

Measuring the Progress and Outcome of Forensic Mental Health Patients

Gregg S. Shinkfield, BSc, MSc(Hons), PGDipClinPsych

Centre for Forensic Behavioural Science

School of Health, Arts and Design, Swinburne University of Technology

Melbourne, Australia

This thesis is submitted for the degree of

Doctor of Philosophy

2017

“Why collect mental health data? Why collate mental health data? Why measure outcomes? Simply: because quality data supports service users on their journey to recovery.”

*Dr Mark Smith
(Te Pou, 2010)*

GENERAL DECLARATION

In accordance with Swinburne University of Technology regulations, the following declarations are made:

I, Gregg Shinkfield, hereby declare that this thesis contains no material which has been accepted for the award to the candidate of any other degree or diploma at any university or equivalent institution, except where due reference is made in the text of the examinable outcome. To the best of the candidate's knowledge this thesis contains no material previously published or written by another person, except where due reference is made in the text of the examinable outcome. Where the work is based on joint research or publications, the thesis discloses the relevant contributions of the respective workers or authors.

As the candidate, I bore principal responsibility for the ideas, research design, implementation and writing the thesis, under the supervision of Professor James Ogloff. In completing this thesis, I worked within the Centre for Forensic Behavioural Science (Swinburne University) and the Thomas Embling Hospital (Forensicare).

Gregg Shinkfield
Doctor of Philosophy Candidate
School of Health Science
Faculty of Health, Art and Design
Swinburne University of Technology

Signature: 

Date: 13th October 2017

COPYRIGHT NOTICES

Notice 1

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular, no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from the thesis.

Notice 2

I warrant that I have obtained, where necessary, permission from the copyright owners to use any third-party copyright material reproduced in the thesis (such as artwork, images, unpublished documents), or to use any of my own published work (such as journal articles) in which the copyright is held by another party (such as publisher, co-author). Documents confirming permission to reproduce such work has been presented in appendix I, J, and K of this thesis.

Gregg Shinkfield
Doctor of Philosophy Candidate
School of Health Science
Faculty of Health, Art and Design
Swinburne University of Technology

Signature: 

Date: 13th October 2017

In the case of chapters two, six, seven and eight, my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status	Nature and extent of candidate's contribution
Chapter Two	A review and analysis of routine outcome measures for forensic mental health services	Published	Reviewed literature, prepared and revised manuscript (85%)
Chapter Six	Use and interpretation of routine outcome measures in forensic mental health	Published	Reviewed literature, obtained ethics approval, collected and coded data, conducted analysis, prepared and revised manuscript (85%)
Chapter Seven	Monitoring risk, security needs, clinical and social functioning within a forensic mental health population	Submitted / Under Review	Reviewed literature, obtained ethics approval, collected and coded data, conducted analysis, prepared and revised manuscript (85%)
Chapter Eight	Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital	Published	Reviewed literature, obtained ethics approval, collected and coded data, conducted analysis, prepared and revised manuscript (85%)

Signed: _____



Date: _____

13th October 2017

ACKNOWLEDGEMENTS

Completing this thesis, within time and without too much undue heartache, has been an incredible testament to the support of my family. Nat, Ben and Daniel; I love you all so very much! Nat, I can't say thank you enough. You have been my rock and my best friend. I look forward to a world without late night writing sessions, to be replaced with adventures yet unwritten. And to my boys, Ben and Daniel, you are my constant inspiration. Whatever you do in life and wherever it takes you; please, please follow your dreams. They are there to be taken, you just have to be brave enough to rise up to the challenge and chase them (and believe me, they don't always come easily... but the rewards can be great!)

My warmest appreciation goes to my supervisor, Professor James Ogloff, for his assistance and continued enthusiasm throughout the past six years. It has been a lengthy journey, working full time and chipping away at this thesis on a part-time basis. Thank you for you for sticking in there with me and for your ongoing advice and support. I also owe thanks to Dr Douglas Bell, whose enduring concern for the needs of forensic mental health clients and his desire to ensure that the best tools were being used to monitor these needs drove the impetus for this study. I also thank the members of the Forensic Mental Health Information Development Expert Advisory Panel for their ongoing contributions to the field of routine outcome measurement in Australia's forensic mental health system.

