PREDICTING LENGTH OF STAY IN A MALE MEDIUM SECURE PSYCHIATRIC HOSPTIAL

by

VICTORIA LOUISE WILKES

A thesis submitted for the Continuing Professional Doctorate in Forensic Psychology Practice (ForenPsyD)

Submitted to the Centre for Forensic and Criminological Psychology
School of Psychology
University of Birmingham
Edgbaston
Birmingham
B15 2TT
May 2012

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

This thesis examines factors associated with length of hospital stay for mentally disordered offenders, detained within the medium secure psychiatric estate. Following an introduction, Chapter two presents a systematic review examining the current literature on factors that predict length of stay for patients detained in medium secure hospitals. Mixed results were found. There was limited convergence across clinical and forensic variables investigated, but greater consensus on what is not associated with length of stay. The limited research available and inconsistencies found indicates the need for further research. Chapter three comprises an empirical research study, investigating which variables within a population of male mentally disordered offenders predict length of stay within a regional, medium secure psychiatric hospital. Preliminary analyses revealed statistically significant relationships between length of stay and nine variables. Effect sizes were small to medium. Logistic regression revealed a statistically significant relationship between length of stay of two years or more and having a diagnosis of schizophrenic disorder. Chapter four presents a critical review of the Historical, Clinical, Risk–20 Violence Risk Assessment (HCR-20) (Version 2), a widely adopted risk assessment framework utilised within forensic psychiatry and the standardised measure used within the empirical study. The review explores the literature on the reliability and validity of the HCR-20, and considers its strengths and limitations. A discussion of the work presented concludes the thesis.

DEDICATION

This thesis is dedicated to my parents, Roderick and Marie Wilkes. Thank you for all your guidance, support and love, and your endless confidence in me. I am so grateful to you both, for everything.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr Jessica Woodhams, my academic supervisor, who has encouraged me throughout this process; for providing me with much positive reinforcement, constructive criticism and support along the path to submitting this work.

I would also like to express my thanks to Sue Hanson, for all her support and encouragement as I have undertaken this project; for always being at the end of the telephone or to respond to emails at all hours – you really are a treasure.

And finally, I would like to acknowledge my husband, Richard. Thank you for all your love and support, for taking care of me, for the constant conveyor belt of cups of tea, and the numerous requests for late night proof reading. I could not have done this without you.

CONTENTS

Contents		**i**
List of Appendices		**iii**
List of Tables		**iv**
List of Figures		**v**
Chapter 1:	**Introduction**	**1**
Chapter 2:	**Predictors of Length of Stay in Male, Medium Secure (Forensic) Psychiatric Hospitals: A Systematic Literature Review**	**8**
	Abstract	**9**
	Introduction	**11**
	Method	**18**
	Results	**25**
	Discussion	**45**
	Conclusions	**55**
Chapter 3:	**Predicting Length of Stay in a Population of Male, Mentally Disordered Offenders Detained in a Medium Secure Psychiatric Hospital: An Empirical Study**	**58**
	Abstract	**59**
	Introduction	**61**
	Method	**70**

	Results	**72**
	Discussion	**91**
Chapter 4:	**The Historical, Clinical, Risk – 20 (HCR-20) Violence Risk Assessment Scheme (Version 2): Critique of a Psychometric Assessment**	**105**
	Introduction	**106**
	Background to the HCR-20	**108**
	Properties of Psychological Tests	**117**
	Limitations	**129**
	Conclusions	**134**
Chapter 5:	**Discussion**	**136**
References		**146**
Appendices		**156**

LIST OF APPENDICES

Appendix A.	Electronic Database Search Syntax	**157**
Appendix B.	References of Included Studies	**159**
Appendix C.	Quality Assessment Scoring Sheet: Cohort Studies	**160**
Appendix D.	Data Extraction Form	**162**

LIST OF TABLES

Table 1.	Inclusion/ Exclusion Criteria (PICO)	**21**
Table 2.	Characteristics of Study Participants	**36**
Table 3	Data Extraction of Included Studies	**37**
Table 4.	Characteristics of Included Studies	**41**
Table 5.	Strengths & Limitations of Included Studies and Quality Assurance Scores	**43**
Table 6.	Length of Stay in Years for Study Sample	**72**
Table 7.	Kolmogorow- Smirnov Test Results	**75**
Table 8.	*p* Values for Mann-Whitney U Tests for Length of Stay and Age on Admission	**78**
Table 9.	Spearman's Rho Correlation Coefficients for Length of Stay and the HCR-20 Items Scale Totals and Full Total	**83**
Table 10.	Logistic Regression Predicting Length of Stay	**90**

LIST OF FIGURES

Figure 1. Flowchart of Study Selection Process **22**

CHAPTER ONE.

INTRODUCTION

"Secure mental health services provide accommodation, treatment and support for people with severe mental health problems who pose a risk to the public."

(Centre for Mental Health, 2011, p.6)

Mentally disordered offenders are defined as "people who have a disability or disorder of the mind and have committed or are suspected of committing a criminal offence" (Centre for Mental Health, 2011, p.14). By the end of 2008, the number of mentally disordered offenders detained within hospital was 3937, the highest figure recorded in the previous ten years, of which 3460 were men (Ministry of Justice, 2010). The increasing trend for admission to secure services has seen a rapid expansion in secure service provision, which is met by both the National Health Service (NHS) and the independent health provider sector, at a cost to the NHS of £1.2 billon in England (2009/10), some 18.9% of the national expenditure for adult mental health services (Centre for Mental Health, 2010).

One area of significant growth has been that of medium secure services, a bi-product of the reduction in the availability of beds within the High Secure Hospital estate (Abbott, et al., 2005). The role of a medium secure mental health service can be said to be multi-faceted. Medium secure services were designed specifically for individuals who are considered to "pose a serious danger to the public" (Centre for Mental Health, 2011, p.12). At the individual level, they exist to provide treatment and support for mentally disordered offenders detained within a secure environment. Individuals who are detained in these services are typically regarded as posing too high a risk for them to receive treatment within general psychiatry services (Walker, Craissati, Batson, Amos & Knowles, 2012), but for whom detention in prison would

be detrimental to their mental health and access to regular mental health treatment difficult. Secure services have the joint responsibility of clinical and risk recovery, by offering a safe and secure environment in which mentally disordered offenders can recover from their illnesses, whilst addressing their associated risk issues. In this regard, secure services also serve an important function within our wider society, through the detention and treatment of a group of patients who, as with many of their prison-based counterparts, will at some point be discharged into the community. It is therefore incumbent on the forensic mental health services to address the serious risk issues that are a central feature of this population, so that they are able to lead law abiding, safe and successful lives within a fully integrated society when discharged.

In addition to this, society and the political system increasingly require that public services perform with maximum efficiency, whilst ensuring that patients are fully rehabilitated and safe to return to live in the community. This brings a tension to the work of forensic mental health services, which hold joint responsibility for care and treatment of patients within a health system, combined with detention and risk rehabilitation of a criminal justice based approach and the need to calm public fears about the risks posed by mentally disordered offenders. This tension arises because secure care is regarded as a high cost, low volume sector of the NHS (Walker, et al., 2012), catering to the needs of a small number of people. Sixty-five per cent of medium secure care is provided by the NHS at an average cost of £176,000 per patient, per year. The Centre for Mental Health (2011) have proposed that less than one third of those detained at the levels of high and medium security stay for less than two years, with nearly 50% of those detained, doing so for over 5 years, with the inevitable associated high costs. These length of stay figures from the Centre for

Mental Health (2011) contrast sharply with the guidance outlined within the Glancy Report (1974), which recommended a ceiling of eighteen months for those detained at the level of medium security, following which consideration for an alternative placement would be advised. Yet numerous studies have all reported trends within secure services where this recommended period of hospitalisation is regularly exceeded (Edwards, Steed & Murray, 2002; Kennedy, Wilson & Cope, 1997; Ricketts, Carnell, Davies, Kaul & Duggan, 2001; Shah, Waldron, Boast, Coid & Ulrich, 2011).

With the national financial downturn, there is concern that the continued expansion of secure services cannot be sustained at current levels (Centre for Mental Health, 2011). Pressure has already been applied by the requirements set by local NHS commissioning teams for specialist services to reduce their costs (West Midlands Strategic Commissioning Group, 2010), whilst at a national level, the NHS is seeking to make £15 -20 billion in efficiency savings by 2015. One method by which some of these savings can potentially be achieved is through the reduction in length of stay within hospital. Parsons (2006) has suggested that length of stay is now becoming a predominant means of cost control, and regularly regarded as a measure of performance of services. For the purpose of this thesis, the definition of length of stay as suggested by Nagtegaal, Horst and van der Schönberger (2011) has been used, and is described as "a discrete period of inpatient hospitalisation for patients receiving healthcare treatment" (p.218).

Length of stay within psychiatry services is by no means a new issue and has generated interest in the empirical literature for some time. Until more recently, research has primarily explored length of stay and patterns of admission within

general adult services, at least as far as the field of psychiatry is concerned (Creed, Tomenson, Anthony & Tramner, 1997; Faulkner, Tobin & Weir, 1994; Goldney, Fisher & Walmsley, 1998; Moran, Fragala, Wise & Noak, 1999). However, interest in length of stay within the forensic psychiatric sector is growing and necessary, given the current fiscal climate. To date, much of the research within the forensic field has followed that found within general psychiatry, with the focus being on describing patterns and trends of admission and patient characteristics (Shah, et al., 2011). With the pressure of needing to reduce length of stay, research appears to be turning towards increasing the understanding of the reasons that might underpin the duration of admissions to hospital, with the potential that this can influence the structure of treatment interventions and care pathways (Parsons, 2006; Walker, et al., 2012).

Despite awareness of the high costs associated with secure care provision and the requirements to reduce these, only one paper has, in recent years, specifically attempted to offer insights into which factors might be associated with length of stay within a British, mentally disordered population (Shah, et al., 2011). As a clinician working within the medium secure forensic psychiatry field, a challenge exists in providing high quality, therapeutic services to individuals who have the twin needs of mental health recovery and risk reduction, whilst being mindful of the increasing pressures associated with on-going efficiency drives. Therefore, further research is becoming essential to our understanding of the factors that are associated with length of stay for this population, to aid effective and meaningful care and treatment planning, and so that those detained within services remain no longer than is absolutely necessary.

Thesis Aims

This thesis aims to examine the issue of length of hospital stay for mentally disordered offenders, with a specific focus upon those who are detained within the medium secure psychiatric estate. It comprises a systematic literature review of the existing available literature, an empirical research study investigating factors that are predictive of length of stay for a sample of male mentally disordered offenders who have at one time been detained within a regional medium secure hospital, and a critique of the Historical, Clinical, Risk – 20: Assessing Risk for Violence (HCR- 20) (Version 2) (Webster, Douglas, Eaves & Hart, 1997), a framework for assessing long term potential for violent recidivism.

Chapters two and three examine directly the factors that are associated with length of stay. Within chapter two, a systematic literature review examines the current literature on understanding and explaining length of stay within the medium secure estate within forensic psychiatric services. The review considers the extent to which the current literature identifies any factors that have predictive ability for determining length of stay, and to what extent there is commonality across the studies that exist for mentally disordered offenders who are detained within medium secure services. The empirical research study presented in chapter three investigates this question within a population of male mentally disordered offenders, attempting to establish the nature of the relationship between socio-demographic, clinical and forensic variables and length of stay.

Chapter four presents a critical review of the HCR-20 (Version 2) (Webster, et. al., 1997). The HCR-20 is a widely adopted risk assessment framework utilised within

forensic mental health services (Douglas and Reeves, 2010). The HCR-20 is the standardised measure incorporated within the empirical study, as it provides information regarding clinical and forensic variables that were investigated in terms of their relationship to length of stay. The review explores the background to the development of the assessment framework. A critique of the assessment is offered through a review of the empirical evidence of the reliability and validity of the HCR-20, and considers its strengths and limitations.

The thesis concludes in Chapter five with a discussion of the work presented, drawing together the main findings and considering implications for future research and applied practice.

CHAPTER TWO.

PREDICTORS OF LENGTH OF STAY IN MEDIUM SECURE (FORENSIC) PSYCHIATRIC HOSPITALS.

A SYSTEMATIC LITERATURE REVIEW

ABSTRACT

Aim

The purpose of this paper is to present a review of the literature on length of stay for detained, mentally disordered offenders within the medium secure forensic psychiatric system and to detail what, if any, socio-demographic, clinical and forensic variables predict length of stay.

Methodology

Scoping methods were used to establish the necessity for this review. A literature review was then conducted using a systematic approach. Inclusion and exclusion criteria were applied, and data extracted and synthesised from the included studies. Of a total of 122 articles found, five were subject to quality assessment. All five articles were included in the review and were subjected to data extraction. A qualitative data synthesis approach was then completed and the results reported.

Results

Five studies were included within the review. All studies followed a retrospective cohort research design approach. The studies exposed disagreement regarding factors that are associated with length of stay, and very few findings were supported across all of the papers. There was limited convergence across clinical and forensic variables. There was greater convergence between the studies on the factors that were not found to be related to length of stay.

Conclusions

Firm conclusions about what factors predict the length of time a person might be detained within a medium secure psychiatric hospital cannot be confidently drawn from the review. The small number of studies reviewed presents a significant limitation to generalising from any results identified, however they do point to the need for further research into this subject. Limitations of the review are considered.

INTRODUCTION

This review, using a systematic approach, examines relevant literature pertaining to the factors that are relevant in understanding length of stay for mentally disordered offenders detained within the forensic psychiatric system, specifically at the level of medium security. Length of stay can be considered as reflecting "a discrete period of inpatient hospitalisation for patients receiving healthcare treatment" (Nagtegaal, Horst & van der Schönberger, 2011, p.218).

Background

In an era of restricted financial resourcing, policy makers and healthcare providers are now asked to consider, develop and employ methods by which treatment interventions can be delivered not only effectively but also efficiently (Centre for Mental Health, 2011; Department of Health, 2008) and those working within the healthcare arena are witnessing the introduction of frameworks and policies that specifically demand the balancing of high quality care and fiscal economy. Length of stay is an example of one of the significant issues that has faced healthcare service providers, policy makers and consumers in recent years. Research has proposed that longer inpatient hospitalisations are usually associated with greater cost of treatment (Compton, Craw & Rudisch, 2006). As a result, reducing the time spent in hospital by patients is becoming a key factor associated with controlling healthcare costs.

Changes to the financial structures of the health service in England, driven by political reform on the need to improve efficiency and effectiveness (Centre for Mental Health, 2011; Department of Health, 2008), have resulted in a shift from

historical contract-based funding to a new system whereby all healthcare providers, NHS or private, are paid according to the activity undertaken within their services. This system is known as Payment by Results and was a central element of the Labour Government's NHS Plan (Department of Health, 2000). It was first applied to the delivery of physical health care services and since 2003 Payment by Results has been implemented across key clinical areas, including accident and emergency. It established standardised prices and tariffs for key medical interventions, for example, surgery for hip replacements, and non-clinical activities such as cleaning and provision of catering, based upon the average cost of a specific procedure across all NHS hospitals. Therefore, the more activity a service provides, the more income it will generate, and vice versa. One of the postulated advantages of the Payment by Results scheme has been the ability of services and commissioners to reinvest any savings made when procedures are completed that cost below the national tariff. From a clinician's perspective, a concern about the Payment by Results scheme for mental health is its potential rigidity in the context of a person-centred care approach, and where inflexible time parameters for treatment may be dictated for specific patient groups by external forces. Similarly, concerns have been raised about the potential for the funding process to be manipulated, specifically as a way of reducing costs (Fairbairn, 2007).

Mental Health Services in England will adopt the Payment by Results scheme from 2012 (Department of Health, 2011). However, difficulties with this have been encountered due to some of the problems assigning a cost to mental health treatment interventions, not least the costs associated with hospitalisation that may add up to years, rather than days. This is because unlike physical health interventions, mental

health is a less precise science involving many invisible elements that aid a person's recovery and therefore length of stay can be difficult to predict. Despite this, Mental Health Services have not escaped the introduction of other frameworks that have perhaps been a preparatory step to the introduction of Payment by Results. The Commissioning for Quality and Innovation Payment Framework (CQUINs) (Department of Health, 2008) is an example of one such framework. Introduced by the Department of Health, CQUINs provides a structure for agreeing local quality improvement schemes. It links performance to payment (a local payment by results), as a means of rewarding excellence and achievement of quality improvement goals, whilst also supporting efficiency drivers. For example, within the West Midlands Strategic Commissioning Group Secure Services Strategy for 2010 to 2015, a key target is the reduction by 5% in the average length of stay (West Midlands Strategic Commissioning Group, 2010), which is described as equating to an annual cost saving of circa £5.5.million. This perhaps presents a context within which it is possible to see why length of stay has become such a controversial issue.

Research has previously attempted to identify which factors are associated with differences in length of stay, as a way of calculating and understanding clinical improvement and trends in clinical practice patterns (Trauer, Callaly & Hantz, 1999). However, in the last decade or so research has become focused on length of stay as an indicator, or outcome measure, of efficiency and quality of inpatient psychiatric care (Castro, Cockerton & Birke, 2002; Creed, et al., 1997; Faulkner, et al., 1994; Goldney, et al., 1998; Moran, et al., 1999). However, most of our understanding about what predicts length of stay is based upon findings from general psychiatry.

Even within the general psychiatry research base, there is a lack of consensus as to what factors consistently or reliably predict the length of time for which psychiatric patients will be detained. One of the reasons for this is the variety of assessment tools used to predict length of stay, and the wide range of psychiatric populations on whom the research has been based (Huntley, Cho, Christman & Csernensky, 1998). For example, several research papers describe the use of the Brief Psychiatric Rating Scale (BPRS, Overall & Gorham, 1962) (Anderson, Crist & Payne, 2004; Biancosino, Barbui & Grassi, 2005), the Global Assessment of Functioning (GAF, APA, 2000) (Compton, et al., 2006), the Health of the Nation Outcome Scales (HoNOS, Wing, Curtis & Beevor, 1996) (Goldney, et al, 1998; Trauer, et al., 1999), or the Millon Clinical Multi-Axial Inventory (MCMI, Millon, Millon, Davis & Grossman, 1994) (Piersma & Boes, 1997). Others rely on, or include demographic or clinically relevant variables (i.e. diagnosis, severity of illness (Creed, et al., 1997), use of seclusion/ restraints (Compton, et al., 2006), level of therapeutic engagements (Castro, et al., 2002), and history of previous employment (Moran, et al., 1999).

Where the research base is lacking is the application of this empirical investigation with a forensic population and what outcomes or predictors are associated with length of stay (Smith, White & MacCall, 2004). This is further complicated by a general lack of consensus about definitions and approaches to standardised outcomes (Fitzpatrick, et al., 2010). Fitzpatrick, et al. (2010) highlighted the diverse approach to the use of outcomes within forensic mental health, reflecting the dominance of public safety based outcomes, in addition to the wide range of clinical and rehabilitation outcomes measured. Cohen & Eastman (2000) considered the way in which appropriate outcomes should be measured for mentally disordered offenders. For example, they

postulate that demonstrating valid and reliable outcomes for individuals with psychiatric disorders is difficult because of the complexity of the disorders themselves, and also because of the wide range of interventions that are used to assist with symptom amelioration, such as pharmacological and psycho-social interventions. Where attempts have been made to look at length of stay within a forensic setting the general findings suggest that there is evidence for poorer long-term outcomes when associated with a shorter length of stay, and that successful treatment, as defined by a move to conditions of lower security including the community, is predominantly associated with longer periods of hospitalisation (Ricketts, et al., 2002).

Length of stay within forensic (secure) services potentially presents a challenge to the Payment by Results approach. This is because unlike generic psychiatric services, forensic services are remitted to manage individual recovery and rehabilitation whilst twinned with a public protection (policy) agenda (Department of Health, 2007b). These two philosophies can at times feel as though they are in competition with each other. Professionals are responsible for assisting service users in their recovery from mental illness and preparing them for community-based independence. However, final decisions regarding detention and discharge are for some service users, out of the jurisdiction of the treating clinicians, but lie instead with Government departments, such as the Ministry of Justice. For example, service users detained under a restricted hospital order (Section 37/41 of the Mental Health Act 1983) are subject to the Act itself but also to further restrictions by the Ministry of Justice whereby the Department, and ultimately the Minister for Justice, have overall responsibility for decisions regarding discharge from hospital. This is in direct contrast to those service users detained under Section 3 of the Mental Health Act 1983, where decisions

regarding discharge from hospital rest with a tribunal panel, based on the recommendations made and argued for by the treating clinical team. Thus, one of the difficulties faced by forensic services is in managing the competing interests of reducing length of stay and achieving subsequent financial savings, and the practicalities of managing and treating a complex population which is subject to a legal framework that can override recommendations regarding hospitalisation. Therefore as scrutiny increases over spending, as compared with the quality of service delivery, it is suggested that factors that help clinicians to identify variables which predict differences in length of stay, and indirectly successful clinical outcomes will be of continued interest, if not growing importance within forensic psychiatry. This suggests that further investigation into this subject is timely and warranted.

The Current Review

Scoping exercises and searches for existing systematic reviews and meta-analyses on this subject were conducted and resulted in poor returns. Whilst length of stay has been considered within the wider psychiatric and physical health systems, there is a lack of research pertaining to the timeframes for forensic psychiatric admission, and the variables associated with this. Therefore, a review of the literature specifically focussing on the medium secure psychiatric estate was deemed a valuable addition to this area.

Aim Of The Current Review

This review examined the literature pertaining to the length of stay for mentally disordered offenders detained within medium secure psychiatric hospitals. In particular this review had the following objective:

To determine what factors may be associated with or predictive of lengths of stay for male, mentally disordered offenders in conditions of medium security.

METHOD

Sources of Information

A scoping exercise was conducted to ascertain the existence and extent of any earlier reviews of the subject. Searches from the following organisations and libraries were conducted on the 6th January 2012: the Cochrane Database of Systematic Reviews (CDSR), DARE (the database for abstract reviews of effectiveness), the Centre for Evidence Based Medicine, Bandolier, The Joanna Briggs Institute and also through PubMed Clinical Queries. A search was also completed using the Google search engine. Two reviews that looked specifically at the issue of length of stay or hospitalisation within a psychiatric population were identified. However, neither of these revealed any investigation or results of reviews concerning a forensic psychiatric population. Rather their focus was upon general adult psychiatry, suggesting that a review addressing the issue of length of stay within the forensic psychiatric population was warranted and may be of value to the literature base.

Search Strategy

An electronic search was completed in one sitting on the 8th February 2012. Four electronic databases were searched: PsycINFO, EMBASE, CINAHL, and Web of Science. Date parameters of 1987 to the present were set (February Week 1 2012) with the exception of EMBASE (where the nearest date parameter was 1980 to 2012 Week 5). Each of these databases has provision for criterion-based limits when searching. Searches were not restricted by language or document type at the electronic search stage. Additional limits such as age of subjects to adult (18-64 years) were set where possible. Where limits for age could not be set electronically, this was applied

when articles were reviewed according to the PICO strategy. All searches were saved.

Search Terms

Keyword and search terms associated with mentally disordered offenders, inpatients and treatment duration were used during the searches (see below). Mapping to subject headings was employed where available to maximise inclusivity of available literature, as well as keyword searching, to account for variation in coding across databases. Wildcard options were applied in the databases, again to maximise article sourcing.

(offend*) OR (criminal*) OR (delinquent*) OR (convict*)

AND

(patient*) OR (mental* ill*) OR (mental* disorder*) OR (inpatient*) OR

(psychiatric patients) OR (mentally ill offenders)

AND

(forensic unit*) OR (forensic hospital*) OR (secure unit*) OR (psychiatric hospital*)

AND

(length of stay) OR (treatment duration) OR (length of treatment) OR

(length of admission) OR (psychiatric hospitalisation) OR (inpatient admission) OR

(psychiatric admission) OR (psychiatric detention)

A full list of search syntax can be found in Appendix A.

Study Selection

Selection of studies involved the searches conducted via electronic database resources

and the application of inclusion/exclusion criteria. Inclusion/ Exclusion criteria were formed on the basis of the research question and earlier reviews of the literature. The criteria can be found in Table 1.

The electronic database search completed on the 8th February 2012 generated a total of 122 citations across the four databases. After accounting for duplicates (*n*=22) and non-English Language articles (*n*=5), a total of 95 articles remained for review. All abstracts and titles were checked for relevance against identified inclusion/ exclusion criteria. Eighty-two articles were excluded at this point. The full text article was reviewed where these abstracts provided insufficient information. A further eight articles were excluded by this process, leaving five remaining articles for quality assessment. All of these articles were included following the quality assessment process (See Appendix B). Figure 1 provides an overview of the process of study selection.

