START is its consideration of 20 dynamic risk items as both a vulnerability and a strength, which subsequently allows practitioners to produce a specific risk estimate for each of the seven risk domains
- As the majority of the existing literature on the START is in relation to males, this study looked at females in medium secure forensic mental health services

Aims

This study aimed to extend the current knowledge of the START by:

i. examining the predictive validity of the START strength scores, vulnerability scores, and Specific Risk Estimates (SREs) for risk incidents over a three-month period,
ii. examining the relationship between START strength scores and vulnerability scores for number of risk incidents over a three-month period,
iii. examining if there is a significant difference between SRE rating and number of risk incidents, and
iv. identifying whether strength scores or vulnerability scores have incremental validity over one another.

Method

- 101 participants were gathered from four medium secure female forensic units across the United Kingdom.
- Current residents of these units who had completed an initial START assessment and had been a resident of the unit for at least three months were included in the study.
- Incident data was gathered for the three months following initial START assessment.
- Incident data was coded in line with the seven risk domains of the START as described in the START manual (Webster et al., 2009).

Data Analyses

- Predictive validity of strength scores, vulnerability scores and SREs: Area Under Curve
- Relationship between strength and vulnerability scores and number of risk incidents: Linear Regression
- Significance of SRE rating and number of risk incidents: One-way ANOVA
- Incremental validity of strength and vulnerability scores: Block-entry Linear Regression

Results

- Predictive validity:

	Strength Score			Vulnerability Score			SRE		
Outcome	AUC	p	95% CI	AUC	p	95% CI	AUC	p	95% CI
Violence	.561	.307	[.446, .675]	.551	.385	[.436, .667]	.562	.299	[.445, .678]
Self-harm	.580	.302	[.445, .675]	.552	.388	[.439, .666]	.714	<.001	[.612, .815]
Suicide	.659	.194	[.491, .827]	.468	.796	[.281, .656]	.654	.206	[.423, .886]
Unauthorised leave	.366	.272	[.108, .624]	.527	.824	[.287, .767]	.567	.585	[.290, .844]
Substance abuse	.437	.711	[.221, .653]	.640	.045	[.767, .913]	.556	.741	[.258, .894]
Self-neglect	.625	.098	[.480, .769]	.641	.062	[.488, .793]	.595	.208	[.443, .747]
Victimisation	.574	.261	[.446, .702]	.626	.056	[.499, .754]	.554	.416	[.424, .683]

Note. START Strength scores have been inverted for this analysis such that higher scores represent less strength; SRE, Specific Risk Estimate; AUC, Area Under the Curve; CI, Confidence Interval

- Relationship between strength and vulnerability scores and number of risk incidents:
  - **No significant change** in magnitude of risk outcomes based on strength or vulnerability scores
- Significance of SRE rating and number of risk incidents:
  - Significant between-group effects (high, medium, and low) found for **self-harm**, **suicide**, and **unauthorised leave** ($p<.05$)
- Incremental validity of strength and vulnerability scores:
  - To be assessed

Discussion

- Preliminary findings question the predictive validity of START Strength and Vulnerability scores for detained female mentally disordered offenders.
- Total Strength and Vulnerability scores do not give an indication of the number of risk incidents likely to occur.
- Specific Risk Estimates show significant variance between groups, illustrating the clinical utility of rating individuals as high, medium, or low on the seven risk estimates.

Key References

- de Ruiter, C., & Nicholls, T. L. (2011). Protective factors in forensic mental health: A new frontier. *International Journal of Forensic Mental Health, 10*(3), 160-170. doi: 10.1080/14999013.2011.600602
- Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior, 24*(5), 595-605. doi: 10.1023/A:1005575113507
- Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, D. L. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START)* (Version 1.1). Port Coquitlam, BC, Canada: British Columbia Mental Heath and Addiction Services

www.nottingham.ac.uk