I would like to acknowledge and thank my parents, Carol and Colin, and brother, Gareth, who have always given me their unwavering support in the pursuit of my goals and dreams. Despite the tyranny of distance, I always feel that you are with me and 'egging me on' – regardless of what challenges I set myself. I also express my undying and most

heartfelt thanks to my parents-in-law, Cathy and Frank McCall. You have been my life support system, helping to keep the world steady and stable when things felt a little chaotic; particularly when deadlines loomed and the rest of the family needed space to get away from ‘the guy typing on the keyboard’.

Finally, I express my thanks to the clinical staff, ward clerks and patients of the Thomas Embling Hospital, without whom this thesis could not have been completed. To those whom I have worked with over my decade at Forensicare, I sincerely hope that I have contributed at least a little to your journey along the way.

TABLE OF CONTENTS

GENERAL DECLARATION	iii
COPYRIGHT NOTICES	iv
ACKNOWLEDGEMENTS	vi
TABLE OF CONTENTS	viii
LIST OF TABLES	xiv
LIST OF FIGURES	xvi
LIST OF APPENDICES	xvii
ABSTRACT	xviii
PART A: INTRODUCTION AND BACKGROUND	1
Chapter One: Outcome Measurement	1
1.1 Overview of Chapter One.....	1
1.2 What is Outcome Measurement?	1
1.3 Outcome Measurement in Australian Mental Health Services	3
1.4 Challenges in the Collection of NOCC Data.....	8
1.5 Measuring Outcomes of Mentally Disordered Offenders	11
1.6 Overview of Chapters.....	16
1.7 Timeframe and Completion of Thesis.....	16
Chapter Two: Outcome Measurement in Forensic Mental Health.....	18
2.1 Overview of Chapter Two.....	18
2.2 Forensic Outcome Measures	18

2.3	Preamble to Published Paper: “A Review and Analysis of Routine Outcome Measures for Forensic Mental Health Services”	18
2.4	Authorship Indication Form: Chapter Two	20
2.5	Declaration by Co-authors.....	21
2.6	Published Paper One: “A Review and Analysis of Routine Outcome Measures for Forensic Mental Health Services”	22
2.7	Addendum to article one: Supplementary information regarding the DUNDRUM Quartet.	43
PART B: RESEARCH METHODS		45
Chapter Three: The current study		45
3.1	Overview of Chapter Three	45
3.2	Overall Approach to Research Methods	45
3.3	Rationale and Overview of Study Aims.....	45
3.4	Aims of the Research	46
3.4.1	Objective One: Review and analysis of the outcome measurement literature to identify tools designed to monitor risk and clinical/social functioning of individuals receiving forensic mental health treatment.	46
3.4.2	Objective Two: Audit of compliance, precision and reliability of outcome measures used in the Thomas Embling Hospital	47
3.4.3	Objective Three: Evaluation of forensic measures against existing NOCC measures (predictive validity, reliability, useability / utility)... ..	48
3.5	Research Questions	51
3.5.1	Research Questions for Objective One.	51

3.5.2	Research Questions for Objective Two.	51
3.5.3	Research Questions for Objective Three.	52
3.5.4	Summary Questions.	52
3.6	Hypotheses	53
3.6.1	Hypotheses for objective one.....	53
3.6.2	Hypotheses for objective two.....	53
3.6.3	Hypotheses for objective three.....	54
Chapter Four: Method		55
4.1	Overview of Chapter Four.....	55
4.2	Overview of Method	55
4.3	Study Setting	55
4.4	Legislative Context.....	56
4.5	Procedure.....	58
4.6	Objective One: Review and Analysis of the Outcome Measurement Literature ..	58
4.7	Objective Two: Audit of Compliance, Precision and Reliability of Current Outcome Measures Used in the Thomas Embling Hospital.....	59
4.7.1	Data collection tool – audit protocol.....	60
4.7.2	Adherence to NOCC protocols.	61
4.7.3	Precision of ratings.	62
4.8	Objective Three: Evaluation of Forensic Measures and NOCC Measures	63
4.8.1	Measures and materials.....	66
4.8.2	NOCC outcome measures.....	66
4.8.3	Forensic outcome measures	69