Table 1. Inclusion / Exclusion Criteria (PICO)

	Inclusion	Exclusion
Population	Male Only Male and Female Adult (Age 18 to 64 years) Mentally disordered offenders Mental Health / Psychiatric	Female Only Child, Adolescent (Age < 18 or > 65) Non offenders Physical Health
Intervention	Inpatient Setting	Community Settings
Comparator	Forensic Medium Secure Care	General Adult Non – secure setting High Secure Low Secure
Outcomes	Length of Stay Treatment Duration	N/A
Study Type	Any	
Other		Non-English Papers Opinion Papers, Commentaries, Editorials, Unpublished Papers Dissertations Papers published prior to 1987

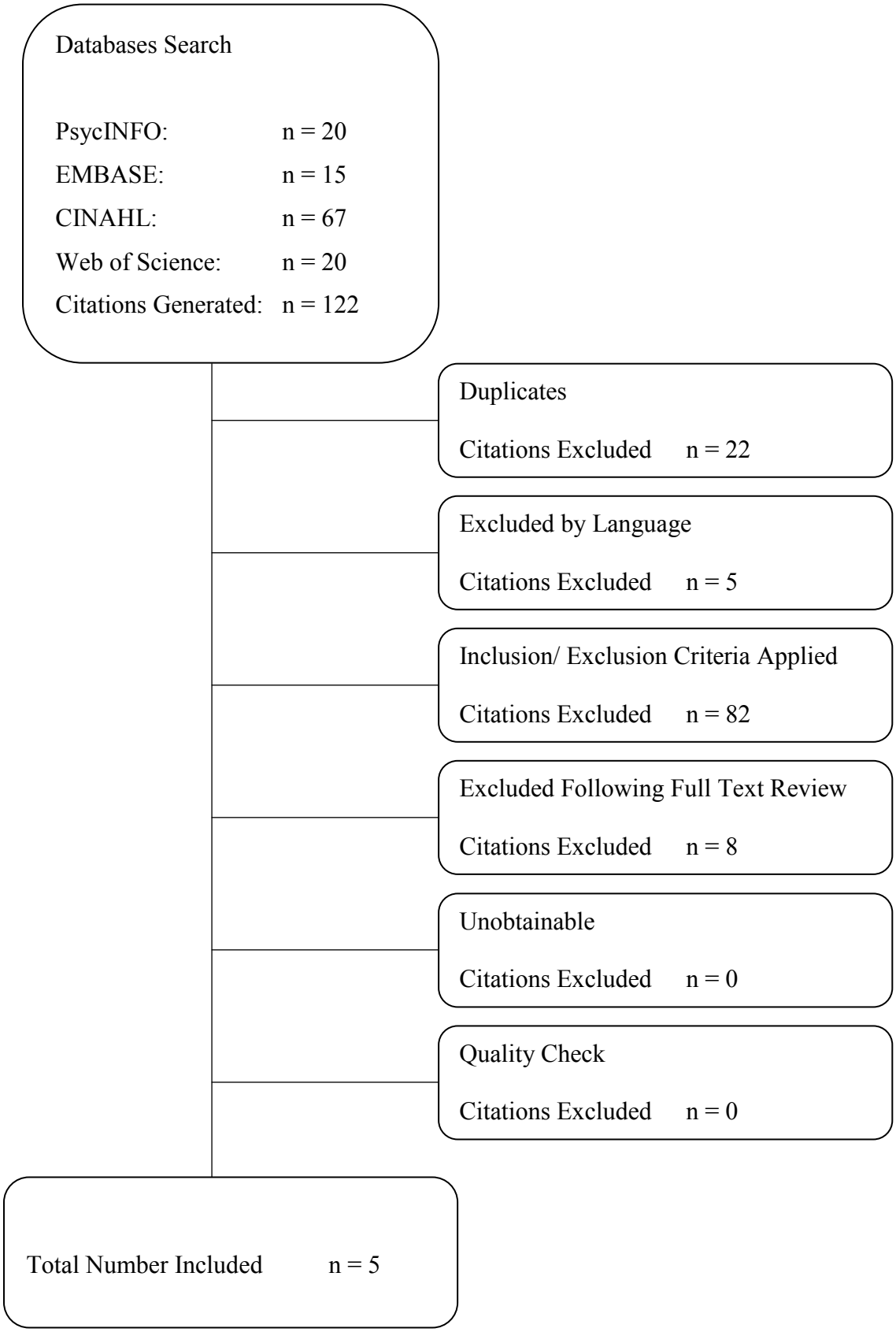

Figure 1. Flowchart of Study Selection Process

Quality Assessment

Subsequent to the inclusion / exclusion process, the methodological quality of included studies was assessed. The quality assessment criteria for the studies were modified from the Critical Appraisal Skills Programme (CASP, 2004) and an example of the scoring proforma can be found in Appendix C.

One researcher completed the quality assessment process. Areas of quality assessment included initial screening questions, appraisal of the study design and sample, bias, measurement of outcomes, appropriateness of statistical analysis used, results and the applicability of any findings. For each item appraised, a three-point scale was applied. A response of 'Yes' was accorded a score of 2, a score of 1 was given to responses of 'Partial', whilst a score of 0 was applied to responses of 'No'. 'Unknown' responses were not scored. This scoring system was adopted from the work of Kmet, Lee and Cook (2004), which set out a structured scoring system initially for systematic reviews. Essentially, the guidance suggests that a score of 2 for a 'Yes' would reflect that the factor being assessed is easily identifiable and appropriate to the subject matter. A score of 0 for a 'No' indicates the absence of the factor or is incomprehensible. The score of 1 for a 'Partial' response reflects a 'half-way house' approach – that the information presented is vague, incomplete or insufficiently defined.

An overall quality score was obtained by adding the scores for each paper. This would yield a maximum total score of 20. Individual scores were then converted to percentages, providing for an overall quality assessment score. Those studies that resulted in quality assessment score of 50% or above were included in the results. The

choice of quality assessment threshold may be considered generous (Kmet, et al., 2004), but was set reasonably low to reflect the limited number of studies available for critical appraisal. The reviewer also considered that the use of summary scores which define the quality of a study as high or low, has been criticised and is therefore to be used with some caution (Centre for Reviews and Dissemination, 2008). None of the articles reviewed at the quality assessment stage were excluded from the review.

Data Extraction

Data extraction was also completed by the sole researcher. Extraction of data from the articles was completed using a data extraction proforma (see Appendix D). The information extracted by the researcher included general information, such as the study title, authorship, date and publication details. Study eligibility was conferred through screening using the PICO structure defined in Table 1. Methodological content was considered, such as clarity of study design and recruitment, outcomes and variables measured, alongside the use of any standardised assessments and appropriateness of statistical analyses. The transparency of result and conclusions, with strengths and limitations was also recorded.

RESULTS

The process of study selection following the quality assessment process produced five studies for inclusion in the review (See Appendix B). Tables 2, 3, 4 and 5 (see pages 36 to 44) summarise the characteristics of each of the studies included, the participant characteristics and the details of the information yielded from the data extraction process, along with the strengths and limitations of the included studies and their quality assessment score.

Descriptive Data Synthesis

The results of the studies were examined from a qualitative standpoint, rather than combined statistically. This was predominately a result of time constraints. However, combining the results into a quantitative analysis may have potentially rendered any results meaningless, due to the difference across the studies in terms of variables and outcomes assessed, the differences within the populations, and the variety of the statistical analyses used across the five studies.

All of the studies reviewed shared the similarity of study design, in that they comprised retrospective cohort studies and were conducted in the United Kingdom. They sampled populations of individuals detained within psychiatric institutions, within conditions of medium security and incorporated male and female participants in their samples. A consistent pattern across all studies was the ratio of male to female participants, with females being a significantly smaller percentage of the cohort for each study. A total of 1,392 participants were recorded across the five studies, with 1,203 (86.4%) males and 189 (13.6%) females, accounted for. The age range reported

across four of the studies was between 15 and 74 years, with one study not defining the age range sampled (Kennedy, et al., 1995). The studies varied in their reporting of the mean ages within each sample, with some not reporting it at all (Castro, et al., 2002; Edwards, et al., 2002; Kennedy, et al., 1995) and the others reporting means of typically 30 years (Ricketts, et al., 2001; Shah, et al., 2011). Ethnicity was reported across all the studies reviewed, although recording patterns differed, with the studies by Edwards, et al. (2002); Kennedy, et al. (1995) and Ricketts et al. (2001) considering a maximum of four ethnicity categories of White, Asian, Afro-Caribbean and Other. The studies by Castro, et al. (2002) and Shah, et al. (2011) demonstrate a wider categorical division of ethnicity that is perhaps more in line with standards in equality and diversity monitoring seen within healthcare organisations.

Variability was seen across the samples in terms of descriptions of the selection of the cohorts used. For example, the study by Shah, et al. (2011) described its cohort as "patients discharged between 1999 and 2008*"* (pp.498). The remaining four studies considered patients that had been *admitted* between two fixed dates. It is only upon further scrutiny of the individual articles that the reader establishes what proportion of the sample were no longer being detained during the study period. This may reflect differences in psychiatric nomenclature. For example, as a clinician it is common practice within psychiatry to hear the time a person is detained in hospital referred to as an *admission*, accounting for not only the point of admittance, but also of discharge. However, for those unfamiliar with medical vernacular, this could be potentially confusing when interpreting results and generalising across studies. A potential discrepancy within each of the studies is the possibility of double counting. It is possible that within their sample groups, individuals may be counted more than once

if they have been admitted and/ or discharged on more than one occasion within the sampling timeframe. None of the studies appear to account for this in their descriptions of their sampling technique. Following on from this, there are also differences in the sampling timeframes used, which vary from four years (Castro, et al., 2002) to sixteen years (Edwards, et al., 2002). Given the lack of clarity of the sample descriptions, this presents further opportunity for double counting to occur because none of the studies make clear what proportion of the sample had more than one episode of care within the sample period.

The five studies each considered length of stay as an outcome. Three studies addressed the specific issue of discharge destination (Castro, et al., 2002; Edwards, et al., 2002; Ricketts, et al., 2001), without recourse to this as an indicator of successful treatment. The studies by Castro, et al. (2002) and Edwards, et al. (2002) also reviewed durability of outcome and forensic outcomes – referencing changes to discharge location and any reconvictions. Scrutiny of the articles revealed the research question and the outcomes that were being measured through the study. Some ambiguity was evident in respect of the research question in the articles by Edwards, et al. (2002) and Ricketts, et al. (2001). In the former, the study's aim is reported as being to "describe the outcome following admission" (p.69). In the latter paper, the authors detail that they are aiming to "identify trends in admission characteristics"(p.79). It is only by reading the articles in more detail that the reader becomes aware of the specific outcomes being measured. A commonality between these two papers is their reference to a former research study by Mohan, Murray, Taylor and Steed (1997) which reviewed trends and patterns in admission across a twelve-year period in a regional secure unit, and their desire to replicate and expand

on the findings in this earlier study. However, a limitation of this is that it relies on the reader having pre-requisite knowledge or a willingness to review the original paper.

The five studies shared commonalities regarding the variables used for statistical analysis. This is perhaps not surprising given that they all used a retrospective case study design, requiring historical records to be reviewed. Consistency was evident in choice of socio-demographic variables, such as gender, age and ethnicity as a minimum for data collection, to the reporting of marital and employment status and residence. Clinical variables also tended to be consistent, in terms of psychiatric diagnosis, history of previous psychiatric contact and admissions, legal status on admission and source of admission (for example, from prison or another hospital or the community). The paper by Castro, et al. (2002) stands alone as the only one which incorporates a variable regarding engagement in clinical therapies (psychological and occupational) as the paper seeks additionally to comment upon their positive relationship with successful treatment outcomes and draws on aspects of the clinical care pathway, such as engagement in psychological and occupational therapies to illustrate this.

Where some of the studies demonstrated variance was in the choice of forensic variables, which may be surprising given the populations sampled. For example, the study by Kennedy, et al. (1995) reported only on the nature of the index offence for its sample, whereas the other four studies included institutional offending, history of previous offending, age at first conviction, and gravity of offence types.

The article by Shah, et al. (2011) is the only one that provides an unambiguous

rationale for their selection of variables within their study and the use of a standardised measure, setting it apart from the other four articles. The measure referenced is the Historical, Clinical, Risk–20: Assessing Risk For Violence (HCR-20) (Version 2) (Webster, et al., 1997). They are explicit in their rationale that they have selected variables that have "previously been found to be associated with length of admission to medium security" (p.499) and that the HCR-20, central to assessment of risk in forensic services, offers an insight into factors that may be associated with "continuing detention" (p.500). Five of the fifteen variables selected within their study are found on the Historical scale of the HCR-20, including previous violence, relationship instability, substance use problems, major mental illness and personality disorder. Whilst there is clear commonality between the papers in their choice of variable, none of the remaining papers present this as clearly as the study by Shah, et al. (2011).

In reporting the key findings of the papers reviewed, it is important to consider the range of statistical analyses reported, as this might have some bearing on the significance of any results found and reflect limitations based upon sample size. Across the five papers we see convergence in the use of statistical tests, from Linear and Logistic Regression and Chi-Square, to ANOVA and *t*-tests, as well as the use of Fisher's Exact test. The choice of analyses reflects the nature of the variables being assessed and the need to accommodate both categorical and continuous data. Essentially they all examine the relationships between a series of variables.

Variance and contradictions are evident in the results reported across the five articles. One area where there is convergence across all five papers is that very few of the

socio-demographic variables, such as ethnicity, gender, age, marital status or employment status, appear to have any association with prolonged length of stay (i.e., length of stay ≥ 2 years). The studies highlight that whilst there may be some evidence to suggest differences in the main characteristics between those who stay in hospital longer than two years from those who are detained for less, these differences or associations between groups do not reach statistical significance. Only in the Edwards, et al. (2002) paper is there evidence of a statistically significant relationship between the socio-demographic variable of ethnicity, specifically being of White ethnic origin, and increased length of stay of more than five years ($R = 0.14$, $p < 0.05$), as identified by logistic regression analysis. Interestingly the study by Shah, et al. (2011) reported a significant relationship between shortened length of stay and being of Black origin, again using logistic regression analysis.

Reviewing the relationship between length of stay and clinical variables, such as having a history of previous psychiatric admissions and the nature of the psychiatric diagnosis, there appears to be more consistency between the articles reviewed. Shah, et al., (2011) reported positive associations between length of stay and the number of previous psychiatric admissions ($F = 5.02$, $p = .026$), as well as having a history of previous forensic psychiatric admissions ($F = 4.07$, $p = .045$). Along similar lines, Kennedy, et al., (1995) reported a statistically significant relationship between long histories of previous psychiatric contacts and prolonged length of stay, although they do not report their statistical outcomes in their published study. Castro, et al. (2002) also reported a statistically significant relationship between previous detentions and length of stay ($r = .152$, $p < .05$).

Shah, et al. (2011) present findings of positive associations between having a diagnosis of psychosis and length of stay ($F = 8.56$, $p = .004$). Inverse relationships between certain diagnostic variables and length of stay were also found within the studies reviewed. Again, Shah, et al. (2011), reported that a diagnosis of affective (mood) disorder was negatively associated with length of stay ($F = 3.91$, $p = .045$), using linear regression.

From a clinician's perspective, a valuable insight into the treatment pathway and its link to length of stay is reported by Castro, et al. (2002). Their study considered therapeutic engagement as a variable associated with length of stay. Their results show that engagement in psychological and occupational therapies was positively related to shortened length of stay for those admitted ($r = .168$, $p < .05$).

Statistically significant relationships between length of stay and substance misuse were also reported (Castro, et al., 2002; Kennedy et al., 1995). Only the latter paper reports its findings within the published article (Chi sq = 8.15, $p < 0.01$), indicating that a history of substance use was associated with increased length of stay. However, the paper by Shah, et al. (2011) reported no statistically significant relationship between substance use and length of stay, as measured through the use of the HCR-20 factor H5: Substance Use Problems.

Turning now to forensic variables, the paper by Shah, et al. (2011) presents an interesting finding that seriousness of index offence was inversely related to prolonged length of stay. In their sample, individuals with a severely violent index offence were increasingly likely to be discharged within a two-year period ($F = 1.26$,

p = .287), although this finding does not reach statistical significance. This study also reported that having a history of offending was not associated with length of admission. This is at odds with the findings reported by Edwards, et al. (2002) and Kennedy, et al. (1995), both of which report significant relationships between prolonged length of stay and severity of an index offence (homicide) (Edwards, et al., 2002; R = 0.22, $p < 0.001$; Kennedy, et al., 199; Chi Sq = 10.50, $p < 0.01$). Other forensic variables were considered in the Shah, et al. (2011) paper, through the application of the HCR-20 risk assessment framework. As noted earlier, it was the only study to have referenced a standardised measure. However, across the specific items that reflect forensic factors, such as having a history of previous violence (item H1) and young age at first conviction (item H2), they remark on the lack of an association between length of stay and issues related to risk, using categorical and linear statistical analyses.

The use of restricted hospital orders (Section 37/41, Mental Health Act 1983, as Amended 2007) were found to be significantly associated with increased length of stay following logistic regression analysis (OR = 3.62, p = .000) (Shah, et al., 2011). This finding replicated that found in the study by Kennedy, et al. (1995).

The source of admission has also been considered within the studies reviewed. Edwards, et al. (2002), Kennedy, et al. (1995) and Ricketts, et al. (2001) found that prolonged length of stay was associated with the source of admission being from a high secure hospital (i.e., Ashworth, Broadmoor or Rampton Hospitals). Kennedy, et al. (1995) reported a statistically significant relationship between increased length of stay and an admission from a High Secure Hospital (Chi Sq = 5.79, $p < 0.05$). Ricketts,

et al., (2001) reported the same finding when compared to those admitted from non-secure hospitals (F ratio = 4.600, p <0.001), despite also acknowledging that their admission rate from the high secure estate had dropped over the study timeframe. Edwards, et al., (2002) reported similar findings but did not stipulate whether this reached statistical significance.

The Quality Assurance Scores (QAS) for each study can be found in Table 5. The QAS achieved were all reasonably high, the highest being 97.5% (Shah, et al., 2011), followed by Edwards, et al., (2002) and Ricketts, et al., (2001) both with 95%, Castro, et al. (2002) with 85% and the lowest achieving 80% (Kennedy, et al., 1995).

The five studies incorporated into this review show similarities in terms of their comparability, within and to other similar studies, including those which have previously looked at length of stay within non-forensic populations, when considering their methodological approaches and their outcomes. One of their strengths was the representation of a mentally disordered population. The studies all considered data from combined male and female participants. It is suggested that this potentially raises questions regarding the generalisability of some of the findings, and the extent to which the findings might have differed if these two sample groups had been considered separately. None of the papers appear to adequately consider what gender related differences may have had on the variables associated with length of stay. This compares directly with some of the findings within the general psychiatry literature, where differences between length of stay for men and women have been found (Compton, et al., 2006).

Differences in quality were evident across the papers reviewed. These are summarised in Table 5. However, a limitation consistent across all the papers was the choice of retrospective case note reviews (albeit an appropriate choice of study design). One of the disadvantages of this approach is the reliance on notes and data that have been compiled historically by others, and the impact of missing data on any final dataset and conclusions drawn from analyses of this. It is also worthy of note that many psychiatric records are compiled using patient self-report. Such records (and therefore potential errors) progress through an admission with the patient, leaving the studies open to bias (Mann, 2003).

With the exception of the Shah, et al., (2011) study there were few examples where attempts to control for confounding factors had been applied or considered. For example, operational factors such as the increase or decrease in bed capacity or changes in service provision (Ricketts, et al., 2001), or more clinically relevant factors such as the extent to which readmissions of individuals within their sample were identified, were factors highlighted in the other studies, but did not appear to have been accounted for. In the study by Shah, et al. (2011) there is reference to the deliberate exclusion from the sample of cases of patients transferred between medium secure units, as this would have potentially biased the results, given their research question. There was also poor description of the sample characteristics in some cases (Kennedy, et al., 1995; Ricketts, et al., 2001), which makes it difficult to understand the choices for variable and outcome selection. As a result, whilst the studies reviewed might be comparable to other studies, it is difficult to comment with any certainty on their generalizability to the wider mentally disordered population. Greater transparency on the processes of establishing variables and outcomes may have

strengthened the quality of the studies and could in future allow for a meta-analytic approach to understanding factors associated with length of stay.

Table 2. Characteristics of Study Participants

Authors / Year	Sample Size	Gender	Age	Ethnicity
Castro, Cockerton & Birke (2002)	166	135 Male 31 Female	15 – 64 $\bar{x}$= Not stated	79 UK-European 47 Afro-Caribbean 23 African 6 European Other 5 Mixed 4 Unknown 2 Asian
Edwards, Steed & Murray (2002)	225	192 Male 33 Female	18 – 64 $\bar{x}$ = Not stated	124 White 79 Afro-Caribbean 13 Asian 9 Other
Kennedy, Wilson & Cope (1995)	100	91 Male 9 Female	Not stated	59 White 36 Afro Caribbean 5 Asian
Ricketts, Carnell, Davies, Kaul, & Duggan (2001)	504	413 Male 91 Female	18 - 64 $\bar{x}$ = 30	396 White 64 Black African / Caribbean 25 Mixed 19 Asian/ Other
Shah, Waldron, Boast, Coid, & Ullrich (2011)	259	234 Male 25 Female	18 – 74 $\bar{x}$ = 30.9	139 Black 71 White 24 South Asian 11 Mixed Race 14 Other

Table 3. Data Extraction of Included Studies

Author / Year	Sampling Timeframe	Sample Structure	Reported Length of Stay	Variables	Statistical Analyses	Attrition Rate/ Drop Outs
Castro, Cockerton & Birke (2002)	Patients admitted between 1995 and 1998	166 former patients	Mean = Not stated Median = Not stated Range = 0 to 27 months	Socio-demographics: age, gender, ethnicity, marital status, parenthood, employment status, educational background, residence. Clinical variables: legal status, diagnosis, admission source, previous hospitalisation, engagement in psychological and occupational therapy programmes Discharge location, durability of discharge, reconvictions, employment, compliance with medication.	Chi-Square, Pearson's Product Moment Correlation, ANOVA	Of 166, only 49 consented to follow-up (29%)

Edwards, Steed & Murray (2002)	Patients admitted between 1983 and 1996	225 Former patients	Mean = 26 months Median = Not stated Range = 9 days to 9 years	Socio-demographics: age, gender, ethnicity Forensic history: age at first conviction, seriousness of offending, index offence type and severity Clinical: age at first psychiatric contact, history of psychiatric contact, psychiatric diagnosis, admission source, legal status on admission, legal category of illness type	*t*-tests, Mann-Whitney Analysis, Logistic Regression	None noted
Kennedy, Wilson & Cope (1995)	Patients admitted between 1987 and 1993	Two groups: 31 'Long-stay' patients (LOS > 24 months) 69 Former patients (Control group LOS < 24 months)	Mean = 8.5 months Median = Not stated Range = Not stated	Socio-demographics: age, gender, ethnicity, marital status, accommodation, employment, source of admission, psychiatric diagnosis, legal status	*t*-tests, Chi-Square (using Yate's Continuity Correction)	None noted

Ricketts, Carnell, Davies, Kaul, & Duggan (2001)	Patients admitted between 1983 and 1999	Two groups: 493 former patients 11 inpatients	Mean = 313.2 days Median = Not stated Range = 2 to 3501 days	Socio-demographics: age, gender, ethnicity, marital status, Source of admission, psychiatric diagnosis, legal status, previous psychiatric history incl. admissions Forensic: Index offence nature and gravity	Chi-Square, ANOVA	4% attrition. 504 of 522 potential case notes sourced

Shah, Waldron, Boast, Coid, & Ullrich (2011)	Patients discharged between 1999 and 2008	259 former patients: of which 26 subject to forensic community follow-up	Mean = 748.9 days Median = 428 days Range = Not stated	Socio-demographics: age, gender, ethnicity Forensic history: history of offending, age at first conviction, seriousness of offending, index offence type and severity Clinical: age at first psychiatric contact, history of psychiatric contact, psychiatric diagnosis forensic/ non forensic, admission source, legal status on admission, legal category of illness type Other: Relationship instability, level of personal support	Logistic Regression, ANOVA, Chi-Square, Linear Regression	Of 26 subject to community follow up only 6 consented to participation (included in 259 sample)

Table 4. Characteristics of Included Studies

Authors / Year	Study Title	Study Type	Research Question	Outcome Indicators	Results / Main Findings
Castro, Cockerton & Birke (2002)	A small-scale study of medium secure provision in the independent sector.	Retrospective Cohort Study	To examine the effects of socio-demographic, behavioural and treatment variables on discharge and independent living at six month follow-up.	Length of Stay Discharge destination Durability of Outcome	Aggression in hospital, absconding & use of substances associated with increased length of stay. Socio-demographic variables not associated with length of stay. Engagement in psychological & occupational therapies significantly positively associated with length of stay. Most patients moved on to conditions of lower security = successful treatment outcome.
Edwards, Steed & Murray (2002)	Clinical and forensic outcome 2 years and 5 years after admission to a medium secure unit.	Retrospective Cohort Study	To describe outcomes following admission to a medium secure unit	Clinical outcomes: length of stay; discharge location; location after 2 and 5 years; Forensic outcomes after 2 and 5 years	Patients remaining after 5 years significantly different in their characteristics than the rest of sample group: typically White in ethnicity, convicted of murder. No other factors significantly associated with length of stay.

Kennedy, Wilson & Cope (1995)	Long-stay patients in a regional secure unit.	Retrospective Cohort Study	To demonstrate the existence of a new 'long stay' population within a medium secure hospital	Length of Stay	Substance misuse, longer history of previous psychiatric care and restriction orders highly correlated with increased length of stay. Association between increased length of stay and seriousness of index offence, discharge difficulties and treatment resistance.
Ricketts, Carnell, Davies, Kaul, & Duggan (2001)	First admissions to a regional secure unit over a 16-year period: Changes in demographic and service characteristics.	Retrospective Cohort Study	To identify trends in admission characteristics.	Length of Stay Discharge Location	Admissions from high secure hospital associated with increased length of stay. Successful treatment outcome (i.e. discharge to community or low secure setting) associated with lengths of stay greater than 2 years.
Shah, Waldron, Boast, Coid, & Ullrich (2011)	Factors associated with length of admission at a medium secure forensic psychiatric unit.	Retrospective Cohort Study	To identify characteristics associated with length of admission To identify characteristics of a group of patients who stay longer than 2 years (with respect to clinical, historical and demographic variables)	Length of Stay	Increased length of stay associated with hospital and restriction orders, previous psychiatric admissions and diagnosis of psychosis. Prior criminal history, seriousness of index offence, diagnosis of affective disorder or personality disorder and substance misuse not associated with prolonged length of stay.

Table 5. Strengths and Limitations of Included Studies and Quality Assurance Scores

Author / Year	Study Strengths	Study Limitations	Quality Assessment Score
Castro, Cockerton & Birke (2002)	Representative sample Comparable to other similar studies Not solely reliant on retrospective case note reviews	Timing of follow-up data collection not consistent Small sample size Sampling timeframe of 3 years Attempts to control for any confounding variables not explicitly stated Attrition rate No standardised measures utilised Self-report techniques used for follow – up.	85%
Edwards, Steed & Murray (2002)	Representative sample Study limitations referenced Study participants blind to research Comparable to other similar studies Sampling timeframe of 13 years	Research question not transparent Participant characteristics not clearly described Reliance on retrospective case note research Choice of follow up period not explained No standardised measures utilised	95%
Kennedy, Wilson & Cope (1995)	Comparable to other studies Study participants blind to research	Small sample size Study limitations not transparent Reliance on retrospective case note research Control group selection rational not transparent Sample characteristics insufficiently described No standardised measures utilised	80%
Ricketts, Carnell, Davies, Kaul, & Duggan (2001)	Large sample size Sampling timeframe of 11 years Comparable to other studies Study participants blind to research	Reliance on retrospective case note research Sample characteristics insufficiently described Attempts to control for any confounding variables not explicitly stated	95%

Shah, Waldron, Boast, Coid, & Ullrich (2011)	Moderate sample size Attempts to control for confounding variables applied Use of standardised instrument (HCR20) Comparable to previous studies Strengths and limitations transparent Sampling timeframe of 10 years	Reliance on retrospective case note research Partial participant blinding to research	97.5%

DISCUSSION

Main findings

This systematic review aimed to report on the literature pertaining to length of stay within the medium secure psychiatric estate, with a focus on male populations, and to identify if there were any specific factors that might predict how long individuals remain within forensic psychiatric services.

A total of five studies were included in the review. Results from the five studies varied. There was some consistency between the studies in terms of their findings, but also divergence with regard to what factors associated with a mentally disordered population actually help to understand length of stay within forensic services. The studies appeared to be united in their presentation of results pertaining to variables that showed no relationship with length of stay, especially where socio-demographic data is considered. Two studies stand out from this in their findings using the variable of ethnicity (Edwards, et al., 2002; Shah, et al., 2011), where being White appears to predict prolonged length of stay in the former, and being Black predicts a shorter length of stay in the latter. Suggestions by both sets of authors indicate their assumptions that this finding may be related to the chronicity of the illness within the patient groups that are admitted and the prognoses for recovery. These findings appear unusual amongst other similar reported studies, or even studies that have simply described trends within forensic populations as no correlations were seen amongst any of the other socio- demographic variables and further investigation of this finding may be called for. Edwards, et al. (2002) stipulate that they have made investigation of this a point for further study.

Whilst socio-demographic variables are perhaps less useful in predicting length of stay, the clinical and forensic variables showed more promise when subjected to statistical analyses. One area of consistency appears to be the diagnosis of a psychotic illness, such as schizophrenia, when compared with diagnoses of affective disorders and indeed personality disorders as a predictor of length of stay. The findings from the Shah, et al. (2011) study demonstrated shorter stays for individuals with diagnoses of affective disorders and a personality disorder. Although the results suggesting that having a diagnosis of a psychotic illness, such as schizophrenia, were shown to be statistically significant in the Shah, et al. (2011) study, this is supported to some extent by the work presented by Kennedy, et al. (1995), who conclude from their work that there is an issue of chronicity of illness with psychoses that has a poorer prognosis in terms of treatment resistance than other psychiatric diagnoses. The findings of the papers reviewed also demonstrate the consistency of clinically diagnostic variables and their relationship to length of stay, with the general psychiatry literature. For example, other studies have also found having a diagnosis of schizophrenia to be associated with increased length of stay (Compton, et al., 2006; Hodgson, Lewis & Boardman, 2000; Huntley, et al., 1998). This is similarly the case with research completed in high secure hospitals, where the presence of psychosis was found to be related to increased length of stay (Moran, et al., 1999)

Another factor that could be considered relevant to these findings is that until recently personality disorder has been regarded as a diagnosis associated with 'the untreatable' (Department of Health, 2003), and therefore individuals with personality disorders were often discharged quickly or even excluded from services that were configured

primarily for individuals with mental illness. It is only since 2003 that policy changes have been implemented that mean it is no longer a diagnosis of exclusion (Department of Health, 2003). It may be that work looking specifically at the characteristics of personality-disordered populations as they relate to psychiatric admission prior to and following the change in policy, and as a distinct sample group from those with mental illness, may be of value to the literature.

Three papers converge in their findings that having a history of previous psychiatric admissions and contact was associated with increased length of stay (Castro, et al., 2002; Kennedy, et al., 1995; Shah, et al., 2011). These findings are in keeping with some of the general psychiatry literature (Huntley, et al., 1998). Again this appears to add weight to the issue of illness chronicity as a key theme when trying to understand the reasons behind extended psychiatric hospitalisation.

A clear focus of this review is the setting and population being a forensic one. Therefore, it is perhaps surprising to see a lack of consistency in the studies regarding the relevance of forensic variables in understanding factors associated with length of stay. One of the issues that sets apart those detained within forensic services from those hospitalised within the general psychiatric system is the issue of dangerousness (Department of Health, 2007b). However, the results from the articles reviewed are not united in their support of a relationship between offending and length of stay. Edwards, et al. (2002) and Kennedy, et al. (1995) both found that there was evidence of a relationship between the severity of index offence, typically murder, and increased length of stay. Both have also argued, in addition to Ricketts, et al. (2001), that longer hospital admissions are seen with those who are admitted from the high

secure estate. This is perhaps an unsurprising finding if based on an assumption that those patients considered highly dangerous (by virtue of their offending) would need to be treated within conditions of highest security to begin with. However, other findings suggest that severity of index offence and having a history of offending is inversely related to length of stay (Shah, et al., 2011). This latter finding might appear counter-intuitive to the clinician and criminological researcher, as one might expect more support for the argument that dangerousness would result in a longer detention. However, Shah, et al., (2011) argue that as an independent variable, severity of offence does not itself lead to longer hospital stays. Their finding is associated with individuals within their cohort having had a history of only one incident of severe violence and fewer contacts with the mental health system generally.

The picture presented by these six studies appears confusing in attempting to explain and understand what factors are associated with and potentially predict the length of stay for forensic populations in the medium secure estate. It is difficult to conclude with any confidence whether there are any factors that consistently correlate with length of stay, due to the variance between the study outcomes. Perhaps the only conclusion that can be tentatively stated is that mental health professionals and researchers are developing a clearer understanding of what is not associated with length of stay, i.e., socio-demographic variables, rather than what is.

Strengths And Limitations Of The Review

The aim of the review was to assess the literature base on length of stay in forensic psychiatric hospitals that represent the medium secure estate. One of the requirements of this review was that it should be conducted using a structured, systematic approach

that could be replicated at a later date. Strengths and limitations of the review conducted are now considered.

The small number of studies included in the final review is a potential limiter o**n** the utility of this review. From the outset, difficulties were experienced in sourcing reviews and literature in this area, which whilst useful on the one hand in supporting the need for a review, could also undermine any conclusions that the results are representative of the wider medium secure psychiatric population. The difficulties experienced by the researcher in sourcing citations may have been a contributing factor to the limited number of studies included in the review. Systematic searches were run through a number of electronic databases with the number of citations found varying between being in their tens of thousands to very small numbers registering less than one hundred, and a significant disparity in subject matter. Increasing the number of electronic databases or the use of alternative electronic sources, such as Google, could have resulted in further citations being available for inclusion within the review. For example, although Google was used as a source to establish the existence of any earlier reviews, it was not used specifically for the search of individual citations. With hindsight, the use of Google may have been advantageous from the perspective of triangulation with those citations sourced through established scientific databases. Had this been done, it is possible that other existing cohort studies by well-known proponents within the forensic psychiatry field that did not appear in the systematic search process may have been found. For example, there are cohort studies that consider the outcomes of admissions to medium secure psychiatric hospitals by well-known forensic psychiatry practitioners such as Coid, Kahtan, Gaukt, Cook and Jarman (2001), Maden, Rutter, McClintock, Friendship and Gunn

(1999), and position papers by Taylor, Maden and Jones (1996), which were missing from the electronic search, but which were found through the use of Google. These papers may have been useful additions to the final review.

Initial difficulties within the devised search syntax were identified and were rectified. These difficulties can be summarised as needing to account for differences in international spelling and use of language, hence the use of subject mapping, as well as being specific enough to reduce the likelihood of crossover with physical health subjects. It is also possible that the variations in keyword codings applied to already published articles has resulted in too diverse or too narrow a field in respect of this subject. A retrospective review of the search terms and parameters used specifically within this review suggests that the systematic approach could have been strengthened by the inclusion of additional terms associated with forensic psychiatry and psychology. For example, as well as using terms 'forensic hospital' or 'secure unit', it would have been beneficial to consider historical terms such as 'regional secure unit' and abbreviations including 'RSU' and 'MSU'. Consideration was also given to the search parameters used. The systematic searches were predominantly restricted to keywords within the title and abstracts of the available literature and it may have been advantageous to consider expansion of these search parameters to include 'within text' and supplementary wildcard options. The addition of these terms and search parameters may have increased the final number of relevant citations available for consideration under the PICO structure and within the quality review.

In addition to adopting a broader electronic search strategy, it is possible that had a hand-search of the individual references of a number of the studies included been

conducted, and even some of those excluded as a result of a reasonably tight PICO structure, a wider literature base might have been found. It is possible that the review could have been improved in terms of its quality and findings by utilising this method, as this may have identified any number of relevant citations that were regularly referenced within the studies found, such as the cohort studies by Coid, et al. (2001) and Maden, et al. (1999) mentioned above. Thus a hand-search may ultimately have increased the size of the research base for consideration at the PICO stage or for inclusion within this review. Contacting the authors of the included studies and clarifying any ambiguity found within their articles, or to source any unpublished works or recommended reading that may exist could also have provided additional resources for the researcher conducting the review. Time constraints restricted the use of these approaches. On reflection, if the review were to be repeated, it would need to take into account these limitations and be amended accordingly to make best use of the available literature.

One of the limitations of this review is that the sole researcher conducted both the data extraction and quality assessment processes. It is possible that another reviewer may have scored other articles differently, which may or may not therefore have subsequently been included in the review. However, a limitation of this could have been the potential for a second assessor to exclude studies from the final review, which would have limited even further what was already a small number of remaining articles following the earlier stages within the study selection process. Nonetheless, the availability of a second reviewer would have potentially added weight to the reliability of the decision making process during the quality assessment process.

Turning to the studies themselves, there are a few factors that the researcher has considered which perhaps add to the weaknesses inherent in the conclusions within the articles, and therefore subsequently the wider review itself. One of these is the description the studies provide of their samples. As noted above, the use of psychiatric nomenclature regarding *admission* and *discharge* has the potential to confuse, as both can refer to a discrete period of hospitalisation. It is only through reading the full articles that a researcher can extrapolate the core facets of the sample group. One of the concerns with the use of these terms however is the lack of reference within the articles to the possibility of double counting of participants, as it is possible that their samples include duplications of participants who may have been admitted and discharged on more than one occasions during the sampling timeframe. This is important because several of the articles refer to the variable of previous psychiatric history and admissions as being a potential predictor of length of stay (Castro, et al., 2002; Kennedy, et al., 1995: Shah, et al., 2011). However it is important to note that none of the studies appeared to address the issue of double counting of individual patient's experiencing more than one episode of care within the sample time-frames.

All of the studies included both men and women within their samples. This may reflect the accommodation situation at the time of some of the studies, where units were established to hold both men and women, whereas in more recent times, gender specific units have been developed and opened (Centre for Mental Health, 2010). One must therefore bear in mind the extent to which the overall results are generalisable to the male secure estate, when women have also been a core element of the study samples. This is important given that there is some support for a shorter length of stay

being associated with being female within the general psychiatry literature (Compton, et al., 2006).

A final point for consideration is the choice of research design and statistical analysis used. The five studies reviewed use a combination of both parametric and non-parametric statistics. For example, Castro, et al. (2002) and Edwards, et al. (2002) use t-tests, but then refer to the use of Mann-Whitney analysis, whilst Kennedy, et al. (1995) also use t-tests but then go on to use Chi-Square. Yet none of the papers reference whether tests of normality have been applied, such as the Kolmogorov-Smirnov test. This raises a possible query regarding to what extent the authors of each study had investigated the distribution of their samples before employing their statistical analyses, and raises possible questions about the precision of their conclusions.

Each of the studies has also relied on a cohort-study design. Despite the limitations noted regarding accessing accurate records, the retrospective cohort design does appear to be an appropriate method for these studies. Perhaps an area for consideration in strengthening the methodological approach would be to include the use of more standardised measures. For example, the general psychiatric literature is littered with studies where standardised measures such as the Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1962) or the Global Assessment of Functioning (GAF) (American Psychological Association, 2000). Both have been used and found to be reliable measures for predicting length of stay outcomes in generic adult settings through the formal standardised assessment of the presence of psychiatric symptoms and a person's level of psychosocial functioning (Anderson, et al., 2004; Biancosino,

et al., 2005). In this regard, a structured prospective cohort study might be possible, using an agreed set of standardised measures, such as those seen within the general adult psychiatry literature within a wider range of variables, and analysing their relationship to the length of time a person may be detained within medium secure conditions. It is important to note however, that neither assessment addresses the issues of risk and its relationship to mental illness in a forensic population.

CONCLUSIONS

This review has found that there is conflict in the reported findings regarding what factors predict length of stay within forensic psychiatric populations within the medium secure estate. Of the five studies considered by this review, all reported a number of different findings based on their investigation of socio-demographic, clinical and forensic variables and their relationships to length of hospitalisation. The inconsistency between the findings could be considered as the strongest evidence to support the need for further research in this area. For example, it would be valuable to consider what other variables that have not been previously investigated at all may be relevant, or investigated in only one or two studies, such as the level of engagement in meaningful therapeutic activities, as seen in the study by Castro, et al., (2002). Alternatively, further research could draw more heavily on variables seen within the general psychiatry literature, such as the relevance of institutional controls, including use of seclusion (Compton, et al. 2006), or access to community leave, which has so far not appeared in the empirical literature. It is possible that other factors that might be regarded as less tangible than a clinical diagnosis or a specific section under the Mental Health Act 1983, could add significant value to our understanding of length of stay for mentally disordered offenders. For example, understanding the interaction between a patients' insight into their mental disorder and level of risk, and length of stay.

Further to the development of gender-specific forensic services, it would also be beneficial to consider male and female populations as distinct from each other and report on the differences in length of stay between these two groups. Given the

relatively small numbers of female mentally disordered offenders, as compared with men (Ministry of Justice, 2010), this may be best addressed through a multi-site study, and would potentially have an advantage of being more clinically meaningful through its generalizability to the female secure services estate.

The small number of studies included in the review should be an indicator of the importance of not drawing too many firm conclusions about which factors are relevant. Rather it adds further weight to the need for on-going research in this area, especially when compared with the availability of studies of length of stay for the general psychiatric population. One might consider that a meta-analysis would be a valuable tool that could be employed. The studies that are available do share a common hypothesis – that there are factors that interact and influence length of stay. By bringing all the data together, this would certainly overcome the limitations of sample size and local sampling bias. It would also enable clarification regarding the extent to which any results are truly generalisable across a mentally disordered population detained within the medium secure estate.

As the National Health Service in England moves towards the use of defined care clusters within mental health and Payment by Results, it will become essential for services to demonstrate that clinical care pathways for patients are cost effective and well defined in terms of the services offered, their effectiveness and the timeframe required for patients to recover. The eighteen-month timescale as originally indicated by the Glancy Report (HMSO, 1974) remains a feature of the clinical care pathway within medium secure services. Therefore, it would appear essential that if clinicians are to adhere to this ‘standard’, they require robust evidence which will assist them to

identify those individuals for whom the two year period is a realistic treatment target, and for those where it may be unfeasible and require a differently configured service. Therefore, it is proposed that further research on length of stay is of increasing relevance, as clinicians strive to be evidence based practitioners, working within a system that is supported by valid and clinically meaningful research. The following cohort study in Chapter Three explores a series of factors that may be pertinent to a better understanding of length of stay within a male, mentally disordered offender population, detained within conditions of medium security, building upon the issues and previous studies considered within this review.

CHAPTER THREE

PREDICTING LENGTH OF STAY IN A POPULATION OF MALE MENTALLY DISORDERED OFFENDERS DETAINED IN A MEDIUM SECURE PSYCHIATRIC HOSPITAL

AN EMPIRICAL STUDY

ABSTRACT

Aim

Research has proposed that longer inpatient hospitalisation is associated with increased treatment costs. As a result, the issue of length of stay is increasingly viewed as a key factor for economic control within health services. The aim of this paper is to investigate what factors are predictive for length of stay within a male, mentally disordered population who had previously been detained at a medium secure hospital in the West Midlands.

Methodology

This retrospective, archival study involved a cohort of 198 adult males who had at one time been discharged from the medium secure hospital. Discharge summary reports and HCR-20 risk assessment reports were sourced for all participants who had been discharged within a ten-year period up to and including June 2011. Data collected was analysed using non-parametric statistical tests within SPSS.

Results

The results indicate that some of the variables investigated have predictive value in relation to length of stay within the population studied. Across a series of socio-demographic, clinical and forensic variables, nine variables were found to have a statistically significant relationship to length of stay, with diagnosis of a schizophrenic disorder being the strongest predictor of length of stay over two years. Effect sizes were small to medium.

Conclusions

This study concludes that length of stay may not be a reliable outcome of effective forensic mental health care. Divergence over what factors consistently explain length of stay for mentally disordered populations remains across the reasonably limited literature base. It suggests that there may be other factors more pertinent to the individual experience of recovery that influences the duration of hospital admissions, and there is continued need for investigation of this.

INTRODUCTION

Forensic psychiatric services provide care and treatment to individuals for whom their mental disorder is associated with risk of harm to others, or themselves (Department of Health, 2007). Secure hospitals admit individuals whose mental illness is too severe for them to be managed within the criminal justice system and who are deemed too risky for general psychiatric environments (Walker, et al., 2012). Services are configured across varying levels of security, ranging from the high secure estate, through medium and low secure units, to community and forensic liaison services. Secure services are charged with balancing the healthcare needs of patients with a public protection agenda. Thus, clinicians are responsible for assisting patients in their recovery of their mental illness, preparing them for community-based independence, whilst helping them to desist from offending (Department of Health, 2007b).

The medium secure estate was established with an objective that patients would remain within the service for between eighteen months (Glancy, 1974) to two years (Reed, 1997). However, research has shown increasing numbers of patients who are being detained in medium secure conditions or higher, for longer than the two years originally suggested (Edwards, et al., 2002; Rutherford & Duggan, 2007). Additionally, there has been a year on year increase in secure hospital admissions (Rutherford & Duggan, 2007), with reports that there are now approximately 5,000 patients detained within the medium secure estate, costing on average £200,000 per person, per year (Walker, et al., 2012). Walker, et al. (2012) report that although the medium secure estate "accommodates only a small proportion of psychiatric patients", it "consumes one percent of the entire NHS budget" (p.1). They also reflect that medium secure psychiatry services often care for and treat small numbers of patients

for long periods of time, resulting in a "low-volume, high cost" service (p.1), meaning that in terms of output, they are relatively expensive operations. Thomas, et al. (1997) studied the link between service volume provision and length of stay, where service volume refers to the number of people accessing or passing through a service. After accounting for diagnostic differences, they found that providers of high-volume services had reduced lengths of stay compared with "low-volume providers" (p.1397), suggesting that high-volume providers are more efficient in their practice. The study also showed that for those individuals with a shorter length of stay, there was no increase in recidivism or readmission rates post discharge. This is an important finding, not just for the general psychiatry population, but also for forensic psychiatry services, where reducing recidivism is a central feature of their work.

At face value this study suggests that length of stay can be manipulated and thus shortened, without causing serious negative consequences for individuals or society, because the focus is on the service provider, rather than the characteristics of service users. However, it could be argued that this is too simplistic a conclusion and that their findings should be regarded with some caution. A point for consideration is the choice of population and the extrapolation that shorter lengths of stay have no negative effect on recidivism or readmission rates. The study sample does not appear to be a forensic population. Additionally, no reference is made to the relevant base rates for recidivism and readmission as applies to their sample or the wider population, and no data is provided for scrutiny to this effect. Although they refer to a survival analysis of patients who have passed through both the high-volume and low-volume services and the rates of readmission, they do not report the circumstances of the readmissions. This makes it difficult for any conclusions or inferences to be drawn in

respect of the effect that length of stay had on readmission as an outcome. From a commissioner's perspective, the findings by Thomas, et al. (1997) would perhaps be reason for encouragement that length of stay and the associated costs with longer hospital admissions could be reduced through the modification of services, without significant detriment to society as far as risk was concerned. However, from a clinician's perspective, this study does not sufficiently explain the link between the reduced recidivism and readmission rates and shortened length of stay for one to be confident of causality. Rather, the reasons for readmission and recidivism, which are typically low in mentally disordered populations (Rutherford & Duggan, 2007), are likely to be multi-faceted and idiosyncratic and maybe more dependent on the provision of outpatient service and patient engagement, than the length of time spent in hospital.

Interest in length of stay of patients within psychiatric care is reasonably longstanding. Historically, research has focussed on the identification of factors that are associated with length of stay, typically within the general psychiatric population. There has been an emphasis on individualisation of care and the idiosyncratic nature of recovery (Parsons, 2006), and prominence given to monitoring trends in clinical populations and issues of clinical need (Creed, et al, 1997; Trauer, et al, 1999). Similar patterns of research have appeared within the forensic psychiatry literature, with descriptive population studies looking at factors that are relevant to changing clinical patterns (Edwards, et al., 2002; Glorney, et al., 2010; Kennedy, et al, 1995; Ricketts, et al, 2001).

Length of stay is increasingly viewed as a key factor for economic control within

health services (Parsons, 2006). Research has already proposed that longer inpatient hospitalisation is associated with increased treatment costs (Compton, et al., 2006). Thus a shift has taken place where length of stay is becoming a frequently cited performance indicator within services, and an outcome measure by which efficiency and quality of inpatient psychiatric care can be measured (Castro, et al., 2002; Goldney, et al., 1998; Moran, et al., 1999). The introduction of national frameworks emphasising the importance of quality of care, whilst maintaining fiscal efficiency could be seen as the context for this. Forensic psychiatric services are not exempt from scrutiny or the implementation of new processes, such as Payment by Results (Department of Health, 2011) or the Commission for Quality and Innovation (CQUINs) framework (Department of Health, 2008), of which length of stay has been a central feature. Local commissioning groups also decided early on to emphasise the importance of controlling costs through reducing length of stay, as has been noted earlier.

Compared with general psychiatric populations, there is very little in terms of recent findings regarding meaningful outcomes for mentally disordered offenders, especially what outcomes or predictors are associated with length of stay (Smith, et al., 2004). Despite these inconsistencies, length of stay appears to be gaining prominence as an outcome within forensic psychiatry.

Within the literature that does exist, debate and conflict about what factors reliably predict or influence length of stay for mentally disordered populations continues. Research studies in existence differ in their findings regarding length of stay, although there are some commonalities. Differences across the general and forensic psychiatry

literatures are typically seen in the choice of variables measured within the research, although they characteristically fall into three areas: socio-demographic variables, forensic/ offence related factors, and clinical variables. The Historical, Clinical, Risk–20: Assessing Risk For Violence (Version 2) (HCR-20) (Webster, et al., 1997) is a risk assessment framework used widely within forensic secure services, and has been incorporated into more recent studies, with the presence or absence of the factors assessed within the framework being used as predictor variables (Shah, et al., 2011; Smith, et al., 2004). Otherwise, very few studies conducted within the forensic estate incorporate recognised assessment tools or standardised psychometrics, especially when compared with the general psychiatry research.