4.8.4	Staff training and feedback mechanism.....	72
4.8.5	Patient consent	73
4.8.6	Data collection tool	74
4.8.7	Web decision support tools: Web reports portal.....	75
4.9	Study Three: Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population.....	75
4.9.1	Classification of needs by NOCC and forensic based ROM tools	76
4.9.2	Heterogeneity of needs amongst forensic mental health patients.....	76
4.10	Study Four: Comparison of the HoNOS and HoNOS-Secure Within the Thomas Embling Hospital.....	78
4.10.1	Comparison of mean HoNOS scores between forensic and civil mental health patients	78
4.10.2	Correlation of HoNOS and HoNOS-Secure total score and items	78
4.10.3	Predictive ability of HoNOS and HoNOS-secure.....	79
4.11	Data Collection.....	80
4.12	Data Coding Protocols.....	80
4.13	Ethical Approval.....	80
PART C: EMPIRICAL STUDIES		82
Chapter Five: Overview of the Empirical Papers.....		82
Chapter Six: Use and Interpretation of Routine Outcome Measures in Forensic Mental Health		84
6.1	Overview of Chapter Six.....	84

6.2	Preamble to Published Paper: “Use and interpretation of routine outcome measures in forensic mental health”	84
6.3	Authorship Indication Form: Chapter Six	86
6.4	Declaration by Co-authors.....	87
6.5	Published Paper Two: “Use and Interpretation of Routine Outcome Measures in Forensic Mental Health”	88
	Chapter Seven: Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population.....	97
7.1	Overview of Chapter Seven	97
7.2	Preamble to Submitted Paper: “Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population”	97
7.3	Authorship Indication Form for Chapter Seven.....	100
7.4	Declaration by Co-authors.....	101
7.5	Submitted Paper Three: “Monitoring Risk, Security Needs, Clinical and Social Functioning Within a Forensic Mental Health Population”	102
	Chapter Eight: Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital	130
8.1	Overview of Chapter Eight.....	130
8.2	Preamble to Published Paper: “Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital”	130
8.3	Authorship Indication Form: Chapter Eight.....	132
8.4	Declaration by Co-authors.....	133

8.5	Published Paper Four: “Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital”	134
PART D: DISCUSSION		154
Chapter Nine: Integrated discussion.....		154
9.1	Overview of the Research	154
9.2	Overview of Main Findings	159
9.3	Main Findings of Each Study.....	162
9.3.1	Study one.....	163
9.3.2	Study two	164
9.3.3	Study three	166
9.3.4	Study four.....	169
9.4	Integrated Interpretation of Findings.....	172
9.5	Limitations of the Research.....	183
9.6	Implications of this Research	186
9.7	Future Directions.....	194
9.8	Conclusion.....	197
PART E: THESIS REFERENCES.....		200
PART F: APPENDICES		229

LIST OF TABLES

Tables contained within the body of the thesis

Table 1: NOCC measures completed with adult users of mental health services	7
Table 2: NOCC consumer-rated measures used each state/territory	7
Table 3: Summary of HoNOS psychometric properties.....	67
Table 4: Summary of LSP-26 psychometric properties	68
Table 5: Summary of BASIS-32 psychometric properties.....	69
Table 6: Wording changes between the HoNOS and HoNOS-Secure of item two (Non-accidental self-injury).	182
Table 7: Number of patients on acute, subacute and rehabilitation wards who completed the BASIS-32 and CANFOR.	192

Tables contained within article one

Table 1: Comparison of treatment needs for mainstream and forensic mental health consumers	25
Table 2: Forensic Measures Identified	27
Table 3: Domains assessed by each of the 18 tools identified	28
Table 4: Summary of forensic tools meeting criteria at each level of the hierarchy	31
Appendix A: Description of forensic outcome measures.....	36
Appendix B: Psychometric properties of the CANFOR, DUNDRUM quartet, HoNOS-Secure, IMR, MHRM and START	40

Tables contained within article two

Table 1: Descriptive data of patient files audited.....	92
---	----

Table 2: Availability of completed outcome measures forms in clinical files	92
Table 3: Inter-rater agreement (HoNOS Items)	93
Table 4: Inter-rater agreement (LSP-16 Items)	93