Socio-demographic variables are widely reported within studies investigating length of stay. Rarely have they predicted length of stay for a forensic population to the point of statistical significance, and in a number of cases no association has been proved at all (Castro, et al., 2002; Kennedy, et al., 1995; Smith, et al, 2004). However, some studies have reported alternative findings regarding socio-demographic variables. In a study by Edwards, et al. (2002), being of White ethnic origin was a significant predictor of length of stay over five years. Consistent with this, Shah, et al. (2011) found that being of Black ethnic origin was significantly related to discharge prior to two years.

Diagnosis is a factor widely researched with regard to its relationship with length of stay in both the general and forensic psychiatry literatures. There appears to be some congruence across the research that diagnoses of personality disorder and affective disorders are associated with shorter hospital admissions (reduced length of stay)

(Shah et al, 2011). In contrast, having a diagnosis of a psychotic illness such as schizophrenia has been found to be related to longer hospital admissions, although not necessarily at the level of statistical significance (Compton, et al., 2006; Huntley, et al., 1998; Kennedy, et al, 1995; Moran, et al., Shah, et al, 2011, Smith, et al., 2004; Wright, O'Neill & Kennedy, 2008).

Other clinically related variables differ in their prominence within the length of stay literature. For example, a persons' history of previous psychiatric admissions, or length of psychiatric contact, and occasionally age at first psychiatric contact have been investigated in their relationship to length of stay. The number of previous admissions to hospital a person has, has been found to be positively correlated with length of stay (Castro, et al., 2002; Shah, et al., 2011), although the study by Edwards, et al. (2002) found no association. Kennedy, et al. (1995) found that a long history of contact with psychiatric services was associated with increased length of stay, a finding that was later contradicted by Smith, et al. (2004). Edwards, et al. (2002) reported that being older on admission was a characteristic of members of their sample who remained in services for over a five-year threshold.

Forensic variables, such as history of offending or the nature of an index offence, have also been considered within the literature, with mixed results. Convictions for a serious offence such as murder, or significant violence as an index offence, have been shown to be associated with increased length of stay (Edwards, et al., 2002; Kennedy, et al., 1997). Other studies contradict this, with results showing no statistically significant relationship between length of stay and index offence (Smith, et al., 2004), or conversely a negative relationship between length of stay and severity of offending

(Shah, et al., 2011). The same has been reported for having a history of previous offending (Shah, et al., 2011). Another study conducted in a high secure setting found that young age at the point of offending predicted increased length of stay (Moran, et al., 1999). Some studies have also considered institutional risk as a factor that impacts on length of stay (Castro, et al., 2002; Kennedy, et al., 1995), suggesting that offence paralleling behaviour whilst in hospital, or problematic behaviour during admission, are relevant in predicting increased length of stay.

Medico-legal variables are also prominent within the literature for forensic populations. For example, length of stay has been investigated in terms of the legal status of patients. Detention under a hospital order (Section 37 and Section 37 with restrictions (41)) has been shown to be related to increased length of stay (Kennedy, et al., 1995; Shah, et al., 2011).

Within a high secure population, having a history of prior employment was reported as having the highest predictive relevance to length of stay (Moran, et al., 1999), with the presence of an employment history being associated with a shorter length of stay. Similarly, within the same study, poor education history has been shown to be related to increased lengths of stay.

It is clear from the existing literature that inconsistencies exist between studies, making it difficult to conclude with any confidence what factors predict length of stay. This in turn complicates the extent to which any earlier findings can be generalised across forensic populations. This has implications at the clinical delivery level, as clinicians do not have a reliable set of factors to assist the identification of individuals

who are likely to remain in services longer, potentially reducing their ability to accurately allocate to appropriate care pathways. At the service level and commissioning level, this also makes it difficult to marshal financial resources in the current economic climate. It is the aim of this paper to investigate which factors, common within a medium secure, male, mentally disordered population, can be used to predict length of stay.

In particular, this study aims to identify whether a series of specific clinical, risk or socio-demographic indicators, or any combination of these, predict outcome in relation to the length of stay of persons detained within a male medium secure psychiatric hospital. Variables under investigation in this study include ethnicity, age at the point of admission, age at first contact with psychiatric services, diagnosis, legal status, history of previous psychiatric admissions, index offence and history of offending, and employment history, including the relevance of employment status at the point of admission. Moreover, an advantage of this study is the inclusion of a standardised assessment of risk of violence in the form of the HCR-20 (Webster, et al., 1997) as with the exception of the Shah, et al., (2011) paper, there is an absence of standardised assessment measures being used within the existing literature.

Hypotheses investigated in the following study included:

- That a diagnosis of psychosis would be associated with increased length of stay when compared with a diagnosis of personality disorder.
- That early onset of illness (as defined by early age at first contact with psychiatric services) would be associated with increased length of stay.
- That a history of increased psychiatric admissions would be associated with increased length of stay.
- That severity of forensic history would be associated with increased length of stay.
- That being detained under a criminal section of the Mental Health Act, 1983 (hospital order, with and without restrictions) would be associated with increased length of stay, when compared with civil and transitional sections.
- That a history of unemployment and unemployment at the time of admission would be associated with increased length of stay.
- That increased length of stay would be associated with high scores across the HCR-20 risk assessment tool, including scale scores (Historical, Clinical and Risk) and total score.

METHOD

Sample

This retrospective, archival study involved a cohort of 198 individuals who had at one time been discharged from a male Medium Secure Psychiatric Hospital in the West Midlands. All 198 participants were former inpatients, discharged from the hospital within a ten-year period up to and including June 2011. Where cases indicated more than one hospitalisation for an individual over the time period, the most recent admission data was used to reduce a potential bias of double counting within the sample. For the purpose of this study, discharge is defined as a discrete episode of care from admission to hospital to the point of discharge into any other setting, including hospital, prison and the community.

Ethics

The research project was conducted in accordance with national NHS guidance and policy in respect of the use of clinical data and confidentiality, and in line with professional codes of conduct as directed by both the British Psychological Society (2009) and Health Professions Council (2007), and the University of Birmingham's Code of Practice for Research 2010-2011. The Research and Development Unit for the NHS Trust, in which the hospital is located, gave approval for the study to proceed. The research proposal was submitted to and approved by the South Birmingham Research Ethics Committee (NHS National Research Ethics Service). Issues of consent and confidentiality were considered and approval given for the use of existing clinical data without seeking informed consent from sample participants, with patient identifiable data being anonymised at the point of transcription from the

clinical record.

Procedure

All data were compiled by the author, who is a current member of the clinical care team within the identified organisational setting. A comprehensive list of all patients discharged from identified hospital within the timeframe specified was established from an administrative database. Discharge summary reports and HCR-20 risk assessment reports were sourced for all names on the list. Where either one of the two required documents was not available, cases were excluded from the final data set. With 384 cases identified, 186 were excluded due to lack of available records, leaving a sample of 198.

Information was extracted from these documents, comprising socio-demographic data (including age on admission and ethnicity), clinical diagnosis and psychiatric history, legal status, index offence, criminal and employment history, and scale data from the HCR-20. Variables selected for the study were chosen considering their representation in the earlier literature. Length of stay was calculated for each case using the date of most recent admission to the date of discharge from hospital.

Treatment of Data

Data was collated, anonymised and entered into the Statistical Package for the Social Sciences (SPSS), Version 19.0, for analysis. Post-hoc power analyses were completed using G* Power 3.1.5. (Faul, Erdfelder, Lang & Buchner, 2007).

RESULTS

Preliminary analyses were completed using descriptive statistics. All 198 participants in the sample were male. Forty-three percent were White (*n*=85), 36% were Black (*n*=72), and 16% Asian (*n*=31). Three (1.5%) participants were of Middle Eastern origin, with seven (3.5%) reporting being of Mixed race.

Length of stay was calculated by summing the number of days each individual was in hospital from the point of admission to the point of discharge. The mean length of stay was 765 days (*SD*=665.0, *Md*= 645.0, *Range 6* - 4151), with 55% of the sample (*n*=109) staying less than two years and 45% (*n*=89) staying over two years. The duration of stay in number of years across the sample is shown in Table 6.

Table 6. Length of Stay In Years For Study Sample (*N*=198).

Length of Stay	***n***	**%**
Up to 12 months	46	21
13 to 24 months	67	33
25 to 36 months	46	23
37 to 48 months	21	11
49 to 60 months	7	3.5
61 to 72 months	4	2
71 to 84 months	7	3.5
85 to 96 months	2	1
97 months and over	2	1

The results indicated that age on admission ranged from 17 to 65, with the median being 33 years. Analysis of the data when categorised into age-related groups showed that at the point of admission, 13 men were under the age of 21 (7%), just over one-third (33%, *n*=65) were aged between 21 and 30, 60 (30%) were between the ages of 31 and 40, with the remaining 60 participants (30%) aged 41 and older. Eighty-four per cent (*n*=166) of the sample was under the age of 30 at the point of their first contact with psychiatric services.

The majority of the sample had received a formal diagnosis of a schizophrenic disorder (85%, *n*=168), with 9% having a diagnosis of an affective disorder (*n*=17), whilst 5% (*n*=10) had a diagnosis of personality disorder. Three cases (1%) were recorded as having no established diagnosis or a non-major mental illness diagnosis.

Regarding legal status on admission, half of the cohort (50%, *n*= 98) were admitted under transitional arrangements of transfer from prison, under Sections 47/49, 48/49 or 38 of the Mental Health Act 1983, as amended 2007. Forty-nine (25%) participants were admitted under a Criminal Section, including Section 37/41, Section 37, CPIA or Notional 37. Thirty-one (15%) individuals were admitted under a Civil Section (Section 2, Section 3 and CTO). Twenty (10%) individuals were admitted on an Informal basis. The majority of participants within the sample had a violent index offence (80%, *n*=158), 18 (9%) had an index offence of arson, with the remaining group having index offences ranging from driving and drug offences, acquisitive offences and breaches of orders (11%, *n*=22). Within the sample, 88% (*n*=173) had a prior history of offending.

Consideration was also given to the relevance of employment status. The results suggest that within the sample of 198 participants, 80% (*n*= 158) had held some form of employment or full time educational role in the past. In comparison, at the point of admission, this figure had decreased sharply, with only 2.5% (*n*=5) of the participants having been in employment or full time education.

The majority of the sample was discharged from hospital into the community (61%, *n*=121), with 22% (*n*=44) of the sample progressing to conditions of Low Security. Six (3%) participants were discharged into the care of a local non-secure hospital. One (0.5%) individual was moved into the High Secure Estate, with 10 (5%) being discharged to other medium secure hospitals. Eight per cent (*n*=15) of the sample was transferred back to the Prison Service. One individual (0.5%) was subject to international deportation.

Predictors of Length of Stay: Preliminary Analyses.

A Kolmogorov- Smirnov test revealed that the continuous data was significantly different to a normal distribution (see Table 7). As a result, measures of central tendency have been reported in terms of their median and range, rather than the mean and standard deviation (Dancey & Reidy, 2011). Data analysis was completed using non-parametric statistical tests, including Spearman's Rho, Mann-Whitney U and the Kruskall-Wallis test, when exploring differences between variables. Bivariate correlations using Spearman's Rho were completed to consider the relationship between length of stay and a series of continuous independent variables. Effect size for the Spearman's Correlation Coefficient test is demonstrated though the use of *r* (Ferguson, 2009). The approximation of *r* is used for effect size for the Mann-

Whitney U tests (Pallant, 2010). Post hoc analysis of statistical power is represented as 1-β (Field, 2009).

Table 7. Kolmogorow- Smirnov Test Results

Variable	Kolmogorov-Smirnov Z
Length of Stay	.127**
Age on Admission	.084*
Age at First Contact with Psychiatric Services	.153**
Number of Previous Admissions	.199**

*p<.05 **p<.001

Length of Stay and Age Related Variables

The relationship between length of stay and age on admission was investigated using Spearman's Rho Correlation Coefficient. A weak, negative but significant correlation was seen between these two variables (r_s =-.194, R_s^2= .004, n=198, p=.006). This suggests that increased length of stay is associated with younger age on admission. The effect size of this relationship is regarded as small (Cohen, 1988). As age on admission ranged from 17 and 65 (Md= 33), this variable was considered for further investigation, to establish to what extent a young age on admission is associated with length of stay.

Firstly, a Kruskall-Wallis Test was applied to compare the relationship between length of stay and age on admission, considered at five levels. These levels were being under 21 years of age (Gp1, n=13), 21 to 30 (Gp2, n=65), 31 to 40 (Gp 3, n=60),

41 to 50 (Gp4, *n*=46) and finally 51 and older (Gp5, *n*=14). As with the Spearman's Correlation, a statistically significant relationship between length of stay and age on admission was found (X_2=15.788, *p*=.003, η_2=.08), but the effect size is small. However, looking at the respective median scores across these groups, the relationship between the two variables may be more complicated than it first seems (Gp1, *Md*=609; Gp2, *Md*=714; Gp3, *Md*=833; Gp4, *Md*=422; Gp5, *Md*=340), as these appear to suggest that being very young at the point of admission, i.e., under 21, is not highly related to length of stay within this sample. This suggests that the results of the Spearman's Correlation should not necessarily be taken at face value.

Therefore, pairwise differences between these groups were investigated further using Mann-Whitney U tests, with a Bonferroni adjustment (*p*<.005) applied to control for Type One errors. Post hoc power analysis was also completed to control for the presence of a Type Two error. Of the tests completed, two showed a statistically significant outcome. The first of these was the comparison of the two groups 21 to 30 (Gp2, *Md*= 728, *n*=66) and 51 and older (Gp5, *Md*=290, *n*=13) (*U* =214, *z* =-2.843, *p* =.004, *r* =.32, 1-β=0.72), with a medium effect size. The second involved the comparison of the groups 31 to 40 (Gp3, *Md*=805, *n*=59) and 51 and older (Gp5, *Md*=290, *n*=13) (*U* = 180, *z* =-2.979, *p* =.0034, *r* =.35, 1-β=0.71), again indicating a medium effect size. Effect size was calculated as an approximation of *r*, using the following equation suggested by Pallant (2010, p.230): *r* = *z* / square root of *N*, where *N* = total number of cases. This approximation of *r* has been applied throughout the study for the Mann-Whitney analyses. None of the remaining eight results reached statistical significance. The results of these tests can be seen in Table 8.

Post hoc power calculations were completed, with mixed results, suggesting that in some cases, the statistical test applied was of sufficient power to determine the presence or absence of effect. Acceptable levels of power (i.e. >0.8) (Cohen, 1988) were established for the Mann-Whitney tests considered for the comparisons between the 21 to 30 years and 31 to 40 years groups (1-β= 0.99), the 21 to 30 years and 41 to 50 years groups (1-β= 0.98) and the 31 to 40 years and 41 to 50 years groups (1-β= 0.98). However in the remaining comparisons across this variable, power calculations did not meet the acceptable level of 80%, suggesting that it is possible that a Type Two error had occurred. It is suggested that limited sample size within the groups analysed may have played some part in limiting the significance of the statistical comparisons completed. For these cases, G*Power indicated that a total sample of 134 would have been required for the group differences to reach statistical significance at the .05 level.

The two initial outcomes appear to confirm the findings from the Spearman's Correlation and the Kruskall-Wallis test. That is, that length of stay is significantly related to age on admission, but that it is the two age groups of 21 to 30 and 31 to 40 which are associated with increased length of stay, rather than the very youngest age group.

Table 8. *p* Values for Mann-Whitney U Tests for Length of Stay and Age on Admission.

	1	2	3	4	5
1. Under 21	-				
2. 21 to 30	.128	-			
3. 31 to 40	.107	.683	-		
4. 41 to 50	.907	.011	.007	-	
5. Over 51	.151	**.004***	**.003***	.229	-

*$p<0.005$

The relationship between length of stay and age at first contact with psychiatric services was first investigated using Spearman's Rho. In this case, there was no significant correlation between these two factors (r_s =-.010, R_s^2= .000, n=198, p=.886), which indicates that the age at which members of the sample had their first contact with psychiatric service was unrelated to how long they were detained in the identified medium secure hospital. However, as this seemed to contradict earlier studies where a relationship between chronicity of illness and length of stay had been established (Kennedy, et. al., 1995), further investigation was warranted.

To establish whether any specific age category might be related to length of stay, based on the age at first contact with psychiatric services, the Kruskall-Wallis test was applied using four groups. These groups included being under the age of 21 (*Md*= 584, *n*=68), age 21 to 30 (*Md*=645, *n*=98), 31 to 40 (*Md*= 711, *n*=24) and 41 and older (*Md*=939, *n*=8). The results of the Kruskall-Wallis test appears to confirm the outcome of the Spearman's Rho correlation, there does not appear to be any significant relationship between the length of stay and the age that participants had

their first contact with psychiatric services (X_2=.843, p=.839, η_2=.004). No further analyses were completed for this variable as a result.

Length of Stay and Ethnicity

A Kruskall-Wallis Test revealed no statistically significant difference in length of stay and ethnicity across five groups (White, *Md*=627; Black, *Md* =72; Asian, *Md*= 567; Middle Eastern, *Md*= 431; Other, *Md*= 936) (X^2 =6.031, p =.197, η_2= .03).

Length of Stay and Employment

A Mann-Whitney U test revealed no significant difference between length of stay and whether participants had ever been employed. The results showed that length of stay in those with a history of employment (*Md* = 633, *n* = 158) and those without (*Md* = 659, *n* = 40) was not significant (*U* = 3016.5, *z* = -.433, *p* = .658, *r* = .03, 1-β=0.99).

No significant difference between length of stay and employment status on admission was found after using the Mann-Whitney U test. Length of stay did not differ between those who were employed on admission (*Md* = 439, *n* = 5) and those who were not (*Md* = 650, *n* = 193) (*U* = 428.5, *z* = -.427, *p* = .669, *r* = .03, 1-β=0.40). It can be seen here that the post hoc power calculation did not meet the acceptable level of 80%, suggesting that it is possible a Type Two error could have occurred. It is suggested that limited sample size within the groups analysed may have played some part in limiting the significance of the statistical comparison completed. For these cases, G*Power indicated that a total sample of 134 would have been required for the group differences to reach statistical significance at the .05 level.

Length of Stay and Psychiatric Variables

When considering the relationship between length of stay and the number of previous psychiatric admissions, a weak negative, significant correlation resulted (r_s =-.164, R_s^2 = .027, n=198, p=.021). This suggests that increased length of stay was associated with a person having fewer previous psychiatric admissions. The effect size here is small (Cohen, 1988).

A Kruskall-Wallis Test found a statistically significant relationship between length of stay and diagnosis, considered at three levels. These levels were having a diagnosis of schizophrenic disorder (n= 168), affective disorder (n=17) and also personality disorder (n=10). The group having a diagnosis of a schizophrenic disorder recorded a significantly higher median score (Md=707.5) than the other two diagnostic groups, which recorded median scores of Md= 239 and Md=354.5 respectively (X_2=12.173, p=.002, η_2= .06), suggesting that it is the diagnosis of a schizophrenic disorder which is associated with longer periods of hospitalisation within this sample.

Pairwise differences between these groups were investigated further using Mann - Whitney U tests, again using a Bonferroni adjustment (p<.017) to control for Type One errors, and a post hoc power analysis to control for Type Two errors. There was a significant difference in length of stay for those with a diagnosis of a schizophrenic disorder compared with a diagnosis of affective disorder (U = 750.5, z = -3.220, p = .001, r =.24, 1-β=0.86). This suggests that, when considering the mean rank data, having a diagnosis of schizophrenic disorder is highly significantly associated with increased length of stay, in comparison to having a diagnosis of an affective disorder, although the effect size is small. Comparing the presence of a diagnosis of a

schizophrenic disorder and a diagnosis of personality disorder and their effect on length of stay, a Mann-Whitney U test revealed no significant difference (U = 592.0, z = -1.567, p = .117, r =.117, 1-β=0.67). Comparing the presence of a diagnosis of an affective disorder and a diagnosis of personality disorder and their effect on length of stay, a Mann-Whitney U test revealed no significant difference (U = 69.0, z = -.803, p = .422, r =.154, 1-β=0.47). The post hoc power analyses indicate the possibility of a Type Two error in respect of the second and third Mann-Whitney test outcomes. Scrutiny of the group sizes again suggests that it is possible that this has been a limiting factor in establishing the presence of a statistically significant outcome. G*Power indicated that a total sample of 134 would have been required for the group differences to reach statistical significance at the .05 level.

The results of the Mann-Whitney U tests appear to support the outcome of the Kruskall-Wallis test. That is, having a diagnosis of a schizophrenic disorder is related to increased length of stay within the study sample, when compared with having a diagnosis of affective or personality disorder. Given the limited power of the subsequent Mann-Whitney tests, it cannot be assumed that there is no relationship between length of stay and having alternative diagnoses of personality disorder or affective disorder. Therefore these results should be treated with some caution. Further analysis using a larger sample would be advised here.

Length of Stay and Forensic Variables

Across five groups, a Kruskall-Wallis test showed no statistically significant relationship between the nature of the index offence and length of stay (X_2=4.475, p=.346, η_2= .022) (Violence, Md= 102, n=138; Sexual, Md=105, n=20; Acquisitive,

Md= 9, *n*=9; Arson, *Md*= 18, *n*=18; Other, *Md*= 11, *n*=11, Other). A Mann-Whitney U test was used to investigate the relationship between length of stay and having a history of offending. The results showed no significant difference between length of stay and not having a history of previous offending (*Md* = 860, *n* = 25), or having a history of previous offending (*Md* = 627, *n*= 173) (*U* = 1816, *z* = -1.294, *p* = .196, *r* = .09, 1-β=0.95).

Length of Stay and the HCR-20.

Spearman's Rho Correlation Coefficients were used to investigate the relationship between length of stay and the scores from the HCR-20 risk assessments for participants. The details of these can be found in Table 9. In summary, no significant correlations were seen across the HCR-20 factors and scale totals with length of stay, with the exception of one variable. This was the factor C2 (Negative Attitudes). The outcome of this correlation showed a weak negative relationship between length of stay and the factor assessed under Negative Attitudes (r_s =-.166, R_s^2 = .027, *n*=198, *p*=.019). The effect size would be considered weak (Cohen, 1988). This suggests that where the factor was considered to be present, i.e., a participant was considered to have negative attitudes, as measured against the factor definition within the HCR-20, this was inversely related to their length of stay.

Weak positive relationships were seen between length of stay and factors H1: Previous Violence (r_s =.135, R_s^2 =.018), H5: Substance Use Problems (r_s =-.119, R_s^2 =.014), C3: Active Symptoms of Major Mental Illness (r_s =.119, R_s^2 =.014), C5: Unresponsive to Treatment (r_s =.109, R_s^2 =.011), but none at the point of statistical significance.

Table 9. Spearman's Rho Correlation Coefficients for Length of Stay and the HCR-20 Items Scale Totals and Full Total (*N*=198).

HCR-20 Factor	r_s	*p*
H1: Previous Violence	.135	.058
H2: Young Age at First violence	-.003	.965
H3: Relationship Instability	.058	.417
H4: Employment Problems	-.069	.333
H5: Substance Use Problems	-.119	.094
H6: Major Mental Illness	.076	.289
H7: Psychopathy	.006	.938
H8: Early Maladjustment	-.089	.212
H9: Personality Disorder	-.040	.577
H10: Prior Supervision Failure	-.034	.643
C1: Lack of Insight	.013	.854
C2: Negative Attitudes	**-.166***	**.019**
C3: Active Symptoms of a Major Mental Illness	.119	.096
C4: Impulsivity	-.090	.210
C5: Unresponsive to Treatment	.109	.128
R1: Plans Lack Feasibility	-.029	.685
R2: Exposure to Destabilisers	-.083	.247
R3: Lack of Personal Support	.023	.742
R4: Non-Compliance with Remediation Attempts	-.024	.741
R5: Stress	-.057	.427
Historical Scale Total	-.074	.300
Clinical Scale Total	-.009	.901
Risk Scale Total	-.037	.602
HCR-20 Total	-.058	.416

Length of Stay and Legal Status

The Kruskall-Wallis Test was used to consider the influence of legal status (i.e., section of the Mental Health Act 1983) on length of stay, using four groups: informal admissions (n=20), civil admissions (n=30), criminal admissions (n=50) and transitional admissions (n=98). The test revealed a statistically significant difference in length of stay across these four groups (X_2 = 17.058, p=.001, η_2 =.086). Inspection of the median scores across the four groups shows that the transitional group had the highest median score (Md=779), with the other groups recording smaller median scores (Criminal, Md=655; Civil, Md=385; Informal, Md=316).

Pairwise differences between these groups were investigated using Mann-Whitney U tests, again using a Bonferroni adjustment (p<.008) to control for Type One errors. Post hoc power analysis was completed to control for Type Two errors. At a first level of analysis, the effect on length of stay by having a civil or a criminal section was considered. The Mann-Whitney U test indicated a significant difference between length of stay and having a criminal section and those having a civil section (U = 463, z = -2.852, p =.004, r=.32, 1-β=0.92). This suggests that increased length of stay is associated with the application of a criminal section, such as a 37/41 (Hospital order with restrictions).

A second analysis considered the difference between a criminal section and having a transitional section, such as a Section 48/49 or 47/49 (where individuals are transferred from prison to hospital conditions) for length of stay. The Mann-Whitney U test revealed no significant difference between length of stay and section (U= 2441.5, z = -.034, p = .973, r = .002, 1-β=0.99).

A third analysis considered the difference between a civil section and having a transitional section, such as a Section 48/49 or 47/49, for length of stay. The Mann-Whitney U test revealed a statistically significant, but weak relationship between length of stay and section ($U = 929$, $z = -3.043$, $p = .002$, $r = .27$, 1-β=0.96).

Three other analyses were conducted comparing an informal admission with those admitted under the three other sections of the Mental Health Act 1983, in terms of length of stay. Two analyses of length of stay for informal admission and firstly criminal, and secondly transitional sections, showed a statistically significant relationship with length of stay at p<.008 (adjusted p). Length of stay was significantly shorter for an informal admission that for a criminal section ($U = 291$, $z = -2.717$, $p = .007$, $r = .32$, 1-β=0.83). The same pattern was seen for an informal admission compared to a transitional section ($U = 577$, $z = -2.891$, $p = .004$, $r = .27$, 1-β=0.88). A Mann-Whitney U test showed no significant relationship between length of stay and having either an informal admission or admission under civil section ($U = 256.5$, $z = -.861$, $p = .389$, $r = .12$, 1-β=0.76). Again it is noted that the post hoc power analysis suggests the possibility of the presence of a Type Two error for this outcome, as power is calculated at below the recommended 0.80 (Cohen, 1988) and may be accounted for by a limited sample size within these groups. G*Power indicated that a sample of 96 would have been required for the group differences to reach statistical significance at the .05 level.

Length of Stay and Discharge Outcome

Length of stay was compared with discharge outcome, considering the three levels of a successful discharge (i.e. being discharged to conditions of lesser security), a

negative discharge (i.e. being discharged to conditions of higher security) and finally a neutral discharge (i.e. a sideways move to another medium secure unit or prison). No significant association was found between any of these discharge outcomes and length of stay (Negative Outcome, *n*=1; Successful Outcome, *n*= 171; Neutral Outcome, *n*= 26) (X_2=3.536, p=.171, η_2= .018).

Predictors of Length of Stay: Secondary Analyses

A binomial logistic regression using forced entry was completed to gauge the impact of nine independent variables on length of stay. Logistic regression was used as the continuous variable of length of stay did not conform to the principles of normal distribution, meaning that multiple regression analysis could not be considered for the secondary analysis (Pallant, 2010). Two categories of length of stay were created, comprising length of stay up to two years, and length of stay of two years and over. This reflected the original premise that medium secure patients would remain within the service for up to two years (Reed, 1997), but also to take account of the findings of other research studies which have shown increases in the numbers of patients detained for longer than this (Edwards, et al., 2002; Kennedy, et al., 1995; Rutherford & Duggan, 2007).

The nine variables that were selected, having demonstrated a statistically significant relationship to length of stay in the preliminary analyses, were age on admission, number of previous psychiatric admissions, C2: negative attitudes (HCR-20 factor), diagnosis of schizophrenic disorder, diagnosis of affective disorder, section on admission: criminal, section on admission: civil, section on admission: transitional and section on admission: informal. Forced entry method was used because it does

not discriminate between variables when they are entered into the regression model and should reduce the likelihood of variability if the model were subject to retesting, as might occur when using stepwise methods (Field, 2009).

Prior to this, it was important to confirm that the assumptions of logistic regression analysis were met (Pallant, 2010). Firstly, consideration was given to the adequacy of total sample size. There were 198 cases and nine independent variables. Peduzzi, Concato, Kemper, Holford and Feinstein (1996) recommend a minimum ratio of 10:1 for valid cases in the smallest criterion group to independent variables. Using this calculation, with length of stay of either up to or two or more years, the inclusion of nine independent variables would require a minimum of 90 cases for each of the dependent variable categories. The distribution of cases into each category was up to 2 years, *n*=109 (55%), and 2 and over years, *n*=89 (45%). Despite having one fewer case than recommended by Peduzzi et al., (1996), the analysis proceeded with nine variables.

Secondly, the presence of collinearity between variables was assessed, to ensure that the independent variables were not strongly related to each other. Field (2009) recommends appraisal of the Variance Inflation Factor (VIF) and tolerance statistic to assess for multi-collinearity, with VIF values of 10 or more and tolerance values of less than 0.1 being a cause for concern. The collinearity diagnostics showed that none of the variables were highly correlated with each other, with all tolerance values being higher than 0.1 and VIF values less than 10.

Goodness of fit tests indicated that that the model performed well with the

independent predictor variables included. The outcome of a Hosmer-Lemeshow Test indicated support for the model by the non-significant chi^2 ($X_{2=}$7.994, p = .434).

Using the intercept-only model as a baseline, a decrease in the likelihood ratio was seen with the application of the full model, from Λ=272.46 to Λ=247.17, suggesting an improvement over the intercept-only model (Field, 2009; Peng, Lee & Ingersoll, 2002). Improvement was also seen between the predicted and observed outcomes with a 12% increase from 55.1% to 67.2% when the predictor variables were included compared with the intercept-only model, continuing to support the application of the regression model.

The full model containing all predictors was statistically significant (X_2 (9, N=198) = 25.29, p<.001). This suggests that the model was able to distinguish between those cases with a length of stay of up to two years, and a length of stay of two years or more. The model explained between 12% (Cox and Snell R square) and 16% (Nagelkerke R squared) of the variance in length of stay. The model was able to correctly classify 69% of cases for whom length of stay was less than two years, and 65% of those who stayed for two or more years, with an overall correct classification rate of 67.2%.

Only one variable made a unique statistically significant contribution to the model (Primary Diagnosis: Schizophrenic Disorder). The strongest predictor of length of stay of 2 or more years was having a primary diagnosis of schizophrenic disorder, recording an odds ratio of 3.348. This suggests that individuals with a primary diagnosis of schizophrenic disorder where three times more likely to have a hospital

admission of two or more years duration, than those with an alternative diagnosis, when controlling for other factors in the model. The results of the logistic regression for the individual predictors are presented in Table 10.

Table 10. Logistic Regression Predicting Length of Stay

	B	S.E.	Wald	df	*p*	Odds Ratio	95% C.I For Odds Ratio (OR) Lower	Upper
Age on Admission	-.014	.018	.612	1	.434	.986	.953	1.021
Number of Previous Psychiatric Admissions	-.008	.036	.045	1	.833	.992	.925	1.064
C2: Negative Attitudes	-.352	.202	3.052	1	.081	.703	.474	1.044
Diagnosis: Schizophrenic Disorder	1.208	.476	6.451	1	**.011***	3.348	1.318	8.504
Diagnosis: Affective Disorder	0[a]							
Legal Status: Civil Section	-.533	.674	.624	1	.430	.587	.157	2.201
Legal Status: Criminal Section	.414	.601	.475	1	.491	1.513	.466	4.917
Legal Status: Transitional Section	.779	.602	1.674	1	.196	2.179	.670	7.089
Legal Status: Informal Section	0[a]							
Constant	-.860	.995	.747	1	.387	.423		

a. This parameter is set to 0 because it is redundant

DISCUSSION

The aim of this paper was to investigate what factors, common within a medium secure, male, mentally disordered population, could be used to predict length of stay. Specifically, this study aimed to identify whether a series of clinical, risk or socio-demographic indicators, or any combination of these, predicted outcome in relation to the length of stay by persons detained within a male medium secure psychiatric hospital.

Summary of Results

The results suggest that only one of the variables investigated, that of having a diagnosis of schizophrenic disorder, had predictive value in relation to length of stay within a population of male, mentally disordered offenders who have been previously detained within a medium secure psychiatric hospital. Across a series of socio-demographic, clinical and forensic variables, nine were found to have a statistically significant relationship to length of stay. Effect sizes were small to medium (Cohen, 1988).

The results of the inferential statistical analyses supported two of the research hypotheses. These were that a diagnosis of psychosis (i.e. schizophrenic disorder) and being detained under Criminal Section of the Mental Health Act, 1983, would predict increased length of stay. An associated finding is that those detained under transitional sections are also more likely to experience increased length of stay than those detained under informal or civil sections.

Nine variables were entered into a logistic regression model to assess to what extent they predicted length of stay of less than two years or two years or more. The full regression model results were statistically significant, indicating that it could distinguish between those who would stay less than two years, and those who would stay for two years or more, with an overall correct classification rate of two-thirds. However, only one variable made a statistically significant contribution to the model, that being the presence of a diagnosis of schizophrenic disorder.

The results did not support the hypothesis that early onset of illness, as defined by early age at first contact with psychiatric services, would be associated with increased length of stay. There was also no support for the hypothesis that a history of increased psychiatric admissions would be associated with increased length of stay. Conversely, having a history of numerous psychiatric admissions appeared to indicate reduced length of stay. The results did not support the remaining research hypotheses: that severity of forensic history, or the lack of an employment history or employment on admission or that high scores across the HCR-20 risk assessment, would be associated with increased length of stay. Although attempts were made to control for Type One errors at the point of statistical analyses, post hoc power analyses indicated that some of the non-significant results might have been subject to a Type Two error. These latter findings suggest that it might be too early to conclude that there is an absence of a statistically significant relationship between length of stay and the variables considered and any formal conclusions regarding factors influencing length of stay should be regarded as tentative. It is possible that the issue of sample size was an influencing factor and would need to be considered when carrying out similar future research.

Although the literature exploring and describing length of stay for mentally disordered populations is limited, there are indications across a number of studies that certain factors appear to be related to the length of time a person may be detained within conditions of medium security. For example, Kennedy, et al., (1995) suggest that factors such as chronicity of a person's illness, issues around diagnosis, the number of admissions, and risk history, are likely to be factors in the length of detention of mentally disordered offenders within medium secure psychiatric services (Kennedy, et al., 1995). This study sought to investigate whether these factors, either individually or to any degree in combination might increase our understanding of length of stay.

The relationship between length of stay and diagnosis has been evaluated within the general psychiatry literature (Creed, et al, 1997), and in studies looking at the population characteristics within medium secure, forensic psychiatric populations (Edwards et al., 2002; Kennedy et al., 1995; Shah et al., 2011). This study found a statistically significant relationship between a diagnosis of a schizophrenic disorder and increased length of stay, and that affective and personality disorders are more likely to be associated with shortened admissions. This finding is congruent with the existing literature for medium secure populations (Shah, et al., 2011). However, this study found no relationship between length of stay and the number of previous admissions to hospital. This was a somewhat surprising result, as other studies have found the relationship between length of stay and previous psychiatric admissions to be positively correlated (Shah, 2011).

This study also found outcomes that support earlier research findings in relation to the impact that legal status has on length of stay. For example, across a number of studies length of stay has been found to be longer for those who are detained under Hospital Orders (Section 37, MHA 1983), especially when accompanied with Ministry of Justice Restrictions (Kennedy, et al., 1995; Shah, et al., 2011). This study found a similar outcome. However, it also indicated that being detained under transitional sections, such as transferred prisoner status (remand or sentenced), was an influential factor for increased length of stay, especially when compared to civil sections or informal status and would warrant further investigation.

The extent to which legal status influences length of stay is an interesting point, and from a clinician's perspective, it is perhaps unsurprising that increased length of stay is associated with detention under a hospital order, with or without restrictions, when compared with those under civil sections. For some service users, final decisions regarding detention and discharge lie outside the jurisdiction of the treating clinicians, because they are detained under a restricted hospital order (Section 37/41 of the Mental Health Act 1983) and therefore subject to restrictions overseen by the Ministry of Justice. Thus the Minister for Justice has overall responsibility for decisions regarding discharge from hospital (Department of Health, 1983). This is in direct contrast to those service users detained under Section 3 of the Mental Health Act 1983, where decisions regarding discharge from hospital rest with a tribunal panel. This represents a potential problem in the value of length of stay as an outcome measure for forensic mental health services, when for approximately 20% of service users the length of their detention is controlled by an external authority (Rutherford & Duggan, 2007). An investigation of the extent to which individuals detained under

criminal sections remain in hospital in excess of the clinical recommendations of their treating professionals, when taking into account the opinions of the Ministry of Justice would potentially contribute to the debate.

The results of the systematic literature review in Chapter 2 of this thesis highlighted the disagreement within the existing literature regarding length of stay and the relevance of a number of variables for mentally disordered offenders within the medium secure estate. Some studies have found statistically significant relationships and associations between variables, where others do not. This study was not immune to such findings. For example, no significant relationship between length of stay and socio-demographic variables, such as ethnicity was found. In other studies, being of Black ethnicity has been found to be associated with decreased length of stay (Shah, et al., 2011) whilst in another, being White was related to increased length of stay, especially when considered alongside other relevant clinical variables (Edwards, et al., 2002). Explanations given for these findings suggest significant differences in the chronicity of the illness across different ethnic groups in the samples investigated. Despite a reasonably diverse sample, no such association was found for this study.

Conflict within the literature has also been seen when considering age related variables. For example, Kennedy, et al., (1995) found a statistically significant relationship between length of stay and the duration of a person's psychiatric history, in terms of first contact with psychiatric services. Findings from a later study by Edwards, et al., (2002) found no relationship between the two variables. This study does not support the finding by Kennedy and colleagues. However, it did find that the age on admission was related to length of stay, although not necessarily in the way

that might have been expected, and it is possible that the lack of statistical power inherent in these analyses may have been a factor here. The study by Edwards, et al., (2002) for example, found that those individuals within their sample who were detained for over five years, were significantly older than those detained for shorter periods of time. This study suggested that it was in fact the group between the ages of 21 and 40 years of age, who were likely to be detained for longer than two years, although the effect size of this test was considered weak, and statistical power was potentially an issue in terms of the sensitivity of the additional group comparisons. Considering the possible explanations for this finding within this sample, it could be hypothesised that service users of greater age may have already had previous admissions and therefore any later admissions were for respite reasons and therefore a shorter admission, rather than for the purpose of high intensity treatment intervention requiring lengthy hospitalisation. It is hypothesised that for the youngest of the age groups considered, length of stay was shorter because they may have been admitted at an earlier stage in the development of their illness and more treatment responsive when compared with their older peers, who may be more treatment resistant.

Earlier studies have also found that factors such as the prevalence of substance misuse problems, a lack of stable relationships and employment factors have also been influential in explaining length of stay (Castro, et al., 2002; Kennedy, et al., 1995; Shah, et al., 2011). These factors were considered in the current study through the use of the HCR-20, a reasonably well known and broadly adopted, standardised assessment framework in forensic mental health services (Douglas & Reeves, 2010; Webster, et al., 1997). It was decided to include the HCR-20 in this study because it would provide insights into a number of clinically and forensically relevant factors,

such as having a history of substance misuse (H5), diagnostic features (H6 and H9) and family (H3 and R3) and employment status (H4) for the study sample. Compared with the other studies above where associations have been found between these factors and length of stay, this study revealed no such associations. For example, employment does not appear to be an influencing factor on how long individuals remained in hospital within the sample used.

Of particular interest was the one reported statistically significant relationship between length of stay and the item assessed under Factor C2 within the HCR-20. This factor is called "Negative Attitudes", and is defined as "the kind of pro-criminal and anti-social attitudes that have some likelihood of eventuating in violence" (Webster, et al., 1997, p.53). At face value, one might hypothesise that for a forensic population, the presence of negative attitudes (as defined within the HCR-20) would be associated with increased length of stay. This is based on an assumption that individuals adhering to such attitudes might be considered difficult to treat because of their anti-authoritarian or anti-social attitudes and therefore be considered at higher risk of violent recidivism. Thus they might be assessed as having more intense intervention needs, therefore requiring longer periods of hospitalisation. However, this study found a negative relationship between length of stay and the presence of this risk factor, suggesting that individuals scoring highly on this item were likely to remain in hospital for less time than their detained counterparts.

In attempting to offer an explanation for this, it is important to consider the clinical biases that might be inherent within the scoring of this item within a forensic setting (Moore, 1996). Mentally disordered offenders differ from the main psychiatric

population by virtue of their risk of offending, usually in terms of their risk of harm towards others and in some cases their proclivity for general offending (Department of Health, 2007b; Walker, et al., 2012). Thus mentally disordered offenders are stigmatised firstly by the shadow of their offending and their alienation within society, and secondly as a result of their mental health issues (Bradley, 2009). It is also important to consider the challenges that clinicians experience engaging and working with mentally disordered offenders (Hodge & Renwick, 2002). At a clinical level, mentally disordered offenders can present as anti-authoritarian, subversive and resistant to opportunities for remediation. However, rather than seeing this as a manifestation of a person's mental health difficulties, such behaviours have the potential to be attributed to inherently negative characteristics of that person's personality. It is hypothesised that where such inferences converge to influence opinions regarding 'treatability', such a person being assessed under the HCR-20 framework is more likely to be coded highly on the C2 item when compared with their more compliant peers, and thus potentially liable to be discharged out of the system within shorter timeframes, whether this is back to prison, or to conditions of long term secure care. This issue also highlights a difficulty of potential double coding within the HCR-20 risk assessment framework, as resistance to treatment opportunities, whether by virtue of a person's capacity to engage or not, should be coded under the alternative item of C5: Unresponsive to Treatment.

Other areas that have shown some relation to length of stay within the literature include variables that would be associated primarily with forensic populations. This includes the nature and severity of index offences and prior offending histories. However, differences between research findings are also evident here. For example,

Edwards, et al., (2002) and Kennedy, et al., (1995) both found that an index offence of serious violence was associated with increased length of hospitalisation. These results have been contradicted by the study by Shah, et al., (2011), which indicated that severity of an index offence was in fact related to shorter stay. The current study found no relationship between length of stay and index offence types, or having a history of prior offending. This could be considered an interesting and somewhat paradoxical result, given that forensic risk is what typically distinguishes a forensic population from general adult psychiatry (Department of Health, 2007b: Walker, et al., 2012), and potentially points the need for further investigation of the relevance of risk related factors in understanding length of stay.

A factor potentially influencing the outcome of this study was the use of a two-year cut off point within the logistic regression model. This was selected to reflect the original stated aims that the medium secure psychiatric estate would work to the objective that patients would remain within the service for up to two years (Glancy, 1974; Reed, 1997), but also to take account of the findings of other research studies which have shown increases in the numbers of patients detained for longer than this (Edwards, et al., 2002; Rutherford & Duggan, 2007). This study found only one variable that predicted length of stay of two or more years. It is possible that alternative analysis using a multiple regression model would have yielded different results, and this may be a point for future research into this subject. However, the influencing factor here was the violation of the normality assumption for a multiple regression.

Methodological Considerations

A number of limitations need to be considered within this study. The first of these is the choice of methodological design. This study involved a retrospective cohort design. Whilst consistent with earlier research designs on this subject, cohort studies are not without their limitations and these will be relevant to the conclusions that can be drawn from this study. For example, this retrospective study relied heavily on the accuracy of the data recorded within clinical records held by the service where the research was undertaken. Attempts to control for errors and missing data were made by including only those cases where the two key clinical documents were available. Nonetheless, there is the potential for bias to be present by assuming consistent quality within the sample timeframe, as well as consistency within the scoring of the HCR-20 risk assessment documents.

Retrospective cohort studies have the advantage of using defined samples, but often require large samples for meaningful conclusions to be established. The sample used within this study (*N*=198) is by no means large, although it appears appropriate relative to similar studies already in existence. For example, the sample in Shah, et al's., (2011) study was 259. Attempts to control for sample size were made, particularly to ensure that sufficient cases would be included for the regression analysis to be meaningful. However, post hoc power analyses indicated that some of the group comparisons tests lacked statistical power and the limited numbers within the groups considered might have been a factor, finding no relationship between variables where one indeed may exist. Additionally, a criticism of this study is the use of nine independent variables within the regression analysis, without fully meeting the threshold for cases within each dependent variable group as recommended by Peduzzi,

et al., (1996). For the group who were detained for over two years, the sample group was one case shy of the required 90 cases. Therefore any conclusions drawn from the regression analysis should be tentative.

Another limitation concerning the sample is the extent to which the results from this study can be truly representative of the particular research site population. As has been aforementioned, cases were selected on the basis of the availability of two key clinical documents – a discharge summary and an HCR-20 risk assessment. Although the timeframe over which discharges was reasonably long, the absence of one or each of these documents resulted in the exclusion of nearly half of the potential sample from the outset. It is difficult to surmise to what extent the research outcomes may have been different if these individuals had been included in the sample. Nonetheless, alongside the study by Shah, et al., (2011), this study does represent a recent investigation into a key area of debate within mental health services currently. Earlier papers have only described the characteristics of a population and its reference to discharge outcome, rather than necessarily seeking to explain the relationships between these characteristics and length of stay (Edwards, et al., 2002; Kennedy, et al., 1995; Mohan, et al., 1997; Ricketts, et al., 2001). Therefore, along with the paper by Shah, et al., (2011), it is hoped that this study represents an addition to the literature and furthers the debate on the utility of using length of stay as an outcome indicator, when there is relatively little agreement on what factors are relevant.

Another advantage of this study is the choice of variables that were considered for analysis. The variables selected were chosen for their relevance to and appearance within previous research. This would to some extent control for researcher bias, who

could have selected variables for their appearance of being connected with length of stay on the basis of her clinical experience. A second advantage could be the potential for generalisation of findings across a wider mentally disordered population when similar or equal variables have been considered. However, across all the various studies that have been previously published, it is possible that there are other relevant factors that have not been considered and represent a gap in the investigation of this issue. For example, qualitative variables such as the nature of the therapeutic relationship or the degree of insight a patient has into their illness and risk, often the intangibles of therapeutic work, may have a bearing on length of stay, but this has been significantly under-represented in the literature to date.

Conclusions and Future Directions

Length of stay is an area of growing interest and importance in the field of forensic mental health, especially where cost control factors are relevant. The results of this study share commonalities with other research available on length of stay. Similarities have been found across the results of this study and those previously in existence. However, it has also shown as many differences. This continues to highlight the variability that exists in our understanding of what factors contribute to the duration of hospital admissions for mentally disordered offenders in medium secure settings.

One factor that might offer an insight into this variance is the consideration of the heterogeneity of mentally disordered offenders generally (Cohen & Eastman, 2000) and what helps a person recover at the individual level. Recovery from mental illness has been described as “a deeply personal, unique process of changing one’s attitudes, values, feelings, goals, skills and roles” (Anthony, 1993, p.17). Therefore, a factor

that is relevant to one person's recovery will differ from another's. What takes one person two years to achieve may take considerably longer for another, for a whole variety of reasons. It is possible that there are more qualitative issues that are of greater influence in understanding the length of time a person may be detained for, and this may essentially reflect the nature of treatment itself and a person's acceptance and response to that treatment. Such variables have not been obvious in the literature to date.

The current literature seems to be indicating that the value of diagnostic labels and legal categories play only a minor role in understanding length of stay, or that they are at the very least inconsistent in their relevance. A potential way forward is for NHS sites to combine their research data into a large study investigating whether there are factors that regularly and consistently predict length of stay. The advantage of a larger sample may be the generalizability of findings. The conundrum of this is that even with a multi-site study, the issue of what predicts length of stay may remain unresolved.

The implication of this at the practice level is whether the focus might then shift from length of stay as being the successful marker of progress and efficiency, to alternative outcomes, such as reductions in readmissions or recidivism rates. At the clinical level, these may be preferable to a treating professional and more consistent with what they would hope for their clients – that a patient would leave hospital having recovered and being able to build a pro-social and good life for the future. However as alternative outcomes, even recidivism and readmissions are complicated factors to consider, given that the base rates for recidivism within mentally disordered

populations are low, especially when compared with a discharged prison population group (Rutherford & Duggan, 2007). Similarly, the rationales for readmissions are equally as complex, when one considers that a short, respite or informal readmission may be a more positive outcome for a patient and service, than a longer readmission following relapse and recall by the Ministry of Justice, or indeed the commission of a violent offence leading to a hospital order with restrictions. For a clinician, the pressure of a two-year ceiling in which to deliver a person-centred care pathway which encompasses both mental health treatment and risk reduction strategies, feels accomplishable for some, but not for others. Therefore, one might suggest that length of stay has gained prominence because it is more easily measurable in the short term, rather than because it is clinically meaningful. It could, and perhaps should be alternatively proposed, that until we can confidently assert what factors or what combination of factors enable us to offer a meaningful explanation of what drives the length of time a person will remain in a medium secure psychiatric service, it would be increasingly beneficial to look towards alternative measures of effectiveness and longer-term, but sustainable efficiencies.

CHAPTER FOUR

THE HISTORICAL, CLINICAL, RISK – 20 (HCR-20): ASSESSING RISK FOR VIOLENCE (VERSION 2.0)

CRITIQUE OF A PSYCHOMETRIC ASSESSMENT

INTRODUCTION

With heightened public knowledge and concerns about violence within society, especially within mentally disordered populations, and the reduction of public confidence in mental health services (Maden, 2007), public protection has increasingly become the *sine qua non* of the world of forensic psychiatry and psychology. Public inquiries into serious cases of violence enacted by mentally disordered individuals, from the Clunis Inquiry (Ritchie, Dick & Lingham, 1994) to the Barrett Inquiry (Robinson, Fenwick & Wood, 2006), have all highlighted the difficulties associated with, and increasingly the importance of predicting and managing the risk posed by such individuals. It is argued that through the development and application of scientific principles in the field of risk assessment, more rigorous approaches to risk assessment and prediction has enabled clinicians to minimise the shortcomings associated with their earliest practices in managing dangerousness (Douglas & Reeves, 2010; Jones & Plowman, 2005; Maden, 2007).