Tables contained within article three

Table 1: Assessment tools grouped by needs areas	126
Table 2: ANOVA – Clinical scales by ward acuity	127
Table 3: Mean scores of patients on acute, subacute and rehabilitation wards for each outcome measurement tools.....	128
Table 4: Proportion of patients with high/low clinical and forensic needs, by ward acuity.....	129

Tables contained within article four

Table 1: Structure and item descriptors of the HoNOS and HoNOS-Secure.....	138
Table 2: Sample and hospital population	143
Table 3: Mean scores state wide and forensic populations	143
Table 4: Correlations between HoNOS and HoNOS-secure items/total score with effect sizes.....	144
Table 5: Logistic regression (initial analysis)	145
Table 6: Logistic regression (repeated analysis)	145
Table 7: Logistic regression (combined scales)	146

LIST OF FIGURES

Figures contained within the body of the thesis

Figure 1: Schematic showing the four categories of high/low clinical and forensic needs, with description of each domain	77
Figure 2: Composition of HoNOS and HoNOS-Secure tools	169
Figure 3: Most effective model for predicting classification of patients (HoNOS + Security Scale).....	172
Figure 4: Mean HoNOS scores obtained by the Forensic and State-wide sample, showing t- statistics of statistical significance at each rating occasion (Graphical representation of data presented in chapter eight of this thesis).....	174

Figures contained within article four

Figure 1: Schematic showing the four categories of high/low clinical and forensic needs, with description of each domain	125
---	-----

LIST OF APPENDICES

Appendix A: Explanatory Statement.....	230
Appendix B: Consent Form.....	234
Appendix C: Ethics Approval – Monash University.....	236
Appendix D: Ethics Approval – Swinburne University of Technology.....	238
Appendix E: Operational Approval – Forensicare	241
Appendix F: Audit Protocol	243
Appendix G: Data Collection Tool (Phase One) – File Audit and Validity Check.....	251
Appendix H: Data Collection Tool (Phase Two) – HoNOS-Secure, LSI-R:SV and CANFOR	257
Appendix I: Copyright Notice – International Journal of Forensic Mental Health.....	260
Appendix J: Copyright Notice – International Journal of Mental Health Nursing	262
Appendix K: Copyright Notice – Journal of Forensic Psychiatry & Psychology.....	268
Appendix L: Comparison of HoNOS and HoNOS-Secure wording.....	270

ABSTRACT

Measuring the progress of mental health service users has been an issue of international concern since the early 1980s. In response to this, the implementation of measures that assess broad areas of clinical functioning and mental health status has become common practice. Known collectively as Routine Outcome Measures (ROMs), these tools have shown good utility as a means of identifying and monitoring the clinical needs and social functioning of patients. However, while much of the ROM literature has focused on tools developed for use with civil mental health patients, there has been less attention paid to specialist fields, including forensic mental health. Forensic mental health patients present with a variety of clinical, social, and forensic/security needs, in addition to those that are commonly encountered with civil mental health clients. As such, monitoring the disparate needs of a forensic mental health population is a complex task. At present, forensic mental health services are generally required to adopt ROM tools that were developed for use with generalist populations. Increasingly, concern has been expressed about this practice, and in particular, the applicability of these measures for forensic mental health populations.

The present thesis sought to address this gap in our knowledge regarding ROM in forensic mental health. This was achieved via completing three main objectives. Firstly, to review and analyse the literature in order to identify outcome measures that could potentially be used in a forensic mental health environment. Secondly, to evaluate the ROM tools currently used for this task in Australia and determine whether the difficulties previously identified with these tools might be a function of the way they are being used in a forensic environment (i.e., whether they are being completed in a reliable and valid

manner), or if there is a deficit in the tools themselves. Finally, this thesis sought to clarify the areas of need present in a forensic mental health population; to examine how these needs differ from civil patients; to ascertain if these needs change over the course of a patient's admission; and perhaps most importantly, to evaluate a selected set of forensic mental health outcome measures against the existing ROM tools to determine if there would be any benefit from adopting such measures in the future.