This review introduces and appraises the Historical Clinical Risk–20: Assessing Risk for Violence (HCR-20) (Version 2) (Webster, et al, 1997). The HCR-20 is a framework for assessing violence recidivism, which has become widely adopted within the field of forensic psychiatry (Douglas & Reeves, 2010; Webster, et al., 1997), for which a key objective is balancing "the twin factors of treating mental disorder and managing and reducing risk" (Department of Health, 2007a, p.9). However, the use of structured professional judgment approaches to risk assessment, such as the HCR-20, is also embedded within forensic practice policy and standards within medium secure services (Department of Health, 2007b, The Royal College of

Psychiatrists, 2007), with its role in the care clustering process for payment by results (Department of Health, 2011) and as a key performance indicator within the CQUINS framework (Department of Health 2008). Thus, from a clinician's perspective, it appears that the role of the HCR-20 is shifting from being a clinical tool for predicting violence and managing risk, to one that also encompasses aspects of outcome measurement.

The theme of this thesis has been to explore the issue of length of stay, for those detained within medium secure services. Whilst the HCR-20 has garnered much interest in terms of research into its scientific validity and reliability over the years (Dernevik, 2008; Douglas, Ogloff & Hart, 2003; Douglas, Ogloff, Nicholls & Grant, 1999; Douglas & Webster, 1999; Macpherson & Kevan, 2004; McKenzie & Curr, 2005; Khiroya, Waever & Maden, 2009; Strand & Belfrage, 2001), rarely does the HCR-20 appear in furthering our understanding of how levels of risk, or the component parts of the framework, may or may not be related to length of stay, as was presented in Chapter 2. In only one of the papers reviewed (Shah, et al., 2011) was the use of the HCR-20 explicitly described in terms of its association with length of stay. With the shift in application of the HCR-20 to processes such as clustering within Payment by Results (Department of Health, 2011), the framework itself may have increasing prominence as a central metric for risk reduction within medium secure services, which has the potential to influence length of stay.

Although this chapter provides an overview of the HCR-20 risk assessment framework, its primary consideration is to present a critique of the scientific and metric properties of the tool.

BACKGROUND TO THE HCR-20 (Version 2)

Historically, clinicians approached the assessment of risk using unstructured, clinical judgments (Heilbrun, Yasuhara & Shah, 2010). These clinical judgement approaches to risk assessment relied solely on a clinician's personal experience, knowledge and clinical ability to assess and predict the risk that a client might pose (Jones & Plowman, 2005). Expertise in risk prediction was therefore the domain of the clinician. However, research has undermined this method, suggesting that as an approach it was fundamentally flawed and did not equate with the argued expertise (Monahan, 1981; Quinsey, Harris, Rice & Cormier, 1998).

In 1968, Goldberg highlighted that levels of training and experience amongst a group of psychologists were unrelated to the level of accuracy in determining risk. Quinsey and Ambtaman (1979) compared the risk prediction abilities of high school teachers with forensic psychiatrists. This study found that the teachers were just as accurate in predicting risk as psychiatrists, with the study also showing that "none of the criteria for expertise in the prediction of violence were met by the forensic psychiatrists" (p.58). Similarly, Monahan and Steadman (1994) reported that clinicians were unable to reach high levels of agreement over risk judgments, despite using the same information. The study concluded, "psychiatrists and psychologists were accurate in no more than one out of three predictions of violent behaviour" (p.5). One of the main criticisms of the clinical judgement approach to risk assessment has been the tendency of clinicians to rate as important symptoms that were not actually predictive of violence. For example, a study by Rice, Harris and Quinsey (1996) asked forensic clinicians to make recommendations about discharge in a group of detained patients.

The outcome of this study indicated that the clinicians were more likely to make positive decisions about those who would have actually been considered to have been the most likely to violently recidivate, whilst continuing to detain the less dangerous. More recently, Heilbrun, Yasuhara and Shah (2010, p.5) go so far as to suggest that "unstructured clinical judgement…is no longer a useful or necessary approach to appraising violence risk."

The actuarial approach to risk assessment developed to counteract the biases inherent in the clinical judgement approach. Actuarial risk assessment approaches involve the use of a highly structured, standardised method with clear scoring processes that provide a probabilistic outcome for the specific behaviour being measured (Grove & Meehl, 1996). Actuarial assessments are usually derived from data from large empirical studies, to which statistical approaches have been applied. For example, meta-analysis and cluster analysis might be used to identify variables that have strong and consistent predictive power in relation to the issue being assessed (Grove, Zald, Lebow, Snitz & Nelson, 2000). In essence, their structures rely on the establishment of strong correlational relationships between factors that have been generated through research. Examples used to predict the risk of re-offending include the Violence Risk Appraisal Guide (VRAG) and the Sex Offence Risk Appraisal Guide (SORAG) (Quinsey, et al., 1998).

A central strength associated with the actuarial approach, when compared with clinically based judgements, is the provision of an explicit coding structure – a structure which, it could be argued, requires minimal expertise and training for administration purposes, and which provides a level of transparency and objectivity to

judgments about risk – something which the clinical approach cannot claim. Another strength attributed to the actuarial approach is that they are sensitive to differing base rates of violence (Quinsey, et al., 1998). Advocates of the actuarial approach have claimed that as assessments of risk, they are superior compared with clinical judgement (Grove, et al., 2000; Mossman, 1994; Swets, Dawes & Monahan, 2000).

Actuarial tools are not without their critics. A key criticism postulates that actuarial approaches are restricted in their applicability, due to the populations upon which they are normed, which at least in the case of the VRAG and SORAG is predominantly a North American maximum security psychiatric population. Quinsey, et al., (1998) argued that the VRAG had good validity for the prediction of violence, irrespective of the population on whom it was being applied, due to cross validation methods applied during the tool's construction. This has since been criticised by Blair, Marcus and Boccaccini (2008) who reported that the correlations for the VRAG and violence in a cross validation sample, were significantly lower than in the original development sample. They further suggested that the correlations were weakened when the cross validation was completed by researchers not involved in the original developmental study.

Criticism has also been levelled at how well such tools may accurately predict risk at the individual, rather than group level (Farrington, Joliffe & Johnston, 2008; Hart, Michie & Cooke, 2007). Hart, et al., (2007) examined margins of error for estimates of risk made with actuarial tools (VRAG and Static 99) and concluded that they had "poor precision" (p. 63). They argued that actuarial assessments were unreliable in estimating violent recidivism at the individual level because the margins of error were

very high. They concluded that such tests should only be used with great caution or not at all, at least when trying to predict risk at the individual level. Further criticisms about the generalisability of actuarial assessments across offender populations have continued, including within the sex offender literature (Craig, Browne, Stringer & Beech, 2004).

Developments in risk assessment attempted to address the concerns that the former approaches to risk prediction posed – thus the structured professional judgement (SPJ) approach evolved. Maintaining its evidence base through adherence to the scientific and professional literatures, SPJ moved away from the limiting, rigid and absolutist process that defined the actuarial approach by not being reliant on reaching an overall predictive score (Webster, Eaves, Douglas & Wintrup, 1995). Rather, it stressed the relevance of 'dynamic' variables, enabling flexibility at the individual, 'clinical' level, from which treatment intervention and management strategies could be more usefully derived at a pragmatic level.

The HCR-20 (Version 1) (Webster, et al., 1995) was developed in Canada and first published in 1995, and later revised in 1997 (Webster, et al., 1997). It was one of the first examples of the SPJ approach, comprising a manualised, structured framework, to guide clinicians when making decisions about a person's risk of future violence, with a view to developing appropriate and relevant treatment interventions and risk management strategies. Effectively, the HCR-20 offered clinicians the best of both worlds - a systematic approach to the assessment of risk, where risk related judgments could still account for the flexibility of a clinician's view, whilst being grounded within the empirical literature (Webster, et al., 1997).

The organisation of the tool is such that it presents the clinician with a "checklist" (Webster, et al., 1997) of characteristics or risk factors that have been empirically linked to the commission of interpersonal violence. Twenty risk variables are spread across three domains – ten historical (static), five clinical (dynamic) and five risk management (dynamic). Already the similarities with the earlier approaches can be seen, in that the tool encompasses historically relevant factors seen within actuarial assessments, such as previous history of violence, age at first violent offence, history of supervision failures and so on. But also present are ten other factors, dynamic in nature, in that they refer to variables which may fluctuate in intensity and relevance according to the individual. These factors may highlight juncture points at which interventions can be targeted to assist in the reduction and management of risk, in a way that the static factors cannot.

Factors considered under the clinical scale include the presence of insight, prevalence of negative attitudes, the presence of active symptoms of mental illness, impulsivity and responsivity to treatment. The clinical scale is coded on the basis of the current presentation of the person under assessment. For example, a clinician would need to consider whether the person being assessed is exhibiting active symptoms of a mental illness such as schizophrenia, which would be congruent with a formal nosological assessment system such as the Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association, 2000) or the International Classification of Diseases (ICD) (World Health Organisation, 2002). This item would therefore link to that of the individual having a history of a mental illness, already considered under the Historical Scale, but reflects the presence of the disorder in the here and now and to what extent it might be associated with a person's risk of violence.

The Risk Scale is the final scale. It is also comprised of five factors. They include the extent to which the person being assessed has made plans which are considered feasible, the presence of a reliable support network and their willingness to access that support, the person's likelihood of being exposed to destabilising factors such as illicit substances and an assessment of the person's likely compliance with on-going attempts at helping them reduce and manage their risks and other related needs. Stress is the final factor considered under this scale, and assessors are guided to consider a wide range of issues and stressors which may undermine treatment gains and increase the likelihood of risk, and the extent to which the person being assessed manages stress. Stressors may include financial concerns, housing and accommodation issues, interpersonal conflict or loss. This item does to some extent highlight the individualistic nature of the dynamic scales, as the assessor can consider those stressors which are most relevant to the client.

Each factor is assessed in terms of its applicability to the individual being assessed, drawing upon interview and collateral information and is then coded as being absent, partially or possibly present, or definitely present. The HCR-20 has no clinical cut-off or composite score, as is seen in many traditional psychometric tests. Instead, clinicians are encouraged to use their discretion to reach a categorical estimate of a person's risk – hence the descriptors of low, medium and high risk are used. This is a fundamental difference between the HCR-20 and many of its actuarial predecessors, and underpins the concept of the SPJ approach. Rather than rely on a composite score to define risk outcome, the onus is upon the clinician completing the HCR-20 to pull together the background information collated prior to the coding of the twenty items within the framework, and to then postulate the context within which a potential

violent act will eventuate. In practice, this is done through the use of risk formulation and scenario planning. The use of risk formulation and scenario planning represent best practice principles in establishing a thorough understanding of a person's risk and the circumstances in which this will increase or reduce. Scenario planning is not an explicit requirement of the HCR-20 (Version 2) and the user manual does not give or recommend a formal structure to be adhered to, although the HCR-20 Violence Risk Management Companion Guide (Douglas, Webster, Hart, Eaves & Ogloff, 2001) was designed as a supplementary guide to assist with the development of risk management plans based on factors identified by the completed framework. Rather, risk scenario planning has developed subsequently and is inherent in other SPJ risk assessments such as the Risk for Sexual Violence Protocol (RSVP) (Hart, Kropp, Laws, Klaver, Logan & Watt, 2003). Scenario planning involves the description of the most likely situations in which a given risk behaviour will occur, taking into consideration the nature of the risk behaviour, the severity of the behaviour, the imminence and frequency of the behaviour, and usually involve a narrative or vignette in which the risk behaviour can be described as most likely to occur. Risk management strategies are then developed to provide a comprehensive support structure to reduce the likelihood of the predicted behaviour occurring.

The HCR-20 was developed using logical (or rational) item selection (Douglas & Reeves, 2010). Webster, et al. (1997) reviewed the scientific and professional literature relating to violence, and from this selected a series of risk factors that were evident across this literature base. Douglas and Reeves (2010) contrast this to the typical development of other assessments, including actuarial tools, which rely upon empirical item selection. Empirical item selection is founded on the use of items

where a “statistical association between risk factors and outcome” (Douglas, Skeem & Nicholson, 2011, p.328) has been demonstrated. This method therefore returns only those factors that predict the issue under scrutiny (in this case violence), and they are weighted accordingly. This is a central tenet of the actuarial approach.

Douglas, Skeem and Nicholson (2011) propose that empirical item selection limits the applicability of tools based on this premise, to the sample that it was developed upon. They argue that the empirical approach has the potential to omit factors that were either not representative within the sample or occurring too infrequently to be statistically valid, and it is because of these issues that such tests cannot be generalised to other populations. Using rational item selection, their “goal was to avoid the omission of important risk factors, or inclusion of unimportant factors on the basis of chance associations” (Douglas & Reeves, 2010, p.155). They argued that because the rational item selection approach is not “sample dependent” (p.156), the HCR-20 tool is therefore generalisable to alternate populations, giving support to its use with both male and female populations, in a way that something like the VRAG was not. The manual for the HCR-20 is very clear that its use “ should be restricted mainly to settings in which there is a high proportion of persons with histories of violence and a strong suggestion of mental illness or personality disorder” (Webster, et al., 1997, p.5). Thus, the authors set out their argument for its use across a series of relevant settings, including prisons, psychiatric hospitals and probation populations.

However, the rational item selection is not without its limitations, some of which are consistent with the criticisms that have historically been levelled at the unstructured, clinical judgment approach. As it does not rely on statistical weighting, it can be

argued that tools founded on rational item selection are less predictive than those built on empirical item selection, and fare even less well when subject to the similar cross validation tests (Douglas & Reeves, 2010). This criticism is based upon the 'human involvement' element to the methodological approach taken, and the introduction of bias towards or against different items at the point of selection, even when the theoretical and literature base being considered remains consistent. Douglas, et al., (2011) stress the importance that where this approach is used, the sources of information used must be well stated, something which the HCR-20 authors have been consistent in doing.

PROPERTIES OF PSYCHOLOGICAL TESTS

A good psychological test has been described as one that aims to measure the same 'thing' and have the "same meaning over time and across situations" (Field, 2009, p.10). Kline (1986) proposed that a good test has three key features: reliability, validity, and appropriate norms. This provides criteria by which it is possible to assess the degree to which the HCR-20 conforms to the properties of a good psychometric test and from which conclusions regarding clinical utility can be drawn.

Reliability

Reliability refers to the consistency of a test or measure (Kline, 1998). Three forms of reliability are considered, each essential to good test construction – these being test-retest reliability, internal consistency and inter-rater reliability.

Kline (1998, p.29) described test-retest reliability as an "essential attribute for any good measure, whether a psychometric or not". Essentially, it suggests that if a test were to be applied to just one person, several times over without any changes to the person, it would be expected that the scores would be generally consistent across all the tests**.** Internal consistency, usually measured as Cronbach's Alpha, refers to the consistency between the components of a test, and that they essentially measure the same thing (Kline 1998). Inter-rater reliability concerns itself with the variation that exists between assessors using the same tool or test. Reliable tests are considered to have a minimum coefficient of .70, indicating that different raters agree on the application of the test to the same cases (Kline, 1998).

Douglas and Reeves (2010) suggest that inter-rater reliability is the most important aspect of reliability when evaluating the HCR-20. They argue that because the HCR-20 is "not a measure of a psychological construct", the items do not "hang together" to form a construct (p.162), making internal consistency less of an issue in the way it would be for something like the Hare Psychopathy Checklist – Revised (PCL-R) (Hare, 1991). This is not to say that evaluations of internal consistency have not been completed. Studies evaluating this have produced Cronbach's Alphas at .95 for the full HCR-20 assessment (Belfrage, 1998), with variations across the three subscales ranging from .85 to .96 (Dunbar, Quinones & Crevecoeur, 2005). An interesting point of note here is that these different studies evaluated the original HCR-20 (Belfrage, 1998) and its successor, HCR-20 V2 (Dunbar, et al., 2005), with internal consistency being demonstrated across both tools.

Studies looking at inter-rater reliability are more widely available and again comprise evaluations of the HCR-20 total scores and also of the individual subscales. Douglas and Reeves (2010) emphasise that of the studies that have looked at inter-rater reliability of the total score for the HCR-20, most report coefficients of .80 or more. For example, this is a finding supported by the meta-analysis of the HCR-20 by Nikolova et al., (2006), which considered over 6000 cases within their study across forensic psychiatric, civil psychiatric and correctional settings, and found that the HCR-20 retained its psychometric robustness. Furthermore, Douglas and Reeves (2010) illustrate the generalisability of the assessment across the genders, citing the evaluation completed by de Vogel and de Ruiter (2005), which found coefficients above .70 for both male and female populations.

Reviewing evaluations of inter-rater reliability of the three separate subscales, a wider range of outcomes is seen. Those looking at the Historical scale have been found to produce a coefficient range of between .58 and .97, but with a median value of .86 (Douglas & Reeves, 2010). Lower coefficients are typically seen with the Clinical scale, ranging from .55 to .95, with a median of .74 (Douglas & Reeves, 2010). The Risk scale median has been reported as being even lower at .68, with the lower limit being .47 and an upper limit of .98 (Douglas & Reeves, 2010). The reliability coefficients for the subscales potentially suggest that as stand-alone scales, the Historical scale is more robust than the two scales measuring dynamic variables (Clinical and Risk). Given the emphasis on static variables within the Historical scale this is perhaps unsurprising, particularly if we consider the earlier, well publicised strengths associated with actuarial assessments compared with more subjective based assessments. Douglas and Reeves (2010) postulate that it is because of the subjective nature of the Clinical and Risk scale ratings, that the variations across the coefficients are seen. Equally, they offer the criticism that many of the evaluated studies relied primarily on file-based reviews, in which historical items may be more easily identified than dynamic ones, resulting in greater accuracy for the Historical scale. One might also consider that the assessment of the Clinical and Risk scales requires a greater level of clinical skill from a practitioner in understanding the client's presentation and its relevance to risk, something which can rarely be gleaned from a file review.

In summary, it could be argued that in terms of reliability, the overall pattern of outcomes from the evaluations that have been undertaken appear to demonstrate that the HCR-20 meets the first requirement of a good psychological test.

Validity

Validity refers to the extent to which a test measures the 'subject' that it was designed to measure and that inferences drawn from results must be "appropriate, meaningful and useful" (Gronlund, 1998, p.226). As with reliability, validity can also be divided into several types.

Criterion related validity is concerned with the usefulness of a measure in terms of predicting the criteria being assessed (Kline, 1998). Therefore, in the case of the HCR-20, an assessor would be interested in the extent to which a final risk judgment would predict the outcome of the risk behaviour occurring. There are two recognised forms of criterion validity, these being predictive and concurrent. Concurrent validity refers to the correlation of one test with a "similar test taken at the same time" (Kline, 1998, p.35). An example of this would be a comparison of the HCR-20 outcomes with another violence risk assessment tool, such as the VRAG. Predictive validity is concerned with the success with which an independent measure predicts an event. Thus, in the case of the HCR-20 the event that is being predicted would be an incident of interpersonal violence.

A retrospective study by Douglas and Webster (1999) examined the concurrent validity of the HCR-20 (Version 1), compared with the VRAG and Hare Psychopathy Checklist – Revised (PCL-R) (Hare, 1991) with a group of offenders detained in maximum security. Pearson coefficients were established. The result suggested only a moderate degree of commonality across the three assessment tools. Douglas and Webster (1999) noted that that the correlation coefficient for the Historical scale

was .62 with the VRAG, and .50. for the PCL-R, with the Clinical scale correlating even less strongly. At face value, these results suggest that the concurrent validity for the HCR 20 is not strong, if we are looking for a minimum coefficient of 0.70 (Kline 1998).

At the same time, Douglas, et al., (1999), looked at the predictive validity for the HCR-20 and Hare Psychopathy Checklist - Screening Version (PCL:SV) (Hart, Cox, & Hare, 1995). Douglas, et al., (1999) found that a Receiver Operating Characteristics (ROC) analysis suggested good predictive validity for the HCR-20 when compared to the PCL:SV. The results produced Areas Under The Curve (AUC) of 0.76 to 0.80. Further regression analysis suggested that the HCR-20 increased the "incremental validity of the PCL:SV" (p.925), and that it was "only the subscales of the HCR-20 that predicted future violence" (p.917). Hosmer and Lemeshow (2000) proposed some rules for the interpretation of AUC values. They proposed that AUC outcomes of 0.5 should be considered as no better than chance, AUCs of 0.7 to < 0.8 as being acceptable, AUCs of 0.8 to < 0.9 as being excellent, and an AUC of 1.0 as indicating perfect discriminating power. According to the guidelines of Hosmer and Lemeshow (2000), the AUCs for the predictive validity of the HCR-20 in the Douglas, et al. (1999) study would be considered 'acceptable', rather than 'good'.

Since then, others have continued to investigate both concurrent and predictive validity for the HCR-20. Concurrent validity was considered in a study by Gray, et al, (2003) in which they compared the HCR-20 (Version 1) and the PCL-R for a sample of mentally disordered offenders detained in conditions of medium security. They reported that both the Historical and Clinical scales were independently correlated at

the level of statistical significance with the PCL-R (r=.68, p<.01, H scale and r=.46, p<.01, C scale), and when combined (r=.69, p<.01). One of the drawbacks that they allude to in their study with regard to the correlation of the Historical subscale and the PCL-R, is that the PCL-R is one of the independent items within the H subscale. This introduces the problem of double counting, lending more weight to the outcome than is necessarily justified, a criticism also postulated by Witt (2000). As a result of this close relationship between the two assessments, Gray, et al., (2003) conclude that the outcome is to be expected.

The study by Gray, et al., (2003) also examined the predictive validity of the HCR-20 and the PCL-R using the same original data set. Using ROC analysis, the results produced AUC associations of between .79 and .83, from which they concluded that the HCR-20 was "strongly predictive of all forms of outward aggression" (p.448). They also concluded that the separate Historical and Clinical subscales themselves represented strong predictive qualities, with AUC associations of .74 for the Clinical scale and .73 for the Historical Scale. It could be considered that Gray, et al., (2003) are somewhat overstating their success in investigating predictive validity, as these results would fall within the descriptive category of acceptable discrimination (Hosmer and Lemeshow, 2000).

A meta-analysis of the efficacy of violence prediction completed by Yang, Wong and Coid (2010) showed that across nine risk assessment tools, the HCR-20 continued to demonstrate good predictive ability. They also found that the individual scales also demonstrated predictive efficacy when compared with other risk assessment tools, but did best when combined as other studies have shown. However, as with the Gray, et

al., (2003) study, and the criticism outlined by Witt (2000), the authors of this study also raise the issue that the PCL-R is one of the twenty items within its framework, leading to doubts about its ability to predict violence over and above that of the PCL-R. They set their findings about the HCR-20 within a context that all nine risk assessment instruments evaluated were able to predict violence at levels greater than chance, and that none of the tools were significantly better than the others.

A later similar study by Gray, Taylor and Snowden (2008), again using a mentally disordered population within medium security, found strong and moderate predictive validity for the Historical and Risk subscales respectively, but contrastingly poorer outcomes for the Clinical scale. ROC analysis was again the statistical methodology adopted. This study differs from the earlier study by Gray, et al., (2003) by incorporating an evaluation of the Risk Scale as an independent scale. Their results indicated AUC associations of .70 to .76 for the combined HCR-20 scales, with similar results seen for the Historical scale (AUC =.68 to .77). AUC associations of .63 to .69 were found for the Risk scale. However, compared with the earlier 2003 study, the AUC associations for the Clinical scale are much lower (AUC .54 to .61), contradicting their earlier conclusions regarding the predictive ability of the independent Clinical scale.

Grann, Belfrage and Tengstrom (2000) also evaluated the predictive validity of the Historical subscale and the VRAG, but with specific reference to personality disordered individuals. Their study found that whilst both tools had good predictive validity within a personality disordered population when compared with individuals with a diagnosis of mental illness, the Historical scale performed better than the

VRAG (for the combined groups, Historical scale AUC =.71; VRAG AUC =.68), although the difference between the two assessment tools did not reach statistical significance (p=.1505). The limitation that this study only used the Historical subscale should be noted – although the authors were only concerned with the 'actuarial' aspect of the risk prediction process. Again, it is noted that according to Hosmer and Lemeshow (2000), the AUCs for the HCR-20 from this study would be described as only 'acceptable.'

One area where the predictive validity of the Clinical scale does appear, perhaps, to have more clinical utility however is in the prediction of institutional violence. A study by McKenzie and Curr (2005) looked at the predictive validity of the HCR-20 within a medium secure psychiatric setting. They found that whilst the Historical scale was not independently predictive of institutional violence (AUC= .546), the Clinical scale was "moderately predictive" (AUC= .678) (p.26). Similar findings were seen in the study by MacPherson and Kevan (2004), but this time with a maximum-security population. It could be argued that the outcomes from such studies are useful for assisting clinicians to develop risk management strategies with regard to those clients who may be inclined to institutional violence early on in their admissions.