The objectives of this thesis were achieved via a combination of literature review/analysis, as well as conducting three related empirical studies. Taken together, this research provides a valuable addition to the outcome measurement literature and confirms that ROMs are a useful means of tracking the needs of mental health patients within forensic settings. The present study attests that the ROM tools currently used for this task may be limited in their ability to track areas of need pertinent to a forensic mental health population (Article Two: *Use and interpretation of routine outcome measures in forensic mental health*). However, the findings of study one (Article One: *A review and analysis of routine outcome measures for forensic mental health services*) confirm that there are a range of forensic focused tools currently available in the literature that could be purposed for this task. However, evaluation of the forensic ROM tools identified by this study suggested that many of these tools would not in fact add significantly to the monitoring of the broad clinical, social and forensic needs of this client group (Article Three: *Monitoring risk, security needs, clinical and social functioning within a forensic mental health population*). However, ultimately, the findings of this thesis recommend that there would significant benefit to including the 7-item 'security scale' of the HoNOS-Secure to the existing ROM framework employed in Australia (Article Four: *Comparison of HoNOS and*

HoNOS-Secure in a forensic mental health hospital). Implications for clinical practice and policy are explored.

PART A: INTRODUCTION AND BACKGROUND

Chapter One: Outcome Measurement

1.1 Overview of Chapter One

The opening chapter of this thesis introduces the concept of outcome measurement within the field of mental health. It describes the climate in which outcome measurement has evolved and the rationale for introducing this framework internationally and in Australia. The chapter then discusses the difficulties inherent in mental health outcome measurement and focuses specifically on the limitations of applying outcome measurement tools developed for use in civil mental health services to a forensic mental health environment. The overarching aims and structure of the thesis are described.

1.2 What is Outcome Measurement?

The term “outcome measurement” refers to the process by which change is monitored and tracked in a variable of interest (Jenkins, 1990). In contemporary society, outcome measurement has become a ubiquitous part of life and is routinely applied to a range of disciplines such as economics, business, education, and health, to monitor and track changes pertinent to each of those fields.

Within the domain of health and medicine, monitoring of patient outcomes has also burgeoned over the past 40 years (Gilbody, House & Sheldon, 2003). Health outcome measurement is typically concerned with tracking the effect that a treatment, service or

intervention has on an individual person or population (Cohen & Eastman, 2000; Ovretveit, 1995). The increased focus on Routine Outcome Measurement (ROM) in health and medicine emerged largely in the context of international public concern about poor-quality health care and the need to move towards health policy that seeks to “eliminate unacceptable variations in clinical practice and ensure uniformly high-quality care” for all members of a population (Holloway, 2002, p. 1).

To ensure the quality and effectiveness of any health system, it is imperative to be able to measure the impact of health care provision in a valid and reproducible manner (e.g., Cohen & Eastman, 2000; Jenkins, 1990). While some areas of health care lend themselves more easily to the monitoring of specific outcomes (e.g., blood sugar levels in the case of diabetes or blood pressure for hypertension), patient outcomes in other fields, such as mental health, are generally more difficult to define (Holloway, 2002). This is due, in a large part, to the often subjective nature of outcomes in mental health treatment. Consequently, indicators of progress can be less objective and more difficult to operationalise than markers of change in the treatment of medical conditions. As such, the development of ROMs for mental health services has occurred more slowly than in general medicine, and the most effective means of achieving this remains the subject of debate and research. Markers such as mortality/morbidity, duration of treatment, or economic factors have frequently been used as service level indicators of outcome. However, such factors neglect to account for an individual patient’s subjective experience of the health care they receive, or the impact of their illness on day-to-day wellbeing and clinical/social functioning (Gilbody, House & Sheldon, 2003).