Content Validity is regarded as a non-statistical validity (Anastasi & Urbina, 1997). Anastasi and Urbina (1997) suggest that tests and measures have content validity because test items are selected specifically because they relate to the literature base for the subject being measured. This directly relates to the construction of the HCR-20 using rational item selection. One study, which addresses the relevance of content validity, is that of Dernevik (1998). It concluded that "content validity for the full

scheme" (p.135), was implied, following establishment of .67 coefficiency using Spearman's Rho between nineteen of the HCR-20 items. They omitted item H1: Previous Violence, highlighting that there was already a strong empirical base for the relationship between future violence and prior offending history. Their findings suggest that the items selected for the HCR-20 relate well to the issue of violence risk, which it sets out to measure. This may be further supported given that each of the factors comprising the HCR-20 have their own established empirical base and examples are provided within the assessment manual against each factor.

Within the literature, emphasis appears strong in terms of criterion validity for the HCR-20, especially predictive validity. Searches within the literature for references to construct validity for the HCR-20 have revealed fewer studies, except where studies have looked at its convergence with other assessments such as the PCL-R and VRAG. Despite the absence of evaluations into some aspects of validity, there is some evidence supporting predictive validity at acceptable levels (Hosmer & Lemeshow, 2000). Again, one might suggest that this is not unusual given that, by its very nature, the HCR-20 is designed to assist with the prediction of risk, and the importance of predictive validity has been singularly emphasised (Bonta, 2002). This potentially suggests a weakness of the HCR-20, when considering other aspects of validity.

Face validity has been described as a "public relations concept" (Barker, Pistrang & Elliott, 2002, p.65). It is concerned with the appearance of the test and that to a person undertaking the test, it appears at 'face value' to measure what it is claiming to measure (Gudjonsson & Haward, 1998). In respect of the HCR-20, it could be concluded that the test would appear to assess aspects of violence and that the factors

within the manual are well defined in terms of the applicability to the risk literature.

The dominant methodological analysis for establishing predictive validity is ROC analysis. Using the guidelines suggested by Hosmer and Lemeshow (2000), all the studies considered would indicate acceptable levels of discrimination for predictive power as AUCs ranged between .70 and .83 for the combined totals of the three scales. More variation is seen when the three scales are considered separately, with the Historical scale appearing to be the only one that consistently retains an acceptable level of discriminatory power. The Risk scale has been investigated by only one study and that was found to have low discriminatory power. The Clinical scale has demonstrated mixed results when investigated using ROC analysis, with some AUCs being reported at .74 and others within a range of .54 to 61, which would suggest that this predictive power of the clinical scale alone is little better than chance. Given the evidence, the argument could be made for the composite HCR-20 meeting the second requirement of a good psychological test. The relative predictive validity when the individual scales are examined separately is not as strong.

Appropriate Norms

A potential flaw of the HCR-20 is the absence of reported norms within the technical manual. Norms provide a basis against which test data is compared, and offer an insight into the base rate of a behaviour occurring within a given population. Kline (1993) asserts that without norms, "the meaning of any test is difficult to gauge" (p.49). Therefore, unlike the findings that support the HCR-20 as adhering to the rules of what makes a good psychological test on the basis of validity and reliability, it could be proposed that where the consideration of norms are concerned, the HCR-20

is somewhat lacking.

It is perhaps surprising that given the extensive research that has been undertaken into the HCR-20 as a violence risk assessment framework since its conception in 1995, that no normative data has been published. Without norms, a question arises as to what degree the HCR-20 can claim to be an appropriately standardised assessment. A conclusion that might be drawn from this is that the assessment is not specific or standardised enough to be able to discriminate between different groups in the prediction of violence risk. Despite the lack of reported norms within the user manual, the growing empirical base has begun to provide a wealth of information about its use within various population groups, from the original studies in the North American forensic population (Belfrage, 1998) to recent research in Europe, addressing the validity and reliability of the HCR-20 across difference psychiatric settings (Grann, et. al., 2000; Gray, et. al., 2008; MacPherson & Kevan, 2004), ethnic groups (Snowden, Gray & Taylor, 2010) and genders (de Vogel and de Ruiter, 2005), many of which have continued to demonstrate consistency in its predictive validity and reliability as a measure of risk.

Another explanation for the absence of reported norms is its underlying idiographic approach, which means that outcomes are not compared with a normative sample (Barker, et al., 2002), but instead highlights the individualistic nature of the assessment and reinforces the authors' statements that the assessment was not designed as a formal measurement instrument. This adheres with Kline's (1993) assertion that when individual differences are the issue of measurement, standardisation becomes less important. It would be difficult for example, to consider

to what extent the outcome of an HCR-20 assessment on a mentally disordered offender could be compared with a non-offender sample, as the HCR-20 would be invalidated if used on a sample of people who had not committed a violent offence. Nonetheless, with the advent of the HCR:V3 (Douglas, Hart, Webster, Belfrage & Eaves, 2011), it is possible that there will be increased opportunities for the publication of appropriate norms given the popularity of the tool within psychiatric and correctional settings.

LIMITATIONS OF THE HCR-20 (Version 2)

As has been described, the empirical literature offers support for the argument that the HCR-20 meets two, if not the three main fundamental attributes of a good psychological tool. However, criticisms about the assessment framework still arise and these perhaps reflect issues of clinical utility rather than its metric properties, but which could be argued still reflect the reliability and validity of the assessment overall.

One criticism that could be levelled at the HCR-20 is the lack of a composite score following completion of an assessment. Using the HCR-20 as the example and extrapolating from classical test theory (Kline, 1998), it is logical to suggest that the higher a composite score, the higher the level of risk posed by an individual. This would fit with the patterns seen with other assessments that measure one domain, such as depression and anger. The higher the composite scores on assessments, the higher the level of the attribute being measured is said to be present. From this, a clinician can infer the level of the clinical problem and address this as required. However, the difficulty with composite scores, as seen with assessments such as the VRAG, is that it tells the clinician nothing in terms of what kind of violence might be repeated, at what frequency and what severity. Therefore, the HCR-20, by not claiming a composite score, can to some extent avoid this type of 'judgement'. That is not to say that without a total score, the HCR-20 performs this role any better. Indeed, from a clinician's point of view it does not do this when scrutinising the factor level scores at face value. Take for example the first of the Historical items – Previous Violence. The scoring rules for this item can render a score of 2 for a person with one very severe example of violent conduct, but also the same score for a person with

numerous, but potentially less severe instances of the same behaviour. The same issue could therefore be applied to all the factors being assessed. Following this logic, it would be possible for both individuals to obtain the same overall composite score by the end of the assessment, but this would tell you little about their comparable risk.

Douglas (2008) cites 'over-breadth' of item content as a limitation of the HCR-20. From a clinical utility perspective, this links to the arguments made about the use of a composite score. The HCR-20 factors can appear broad in their definitions. Yet, because of the idiosyncratic nature of violence and risk, depending on which factors a clinician weighs as being relevant to the individual case, those being assessed can again, theoretically, be of similar risk levels, yet have significantly different dynamic risk profiles. An example of this relates to the item addressing childhood maladjustment. This item is based upon the theoretical premise that exposure to violent experiences in childhood, whether as victim or victimiser, increases a person's risk of committing violence. There is a similar inherent difficulty with the scoring here too, in that a person who has been a victim can score the same as someone who has been a victimiser, which would at face value appear to be inconsistent with the concept of violence. However, clinicians are required to provide justifications for their scoring, and therefore account for such variation. The issue of scoring at the item level (and then assuming an eventual composite score) is problematic, because fundamentally people are 'risky' for different reasons.

It could be argued that at the clinical level, reliance on scoring protocols threatens appropriate treatment planning, because logic would suggest that as both individuals in the example cited above scored the same, they would therefore require the same

level and nature of treatment. Investigation of the qualitative element of the history would however, provide evidence to the contrary. It could be argued therefore, that in terms of clinical utility, this is a strength of the HCR-20. By not capitulating to the classical requirement of requiring a composite score, it shows the strength of the structured judgement approach in its valuing of the dynamic nature of individuals and their risk, as well as just the actuarial. Douglas, Ogloff and Hart (2003, p.1372) postulated that "clinical judgment, if made within a structured context, can contribute in meaningful ways to the assessment of violence risk". What it does not take account of however, is that it opens the assessment process to greater bias and misinterpretation because of the subjective nature of the information that may be appraised in reaching a score.

Another criticism of the HCR-20 lies in its construction using rational item selection and the absence of additional factors that are relevant to the commission of interpersonal violence and aggression. It is possible, and very likely probable, that whilst the current structure of the HCR-20 was a good representation of the empirical basis at the time of its construction in 1997, it is less so now fifteen years on. This is a weakness that has been postulated by the authors of the HCR-20 and used as one of the justifications for a revision of the assessment, which has led to decision to publish the HCR-20: Version 3 (Douglas, et. al., 2011). An example of this is the absence of factors relating to psychological and emotional functioning, such as anger, which can have a moderating influence for the commission of violence. Bonta (2002) argued that issues of emotional instability are "largely irrelevant" for risk assessment (p.361). However, there is a significant literature base for the relationship between emotional states and violence towards others (for example, Baumeister & Bushman, 2007;

Steffgen & Gollwitzer, 2007) and between anger and violence specifically (Dodge & Coid, 1987; Novaco, 1994). Within the sex offender literature, emotional instability in the form of emotional loneliness, anxiety, depression and anger have all been linked to the perpetration of sexual violence (Proulx, McKibben & Lusignan, 1996). Both aspects of emotional regulation are relevant to the HCR-20 given its strict definition regarding interpersonal violence, which includes the commission of sexual offences. These examples alone suggest that the HCR-20 has a weak point from a theoretical basis.

Whilst the HCR-20 (Version 2) appears to have an acceptable level of validity and reliability in terms of its psychometric properties, the omission of other salient factors remains a weakness of the framework and has the potential to impact on the formulation of a person's risk. For example, just because a factor may not have a strong relationship with violence in a wider sample group, does not mean the relationship is not significant for those where it is highly pertinent. The authors of the HCR-20 stipulate that clinicians undertaking any assessment using the framework should have "expertise in conduct of individual assessments" and "expertise in the study of violence" (Webster, et. al., 1997, p.17), in which they indicate that assessors need to be "familiar with the professional and research literature on the nature, causes and management of violence". Thus the onus is on the clinician to ensure that they are aware of the relevant research base and that they incorporate any additional aspects of a person's functioning or the nature of previous offending into any decisions that are made at the point of risk formulation and final risk conclusions. It could be argued then, that where clinician's are not sufficiently aware of the available literature and rely solely on the twenty factors within the HCR-20 framework, this has the potential

to undermine the validity and reliability of the tool overall, because the framework is by its very nature, leaves too much room for idiosyncrasies on the part of the assesse, and error on the part of the assessor.

CONCLUSIONS

The HCR-20 may at face value, appear an unusual choice for a critique regarding psychometric testing. It does not market itself as a psychometric tool, along the same lines as, for example, the PCL-R. Rather, it is a theoretical and empirically supported framework, with psychometric properties of validity and reliability embedded throughout.

This chapter focused on a critical review of the HCR-20 Violence Risk Assessment, through an appraisal of the evidence base available addressing its reliability and validity, especially when compared with its forerunners and contemporary competitors. It also attempts to outline a number of limitations of the tool, not least in terms of its clinical utility, through consideration of the contextual application of the HCR-20 and the use of risk formulation.

The work of Kline (1998) has been influential in developing the standards by which modern psychological tests are benchmarked. Using this as a basis, it could be argued that the HCR-20 meets the core conditions to be sufficiently robust as to be called a psychometric. It offers a standardised approach in that all assessors follow a "uniform procedure in terms of administration and scoring" (Gudjonnson & Haward 1998, p.83) leading to "an evaluation within a specified domain" (APA, 1999, p.3). In its earliest stages, the HCR-20 was hailed as showing "promise" (Borum, 1996) with its empirical focus and structured coding system. However, it is not without its limitations, not least because the field of risk assessment and understanding harmful behaviours has developed significantly since the HCR-20 was first published in 1995.

Nonetheless, it remains the predominant framework upon which fundamental, life-changing decisions are based, and upon which clinicians rely in their work with forensic populations (Douglas & Reeves, 2010; Webster, et al., 1997). Without the weight of the empirical literature to provide support, this would be harder to justify.

CHAPTER FIVE

DISCUSSION

The purpose of this thesis was to examine the factors associated with length of stay for male, mentally disordered offenders who are detained within conditions of medium security within the forensic psychiatric estate. The relevance of length of stay to applied clinical practice has begun to garner increasing significance as a result of the financial challenges posed to the NHS by the current state of the national fiscal position, and the heightened pressures associated with maintaining effective clinical practice at a time of increasing financial challenges.

The investigation into length of stay commenced with a review of the current literature regarding length of stay within the medium secure psychiatric estate, presented in Chapter 2. The review considered the extent to which the existing research base was able to explain which factors were pertinent to an understanding of the length of time mentally disordered offenders are detained in hospital, and the extent to which these factors are consistent across populations. The review found that in comparison to the general adult psychiatry literature, length of stay for medium secure forensic populations has attracted limited interest in recent years. Within the literature that does exist, studies have, until more recently (Shah, 2011), tended to focus on describing patterns of admission and changing trends (Kennedy, et al. 1995; Ricketts, et al., 2001), rather than seeking to critically explain or investigate relationships between clinical and forensic variables that might improve understanding of the factors that influence length of stay for certain patients. There was divergence across the papers in terms of their findings when considering social demographic, clinical and forensic variables for each of their samples. The main area of convergence across papers was in relation to factors that demonstrated no association with the length of hospitalisation. This suggests that when it comes to

understanding length of stay, it may not be useful to consider standard variables in isolation from each other, but that issues of chronicity of illness and comorbidity with risk need to be considered in combination. It indicates that despite previous attempts to understand length of stay through describing and monitoring admission patterns and trends, length of stay research could be more meaningful if it focussed on investigating other variables that are inherent to a person's experience of care. For instance, the extent to which clients meaningfully engage in therapeutic treatments to reduce their risk, and the extent to which clinical teams then test change at the individual level, would be salient factors for investigation. A limitation of the review was the small number of papers considered for review, meaning that it was difficult to draw any firm conclusions about the reliability of the factors investigated and their relationship to length of stay for the wider mentally disordered population. The inconsistency of findings within the papers reviewed perhaps offer the best evidence for additional investigation of length of stay within mentally disordered offenders detained within the medium secure estate.

Following the literature review, an empirical study was presented in Chapter 3. This study sought to add to the current knowledge base through further examination of a set of variables and their relationship to length of stay specifically for a male, mentally disordered population detained in a medium secure psychiatric hospital over a ten-year period. Variables investigated in this study included socio-demographic factors, such as ethnicity and age on admission, clinical variables including diagnosis and legal status, as well as psychiatric history, and finally forensic variables such as offending history and the nature of index offence, many of which had been found within the studies reviewed in Chapter 2. However, unlike many of the other studies

reviewed (with the exception of Shah, et al., (2011)) this study also incorporated a standardised assessment framework, that being the HCR-20 (Webster, et al., 1997), which provided additional clinical and forensic information, such as the relevance of a history of substance use, and ratings of clinically relevant items such as insight, treatment compliance and responsivity. Similar to the five studies presented in the previous chapter, the study found relationships between some clinical and forensic factors, such as diagnosis and legal status, but not others, for example ethnicity, or nature of the index offence or the factors as measured by the HCR-20.

The key finding of the study was that a length of stay of two years or more was predicted by having a diagnosis of a schizophrenic disorder, a finding consistent with the earlier studies reviewed (Shah, et al., 2011). Limitations associated with the study involved the use of a retrospective cohort study design, and the implications of the reliance on historical data, which may have contained inaccuracies and inconsistencies. Where possible, efforts were employed to ensure that the study would be sufficiently robust for the conclusions to be meaningful, such as consideration of the sample size for a logistic regression analysis as set out by Peduzzi, et al, (1996). However, it is also noted that the threshold for cases within each dependent variable group did not quite adhere to this recommendation, and post hoc power analysis also pointed to the possibility of the presence of Type Two errors for some of the statistical analyses computed, which are potential limitations of this study, meaning that all conclusions drawn about the reliability of the findings needed to be tentative.

In chapter four, a critique of the HCR-20 (Version 2) (Webster, et al, 1997) is presented. The chapter presents an overview of the development of the tool, but

predominantly considers the available literature on validity and reliability, and considers these in reference to its use in assessing long-term interpersonal violence from a psychometric perspective. A widely adopted, standardised assessment framework used within mentally disordered services (Douglas & Reeves, 2010); Webster, et al., 1997), it has in recent years started to increasingly become a focus for outcome measurement within forensic services (Department of Health, 2008). The extent to which its metric qualities are reliable and valid has therefore, the potential to be of increasing importance if those working at the clinical level become required to use it to demonstrate efficacy of treatment and for it to play a more central role in determining the care pathway for patients, such as through the payments by results scheme (Department of Health, 2011). It was adopted as the only standardised measure within the empirical study presented in Chapter 3, to investigate the extent to which any of the twenty items might contribute to the length of stay for the study sample. The results found no relationship between nineteen of the items and length of stay. Only one item was found to have a statistically significant relationship with length of stay – that of 'Negative Attitudes.' However, the relationship suggests that patients who are rated highly on this item are more likely to experience a short period of hospitalisation and the reasons for this are considered within chapter four. Nonetheless, this raises an interesting dilemma about the merit of using the HCR-20 as an outcome measure within mentally disordered populations beyond its original aim, that of being a framework to assess the long term risk of violent recidivism.

Despite the inconsistencies across the limited research that does exist, length of stay remains an area of growing importance and significance in forensic psychology and psychiatry. This is not least because of the pressures of managing resources whilst

still providing high quality care and treatment for a highly complex client group. It is hoped that this thesis contributes to the current knowledge base regarding length of stay, firstly through the presentation of a review of the existing literature, and secondly through the research study, with its focus on investigating the extent to which a series of variables can predict length of stay. The inconsistencies found within the literature and which to some extent remain unresolved through the empirical study in Chapter 3, suggest that further work is required before clinicians and researchers can assert with confidence a clear understanding of the relevance of any variables that explain the variability of length of stay. This raises a number of potential avenues for further research. It may be that a national approach to investigating length of stay is required, where NHS and independent providers that manage mentally disordered offenders in conditions of medium security combine data for analysis. An advantage would be that a large sample would allow for generalisation, and would to some extent control for local, geographical differences. As suggested earlier however, a multi-site study may not provide any clearer answers to the question of what factors predict length of stay, unless a set of factors which might be considered to be relevant to length of stay can be agreed. Other options include the investigation of variables associated with length of stay within the female mentally disordered population, which as yet remains an under-researched sample.

There is another potential area for investigation: that of the qualitative aspects of a person's detention in hospital, the intangibles that often define a person's engagement in their recovery, as opposed to the legal and clinical labels that we apply. So far this has been an under-represented area within the research, with the exception of the study by Castro, et al (2002), which did consider the relevance of therapeutic

engagement to length of stay. This suggests that there is still a significant void in the literature and consequently our understanding of the importance and relevance of all factors that are central to a person's detention in hospital.

The outcome of any future research could have significant implications at both clinical practice and policy levels. Secure services are adopting defined care clustering processes and the Payment by Results scheme (Department of Health, 2011). Therefore, clinicians will be required to demonstrate that the treatment offered for specifically defined client groups will be effective in terms of clinical utility, but also cost efficient and timely in their delivery. This brings us back to the standard of an eighteen-month to two-year timescale for those detained within medium secure hospitals (Glancey, 1974; Reed, 1997). The rubric of the two year period appears to be gaining momentum at a practical level, yet the research presented in Chapter 2 has shown that this is, for the time being, an unreliable token by which to measure effective clinical and risk recovery. At best, we know that a diagnosis of schizophrenia, combined with a restricted hospital order is likely to result in a hospital admission of two or more years (See Chapter 3; Shah, et al., 2011). We also know that there are external factors that impact on length of stay, such as the influence of Ministry of Justice appraisals of risk and readiness for discharge, and the extent to which commissioners, the Ministry of Justice, and, to some extent, clinicians can be risk averse (Centre for Mental Health, 2011). Each of these, no matter whether unintentional or not, will influence length of stay for some individuals, because of the potential costs of 'getting it wrong' (Moore, 1996).

Standardisation of outcomes that measure clinical and risk recovery across all the tiers

of secure care remains limited (Centre for Mental Health, 2011), but is nonetheless a central feature of how our knowledge and practice can be improved. Measures such as the HCR-20 and the Health of the Nation Outcome Scales - Secure (HoNOS - Secure) (Sugarman & Walker, 2007) are increasingly evident, but there are alternative assessments and frameworks emerging and in existence, that may provide equally, if not better information on the progress of clients, such as those seen within the general psychiatry field (Thornicroft & Tasnsella, 2010).

Within secure services, the shared pathway outcomes framework (National Secure Services Quality, Innovation, Productivity and Prevention Working Group, 2012), introduced as a CQUIN in 2012, addresses a wide range of individual factors including mental and physical health, risk, substance misuse, relationships and leisure time. This may offer a system for explicit goal setting and attainment, which can be measured against timescales and throughput. The proposed Payment by Results framework and the identification of five forensic care pathways may also offer an option by which length of stay can be better understood. By allocating patients to one of the five forensic care pathways or diagnosis-related care clusters, clinicians and commissioners may gain a better understanding of the clinical and risk profiles of the service users for whom that group is comprised. It may assist in helping us to identify specific bio-psychosocial and risk treatment needs and the appropriate treatment pathways to ensure those needs are met. It may also help with the geographical profiling of morbidity associated with clinical and risk factors. As such, it is possible that there may result some standardisation in the care and treatment for individuals within each of these forensic pathways, and an understanding of the length of time this will take and an acceptance of the differences between the patient groups. This

would to some extent, alleviate the pressure of a global length of stay target for such a heterogeneous population, and focus attention on the timescales required for successful and meaningful recovery for the identified groups, for which an appropriate cost can be calculated.

It is also important to consider whether an alternative marker of clinical effectiveness might not be rates of recall and readmission to hospital or recidivism, rather than focussing on the length of time it takes to reach discharge. Recall and readmission do not appear to have the same negative inferences associated with them, as does a lengthy hospital admission. From a clinician's perspective, readmission can at face value sometimes be seen as evidence of unsuccessful treatment. However, readmission under the auspices of respite, or to prevent harm to self or others, even for short periods of time could be reframed as examples of good practice, in terms of meeting the twin requirements of good clinical care and safe and effective practice when managing clients who pose a risk to wider society. It could be suggested that it is usually because of the longitudinal approach to care, so typical of forensic mental health services, that readmissions and reoffending amongst mentally disordered offenders living in the community are as infrequent as they are. It may also be the case that it is preferable from a commissioning perspective to support frequent, but short-term admissions, rather than lengthy admissions which can at times run into a significant number of years, at higher cost.

As a practising clinician, it is also important to consider that what works for one person may not work for the next and that recovery at the clinical and risk level remains a very personal issue (Drennan & Alred, 2012). For some patients, the central

issue of recovery is not how quickly they can progress, but the importance of getting and remaining well and ultimately staying out of hospital. For some clients this can take years. The essential aspect for them, and which the author would argue also for clinicians, is the prospect of patients maintaining a meaningful life in the community, a goal that is central to the themes of risk and clinical recovery (Barker, 2012). In this regard, clinicians, services and commissioners need to feel and be more assured of exactly which elements contribute to the length of stay for a person in medium secure conditions, if the upper limit of two years is to be anything other than an arbitrary target. Until this is the case, there is potentially an argument for reducing the emphasis that length of stay currently holds, and looking for alternative ways of thinking about the cost of mental health treatment with this client group and the way in which we measure effectiveness of our services.

REFERENCES

Abbott, P., Davenport. S., Davies. S., Nimmagadda, S. R., O'Halloran, A. & Tattan, T (2005). Potential effects of retraction of the high secure hospitals. *The Psychiatrist, 29*, 402-406.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)* (4th ed.). Washington DC: American Psychiatric Association

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Anastasi, A., & Urbina, S. (2007). *Psychological testing* (7th ed.). New Jersey: Prentice Hall.

Anderson, S. W., Crist, A. J., & Payne, N. (2004). Predicting inpatient length of stay with the expanded version of the Brief Psychiatric Rating Scale (version 4.0). *Psychiatric Services, 55*(1), 77-79.

Anthony, W.A. (1993) Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal, 16*, 11–23.

Barker, C., Pistrang, N., & Elliott, R. (2002). *Research methods in clinical psychology: An introduction for students and practitioners* (2nd ed.). England: John Wiley & Sons, Ltd.

Barker, R. (2012). Recovery and risk: Accepting the complexity. In Drennan, G. & Alred, D. (Eds). *Secure recovery: approaches to recovery in forensic mental health settings*. Oxon: Routledge

Baumeister, R. F., & Bushman, B. J. (2007). Angry emotions and aggressive behaviors. In G. Steffgen, & M. Gollwitzer (Eds.), *Emotions and aggressive behavior* (pp. 61-75). Ashland, OH, US: Hogrefe & Huber Publishers; US.

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.

Biancosino, B., Barbui, C., & Grassi, L. (2005). The BPRS-E as predictor of length of stay in a residential facility. *Psychiatric Services, 56*(6), 755-756.

Blair, P. R., Marcus, D. K., & Boccaccini, M. T. (2008). Is there an allegiance effect for assessment instruments? Actuarial risk assessment as an exemplar. *Clinical Psychology: Science and Practice, 14*(4), 346-360.

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behavior, 29*(4), 355-379.

Boram, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines and training. *American Psychologist, 51*, 945-956

Bradley, K. (2009). *The Bradley report: Lord Bradley's review of people with mental health problems or learning disabilities in the criminal justice system*. London, UK: Department of Health

British Psychological Society. (2009) *Code of ethics and conduct: Guidance published by the Ethics Committee of the British Psychological Society.*

Leicester: British Psychological Society

CASP Critical Appraisal Skills Programme (2004). *Critical Appraisal Skills Programme: Making sense of evidence. 12 questions to help you make sense of a cohort study*. Milton Keynes: Milton Keynes Primary Care Trust.

Castro, M., Cockerton, T., & Birke, S. (2002). From discharge to follow-up: A small-scale study of medium secure provision in the independent sector. *The British Journal of Forensic Practice, 4*(3), 31-39.

Centre for Mental Health. (2010). *Working towards women's well-being: Unfinished business*. London: Centre for Mental Health, National Mental Health Development Unit.

Centre for Mental Health. (2011). *Pathways to unlocking secure mental health care*. London: Centre for Mental Health, National Mental Health Development Unit.

Centre for Reviews and Dissemination. (2008). *Systematic reviews: CRD's guidance for undertaking reviews in healthcare*. York: University of York.

Cohen, A., & Eastman, N. (2000). Needs assessment for mentally disordered offenders: Measurement of 'ability to benefit' and outcome. *The British Journal of Psychiatry, 177*(DEC.), 493-498.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates

Coid, J., Kahtan, N., Gault. S., Cook. A., & Jarman. B. (2001). Medium secure forensic psychiatry services: Comparison of seven English health regions. *The British Journal of Psychiatry*, 178: 55-61

Compton, M. T., Craw, J., & Rudisch, B. E. (2006). Determinants of inpatient psychiatric length of stay in an urban county hospital. *Psychiatric Quarterly, 77*(2), 173-188.

Craig, L. A., Browne, K. D., Stringer, I., & Beech, A. (2004). Limitations in actuarial risk assessment of sexual offenders: A methodological note. *The British Journal of Forensic Practice, 6*(1), 16-32.

Creed, F., Tomenson, B., Anthony, P., & Tramner, M. (1997). Predicting length of stay in psychiatry. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences, 27*(4), 961-966.

Dancey, C. P., & Reidy, J. (2011) *Statistics without maths for psychology (*5th ed.*)*. Essex: Pearson Education Limited.

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 and Psychotherapy, 12*, 226-240.

Department of Health. (1983). *The Mental Health Act 1983*. London: HMSO

Department of Health. (2000). *The NHS Plan: A plan for investment, a plan for reform.* London: Department of Health

Department of Health. (2003). *Personality disorder: No longer a diagnosis of exclusion.* London: Department of Health.

Department of Health. (2007a). *Best practice guidance: specification for adult medium secure services.* London: Department of Health.

Department of Health. (2007b). *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.* London: Department of Health National Risk Management Programme.

Department of Health. (2008). *Commissioning for quality and innovation (CQUIN) payment framework.* London: Department of Health.

Department of Health. (2011). *Payment by results: Draft 2012-13 mental health guidance.* London: Department of Health.

Dernevik, M. (1998). Preliminary findings on reliability and validity of the historical-clinical-risk assessment in a forensic psychiatric setting. *Psychology, Crime & Law, 4*(2), 127-137.

Dodge, K. A., & Coid, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality & Social Psychology, 53*(6), 1146-1158.

Douglas, K. S. (2008). Drawing upon contemporary risk assessment and management principles in the revision of the HCR-20 violence risk assessment scheme. Message posted to http://kdouglas.wordpress.com/hcr-20/hcr-20/

Douglas, K. S., Hart, S. D., Webster, C. D., Belfrage, H., & Eaves, D. (2008). HCR: Assessing Risk for Violence (Version 3, Draft 1.0). Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., & Reeves, K. A. (2010). Historical- Clinical - Risk- Management - 20 (HCR-20) Violence Risk Assessment Scheme: Rationale, application and empirical overview. In R. K. Otto, & K. S. Douglas (Eds.), *Handbook of violence risk assessment: International perspectives of forensic mental health* (1st ed.) (pp. 147-186). New York; London: Routledge.

Douglas, K. S., Skeem, J. L., & Nicholson, E. (2011). Research methods in violence risk assessment. In B. Rosenfeld, & S. D. Penrod (Eds.), *Research methods in forensic psychology* (1st ed.) (pp. 325-346). New Jersey: Wiley & Sons, Inc.

Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services, 54*(10), 1372-1379.

Douglas, K. S., Ogloff, J. R. P., 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 & Clinical Psychology, 67*(6), 917-930.

Douglas, K. S., & Webster, C. D. (1999). The HCR-20 violence risk assessment scheme: Concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior, 26*(1), 3-19.

Douglas, K. S., Webster, C. D., Hart. S., Eaves, D., & Ogloff, J. R. P. (Eds) (2001). *HCR-20: Violence risk management companion guide*. Burnaby, BC, Canada: Mental Health, Law and Policy Institute, Simon Fraser Univeristy.

Drennan, G. & Alred, D. (2012). Recovery in forensic mental health settings: from alientation to integration. In Drennan, G. & Alred, D. (Eds). *Secure recovery: approaches to recovery in forensic mental health settings*. Oxon: Routledge

Dunbar, E., Quinones, J., & Crevecoeur, D. (2005). Assessment of hate crime offenders: The role of bias intent in examining violence risk. *Journal of Forensic Psychology Practice, 5*(1), 1-19.

Edwards, J., Steed, P., & Murray, K. (2002). Clinical and forensic outcome 2 years and 5 years after admissions to a medium secure unit. *Journal of Forensic Psychiatry, 13*(1), 68-87

Fairbairn, A. (2007). Payment by results in mental health: the current state of play in England. *Advances in Psychiatric Treatment*. 13, 3-6

Farrington, D. P., Jolliffe, D., & Johnstone, L. (2008). *A systematic review of risk assessment devices in the prediction of future violence.* Glasgow: Scottish Risk Management Authority.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis programme for the social, behavioural and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

Faulkner, P. A., Tobin, M. J., & Weir, M. A. (1994). Predicting the unpredictable: Issues for output-based funding in psychiatric services. *Australian Health Review*, *17*(1), 86-113

Ferguson, C. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice. 40,* (5)*,* 532 -538

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications

Fitzpatrick, R., Chambers, J., Burns, T., Doll, H., Fazel, S., Jenkinson, C., Kaur, A., Knapp, M., Sutton, L., & Yiend, J. (2010). A systematic review of outcome measures used in forensic mental health research with consensus panel opinion. *Health Technol Assess, 14*(18)

Glancy Report. (1974). *Revised report of the working party on security in NHS psychiatric hospitals.* London: HMSO.

Glorney, E., Perkins, D., Adshead, G., McGauley, G., Murray, K., Noak, J., & Sichau, G. (2010). Domains of need in a high secure hospital setting: A model for streamlining care and reducing length of stay. *International Journal of Forensic Mental Health, 9*, 138–148.

Goldney, R. D., Fisher, L. J., & Walmsley, S. H. (1998). The health of the nation outcome scales in psychiatric hospitalisation: A multicentre study examining outcome and prediction of length of stay. *Australian and New Zealand Journal of Psychiatry, 32*(2), 199-205.

Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behavior, 27*(1), 97-114.

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.

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 & Clinical Psychology,*

71(3), 443-451.

Gronlund, N. E. (1998) *Assessment of student achievement* (6th ed.) Boston: Allyn & Bacon

Grove, W. M. & Meehl, P. E. (1996) Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychology, Public Policy and Law, 2*, 293-323

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19-30.

Gudjonsson, G. H., & Haward, L. R. C. (1998). *Forensic psychology: A guide to practice*. London: Routledge.

Hare, R. D. (1991). *The Hare Psychopathy Checklist - Revised (PCL-R)* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

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

Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The risk for sexual violence protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, BC, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.

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. *The British Journal of Psychiatry, 190* (SUPPL. 49), s60-s65.

Health Professions Council. (2007). *Standards of conduct, performance and ethics.* London: Health Professions Council

Heilbrun, K., Yashuhara, K., & Shah, S. (2010). Violence risk assessment tools: Overview and critical analysis. In R. K. Otto, & K. S. Douglas (Eds.), *Handbook of violence risk assessment: International perspectives on forensic mental health* (1st ed.) (pp. 1-18). London: Routledge.

Hodge, J. E. & Renwick, S. J. (2002). Motivating mentally disordered offenders. In McMurran, M. (Ed). *Motivating offenders to change: A guide to enhancing engagement in therapy.* (pp. 221-234). Chichester, England: Wiley & Sons Ltd.

Hodgson, R. E., Lewis, M. & Boardman, J. (2000) The prediction of in-patient length of stay for acute psychiatric admissions. *Journal of Mental Health, 9*, 2, 145-153

Hosmer, D.W. & Lemeshow, S. (2000). *Applied logistic regression*. (2nd ed.) John Wiley & Sons, Inc.

Huntley, D. A., Cho, D. W., Christman, J., & Csernansky, J. G. (1998). Predicting length of stay in an acute psychiatric hospital. *Psychiatric Services, 49*(8), 1049-1053.

Jones, J. & Plowman, C. A. (2005) Risk assessment: A multi-disciplinary approach to

estimating harmful behaviour in mentally disordered offenders. In Wix, S. & Humphreys, M. (Eds) *Multi-disciplinary working in forensic mental health care* (pp. 133-150). UK: Elsevier

Kennedy, J., Wilson, C,. & Cope, R. (1995). Long-stay patients in a regional secure unit. *Journal of Forensic Psychiatry, 6*(3), 541-551.

Khiroya, R., Weaver, T. & Maden, A. (2009). Use and perceived utility of structured violence risk assessments in English medium secure forensic units. *Psychiatric Bulletin*, 33, 129-132.

Kline, P. (1986). A *handbook of test construction: introduction to psychometric design*. New York, NY, US: Methuen.

Kline, P. (1993). *The handbook of psychological testing (*2nd ed*).* Florence, Kentucky, US: Taylor & Frances/Routledge.

Kline, P. (1998). *The new psychometrics: science, psychology and measurement.* Florence, Kentucky, US: Taylor & Frances/Routledge.

Kmet, L. M., Lee, R. C., & Cook, L., S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields.* Alberta Heritage Foundation For Medical Research, Canada.

Macpherson, G. J. D., & Kevan, I. (2004). Predictive validity of the HCR-20 violence risk assessment scheme within a maximum security special hospital. *Issues in Forensic Psychology, 5*, 62-80.

Maden, A. (2007) *Treating violence: A guide to risk management in mental health.* Oxford: Oxford University Press

Maden. A., Rutter., S., McClinton., T., Friendship. C., & Gunn. J. (1999). Outcome of admission to a medium secure psychiatric unit. 1. Short and long-term outcome. *The British Journal of Psychiatry*, 175: 313-316

Mann, C. J. (2003). Observational research methods. Research design II: cohort, cross sectional and case-control studies. Emergency Med Journal, *20*, 54-60.

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.

Millon, T., Millon, C., Davis, R. & Grossman, S. (1994) *Millon Clinical Multi-axial Inventory* (3rd ed.) Minneapolis, MN: Pearson Education

Ministry of Justice (2010). *Statistics of mentally disordered offenders 2008 England and Wales.* London: Ministry of Justice

Mohan, D., Murray, K., Taylor, P., & Steed, P. (1997). Developments in the use of regional secure unit beds over a 12-year period. *Journal of Forensic Psychiatry, 8*(22), 321-35.

Monahan, J. (1981). *Predicting violent behaviour: An assessment of clinical techniques*. Beverly Hills, California, US: Sage.

Monahan, J., & Steadman, H. J. (1994). Towards a rejuvenation of risk assessment research. In J. Monahan, & H. J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 1-20). Chicago, US: University of Chicago Press.

Moore, B. (1996). *Risk assessment: A practitioner's guide to predicting harmful behaviour.* London: Whiting & Birch Ltd.

Moran, M. J., Fragala, M. R., Wise, B. F., & Novak, T. L. (1999). Factors affecting length of stay on maximum security in a forensic psychiatric hospital. *International Journal of Offender Therapy and Comparative Criminology, 43*(3), 262-274.

Nagtegaal, M.H., Horst, R.P. & van der Schönberger, H.J.M. (2011). *Length of stay in forensic psychiatric hospitals: Numbers and possible explanations.* The Hague: Boom Juridische Uitgevers

National Secure Services Quality, Innovation, Productivity and Prevention Working Group. (2012). *Shared care pathway*. Leicester: National Secure Services Quality, Innovation, Productivity and Prevention Working Group (QIPP)

Nikolova, N.L., Collins, M.J., Guy, L.S., Lavoie, J.A.A., Reeves, K.A., Wilson, C.M. & Douglas, K.S. (2006). HCR-20 violence risk assessment scheme: A quantitative synthesis of its application, reliability, and validity. Poster presented at the annual conference of the American Psychology-Law Society (Division 41 of the American Psychological Association), St. Petersburg, FL.

Novaco, R. (1994). Anger as a risk factor for violence amongst the mentally disordered. In J. Monahan, & H. J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 21-60). Chicago: University of Chicago Press.

Overall, J. E. & Gorham. D. R. (1962) The Brief Psychiatric Rating Scale. *Psychological Reports, 10*, 799-812

Parsons, T. (2006). Length of stay: managed care agenda or a measure of clinical efficiency? *Psychiatry MMC, 3(6)*, 46-52.

Pallant, J. (2010). *SPSS survival manual (*4th ed*.).* Berkshire: Open University Press

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A.R. (1996). A Simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*(12)*,* 1373-1379.

Peng, C. J., Lee. K. L. & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*(1*)*, 3-14

Piersma, H. L., & Boes, J. L. (1997). The relationship of length of stay to MCMI-II and MCMI-III change scores. *Journal of Clinical Psychology, 53*(6), 535-542.

Proulx, J., McKibben, A., & Lusignan, R. (2006). *Relationships between affective components and sexual behaviours in sexual aggressors.* Florida: Psychological Assessments Resources, Inc.

Quinsey, V. & Ambtaman, R. (1979). Variables affecting psychiatrists' and teachers' assessments of the dangerousness of mentally ill offenders. *Journal of Consulting and Clinical Psychology*, 47, 353-362

Quinsey, V., Harris, G., Rice, M., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk.* Washington DC: American Psychological Association.

Reed, J. (1997) The need for long-term psychiatric care in medium and low security.

Criminal Behaviour and Mental Health, 7, 201-212

Rice, M. E., Harris, G. T., & Quinsey, V. L. (1996). Treatment for forensic patients. In B. D. Sales, & S. A. Shah (Eds.), *Mental health and law: Research, policy and services* (pp. 141-189). New York: Academic Press.

Ricketts, D., Carnell, H., Davies, S., Kaul, A., & Duggan, C. (2001). First admissions to a regional secure unit over a 16-year period: Changes in demographic and service characteristics. *Journal of Forensic Psychiatry, 12*(1), 78-89

Ritchie, J. H., Dick, D., & Lingham, R. (1994). *The report of the inquiry into the care and treatment of Christopher Clunis*. London: HMSO.

Robinson, R., Fenwick, J., & Wood, S. (2006). *Report of the independent inquiry into the care and treatment of John Barrett*. London: NHS London.

Rutherford, M. & Duggan, S. (2007). *Forensic mental health services: Facts and figures on current provision*. London, UK, The Sainsbury Centre for Mental Health.

Shah, A., Waldron, G., Boast, N., Coid, J. W., & Ullrich, S. (2011). Factors associated with length of admission at a medium secure forensic psychiatric unit. *Journal of Forensic Psychiatry and Psychology, 22*(4), 496-512.

Smith, H., White, T., & MacCall, C. (2004). A comparison of special hospital patients and other admissions to a regional low security unit. *Journal of Forensic Psychiatry & Psychology, 15*(4), 660-668.

Snowden, R. J., 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.

Steffgen, G., & Gollwitzer, M. (2007). *Emotions and Aggressive Behavior.* Ashland, Ohio, US: Hogrefe & Huber Publishers.

Stand, S. & Belfrage, H. (2001). Comparison of HCR-20 scores in violent mentally disordered men and women: Gender differences and similarities. *Psychology, Crime & Law*, 7 71-79

Sugarman, P. & Walker, L. (2007) *Health of the nation outcome scale for users of secure and forensic services (HoNOS – secure) (version 2b)*. London: The Royal College of Psychiatrists Research Unit.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*(1), 1-26.

Taylor. P., Maden. A., & Jones. D. (1996). Long-term medium security hospital units: a service gap of the 1990s? *Criminal Behaviour and Mental Health*, 6, 213-229

The Royal College of Psychiatrists. (2007). *Standards for medium secure units: Quality network for medium secure units*. London: The Royal College of Psychiatrists Centre for Quality Improvement.

Thomas, M. R., Fryer, G. E., Rosenberg, S. A., Kassner, C., Dubovsky, S. L. & Shore, J. H. (1997) Examining the link between high volume providers and shorter inpatient stays. *Psychiatric Services, 48* (11), 1396-1398.

Thompson, E. E., Neighbors, H. W., Munday, C., & Trierweiler, S. (2003). Length of stay, referral to aftercare, and re-hospitalization among psychiatric inpatients. *Psychiatric Services, 54*(9), 1271-1276.

Thornicroft, G. & Tansella, M. (2010). *Mental health outcome measures*. (3rd ed.) London: The Royal College of Psychiatrists.

Trauer, T., Callaly, T., & Hantz, P. (1999). The measurement of improvement during hospitalisation for acute psychiatric illness. *Australian and New Zealand Journal of Psychiatry, 33*(3), 379-384.

Walker, J., Craissati, J., Batson, S., Amos, T., & Knowles, P. (2012). *How to get better value for money from psychiatric care units*. Health Service Journal.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence (version 2)* (2nd ed.). Simon Fraser University, Canada: Mental Health, Law and Policy Institute.

Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk.* Vancouver, Canada: Simon Fraser University and British Columbia Forensic Psychiatric Services Commission.

West Midlands Strategic Commissioning Group. (2010). *Secure services strategy 2010 – 2015.* West Midlands Strategic Commissioning Group.

Wing, J. K., Curtis, R. H. & Beevor, A. S. (1996) *The Health of the Nation Outcome Scales: Report of research development July 1993 – December 1995*. London: Royal College of Psychiatrists

Witt, P., H. (2000). Book Review: A practitioner's view of risk assessment: The HCR-20 and SVR-20. *Behavioral Sciences and the Law, 18*, 791-798

World Health Organization (2002). *The ICD-10 Classification of mental and behavioral disorders: clinical descriptions and diagnosis guidelines*. Geneva, Switzerland: World Health Organization.

Wright, B., O'Neill, C., & Kennedy, H. G., (2008). Admissions to a national forensic hospital 1997 – 2003. *Irish Journal of Psychological Medicine, 25*(1), 17-23.

APPENDICES

APPENDIX (A) Electronic Database Search Syntax

PsycINFO 1987 to February Week 1 2012

1. exp Psychiatric Patients/
2. exp Mentally Ill Offenders/
3. (offend* or criminal* or delinquent* or convict*).ti,ab.
4. 3 or 1 or 2
5. exp Psychiatric Hospitals/
6. (forensic unit* or forensic hospital* or secure unit*).ti,ab.
7. 5 or 6
8. exp Treatment Duration/
9. ("length of stay" or "treatment duration" or "length of treatment" or "length of admission" or "psychiatric hospitalisation" or "inpatient admission" or "psychiatric admission" or "psychiatric detention").ti,ab.
10. 8 or 9
11. 4 and 7 and 10
12. limit 11 to ((320 young adulthood or 340 thirties or 360 middle age) and male)

Embase 1980 to 2012 Week 05

1. (offend* or criminal* or delinquent* or convict*).ti,ab.
2. mental* ill*.ti,ab.
3. (forensic unit* or forensic hospital* or secure unit*).ti,ab.
4. exp Psychiatric Hospitals/
5. exp Inpatients/
6. ("length of stay" or "treatment duration" or "length of treatment" or "length of admission" or "psychiatric hospitalisation" or "inpatient admission" or "psychiatric admission" or "psychiatric detention").ti,ab.
7. 1 or 2
8. 3 or 4
9. 5 or 6
10. 7 and 8 and 9
11. limit 10 to adult <18 to 64 years>

Web of Science

Topic=(("length of stay" or "treatment duration" or "length of treatment" or "length of admission" or "psychiatric hospitalisation" or "inpatient admission" or "psychiatric admission" or "psychiatric detention").) AND Topic=(("forensic unit*" or "forensic hospital*" or "secure unit*" or "psychiatric hospital*").) AND Topic=(("offender*" or "convict*" or "criminal*" or "delinquent*")) AND Topic=(("patient*" or "mental* ill*" or "mental* disorder*" or "inpatient*"))
DocType=All document types; Language=All languages;

CINAHL

S12 S4 and S7 and S10 Limiters - Published Date from: 19870101-20121231
S11 S4 and S7 and S10
S10 S8 or S9
S9 ("length of stay" or "treatment duration" or "length of treatment" or "length of admission" or "psychiatric hospitalisation" or "inpatient admission" or "psychiatric admission" or "psychiatric detention")
S8 (MH "Treatment Duration")
S7 S5 or S6
S6 (forensic unit* or forensic hospital* or secure unit*).
S5 (MH "Hospitals, Psychiatric")
S4 S1 or S2 or S3
S3 (offend* or criminal* or delinquent*or convict*)
S2 (MH "Mentally Ill Offenders")
S1 (MH "Psychiatric Patients+")

APPENDIX (B) References of Included Studies

Castro, M., Cockerton, T., & Birke, S. (2002). From discharge to follow up: A small-scale study of medium secure provision in the independent sector. *The British Journal of Forensic Practice*, *4*(3), 31-39.

Edwards, J., Steed, P., & Murray, K. (2002). Clinical and forensic outcome 2 years and 5 years after admissions to a medium secure unit. *Journal of Forensic Psychiatry, 13*(1), 68-87

Kennedy, J., Wilson, C,. & Cope, R. (1995). Long-stay patients in a regional secure unit. *Journal of Forensic Psychiatry, 6*(3), 541-551.

Ricketts, D., Carnell, H., Davies, S., Kaul, A., & Duggan, C. (2001). First admissions to a regional secure unit over a 16-year period: Changes in demographic and service characteristics. *Journal of Forensic Psychiatry, 12*(1), 78-89

Shah, A., Waldron, G., Boast, N., Coid, J. W., & Ullrich, S. (2011). Factors associated with length of admission at a medium secure forensic psychiatric unit. *Journal of Forensic Psychiatry and Psychology, 22*(4), 496-51

APPENDIX (C) Quality Assessment Scoring Sheet: Cohort Studies

QUESTION	Yes	Partial	No	Unknown	Comments
Screening Questions					
Does the study address length of stay in a medium secure psychiatric population?					
Is there a clear research question?					
Study Design					
Is a cohort study an appropriate method for answering the research questions?					
Has the study addressed the research question?					
Are the limitations of the study clearly stated?					
Sampling Bias					
Is the sample representative of a medium secure psychiatric population?					
Is there sufficient description of the sample characteristics?					
Have the authors considered any confounding factors?					
Have any controls been applied to limit the bias of any confounding factors present?					
Outcome Bias					
Is Length of Stay clearly defined as an outcome?					
Is there evidence of appropriate outcome indicators?					
Were the outcomes selected comparable to those seen in other studies?					
Attrition Bias					
Were the participants blind to the research and the outcomes?					
Statistics					
Was the statistical analysis used appropriate?					
Results					
Are the results clearly stated?					
Are the results significant?					
Have any limitations and biases been addressed?					

Discussion					
Are the conclusions of the study clearly stated?					
Are the conclusions supported by the results?					
Applicability of Findings					
Are the results from this study transferable?					
SCORE					

APPENDIX (D) Data Extraction Form

DATA EXTRACTION FORM.

Author
Title
Date of Publication
Name of Publication Source (Journal etc.)

Eligibility of Study

P	Males	Yes	No
	Male and Female Mixed Population	Yes	No
	Age 18 to 64	Yes	No
	Mentally Disordered Offenders	Yes	No
	Mental Health/ Psychiatric Population	Yes	No
I	Inpatient Setting	Yes	No
C	Forensic, Medium Secure Setting	Yes	No
O	Length of Stay, Treatment Duration	Yes	No
Continue to next stage?		**Yes**	**No**

Methodology
Research Question
Study Design
Recruitment Process
Participant Characteristics
Sample Size
Outcomes Measured
Variables Considered
Standardised Measures Used

Statistical Analysis
Statistical Technique Used
Were confounding variables assessed and controlled for?

Results
What were the results?

What were the conclusions?
Limitations of the study
Strengths of the study
Applicability of findings

Quality Rating Score