1.3 Outcome Measurement in Australian Mental Health Services

In line with the international trend towards developing more effective means of monitoring mental health service provision (Trauer, 2003), the assessment of clinical outcomes for consumers of Australian mental health services has been an issue of high importance for service providers, patients, and policy makers since the early 1990's (Stedman, Yellowlees, Mellsoy, Clarke, & Drake, 1997). In 1992, recognising there were deficiencies in Australia's mental health service, the Commonwealth, state and territory health ministers agreed on a National Mental Health Strategy that set out an approach to improve psychiatric services across the nation (Australian Health Ministers, 1993; Stedman et al., 1997). The strategy proposed that improving the quality and effectiveness of treatment for people with a mental illness could only be achieved through generating sound information upon which systematic change could be measured (Pirkis, Burgess, Kirk, Dodson, & Coombs, 2005). This would be accomplished by employing consumer outcome data to drive service development and to assist with the evaluation of treatment programs (Cohen & Eastman, 2000; Ellwood, 1988; Health Research Council of New Zealand, 2003; Jenkins, 1990; Sederer & Dickey, 1996; Slade et al., 2006; Trauer, 2003; Relman, 1988; Walters et al., 1996). Within the National Mental Health Strategy, national targets that were considered necessary to foster positive change across psychiatric services were proposed and subsequently operationalized, and set out in a series of 5-year National Mental Health Plans (Australian Health Ministers, 1993; 1998; 2003; Brown & Pirkis, 2009).

Under the first (1992 – 1997) and second (1998 – 2003) National Mental Health Plans (Australian Health Ministers, 1993, 1998), the importance of outcome measurement

and the creation of a case-mix classification system for Australian mental health services was described. Of the 38 objectives outlined in the first National Mental Health Plan, two objectives were specifically identified in relation to outcome measurement (Australian Health Ministers, 1993).

These objectives were:

- Objective 30: Institute regular reviews of outcomes of services provided to persons with serious mental health problems and mental disorders as a central component of mental health service delivery (p. 29);
- Objective 32: Encourage the development of national outcome standards for mental health services, and systems for assessing whether services are meeting these standards (p. 30).

Underpinning these recommendations was the notion that through introducing routinely completed outcome measurement tools, this would provide a means of standardising the way in which the clinical needs and social functioning of patients could be identified and monitored. Deploying a set of standardised outcome measures was also conceived as a means of providing a consistent platform for developing individual care and treatment plans for patients. In addition, it was envisioned that the use of standardised tools could serve a range of other functions, including capturing longitudinal information regarding the needs and outcomes of individual patients, irrespective of the mental health service with which they had contact. As such, when examined collectively, it was anticipated that the data generated by these tools would provide a measure of overall

service delivery within any given mental health facility and enable comparisons to be made about service delivery across mental health services. The introduction of routine outcome measures was also driven in part by demands from patients to have greater involvement in their treatment. As such, it was proposed that outcomes data would provide a common platform to facilitate discussion between patients and clinicians (Pirkis et al., 2005).

Since the mid-1990s, Australia has made considerable progress towards implementing routine outcome measures throughout the country's mental health network. In the first instance, instruments that had potential for use as a routine measure of patient need and clinical outcomes were reviewed (Andrews, Peters, & Teesson, 1994; Buckingham Burgess, Solomon, Pirkis, Eagar, 1998; Burgess, Pirkis, Buckingham, Eagar, Solomon, 1999) and field tested (Stedman et al., 1997). This was undertaken as part of the Mental Health Classification and Service Cost project (MH-CASC; Pirkis et al., 2005). A national data collection protocol was also developed to ensure that measures would be completed at key transition points in the treatment process. The protocol specified that the measures should be administered whenever a consumer was admitted to or discharged from a mental health service and every 3 months for those individuals remaining in continuing care (Burgess et al., 2015).

Beginning in August 2002, the implementation of the selected outcome measures commenced in Australia (Australian Mental Health Outcomes and Classification Network, 2016). The primary clinician-rated measures for adult service users (aged 18 – 64 years) included the Health of the Nation Outcome Scales (HoNOS; Wing, Beevor, Curtis, Park, Hadden & Burns, 1998) and the Life Skills Profile-16 (LSP-16; Rosen, Hadzi-Pavlovic & Parker, 1989; see Table 1). In addition, three consumer-rated measures were introduced as

a means of assessing consumer perspectives of treatment outcomes. Specifically, these consumer-focused tools were the Mental Health Inventory (MHI; Veit & Ware, 1983), Behaviour and Symptom Identification Scale (BASIS-32; Eisen, Dill & Grob, 1994) and Kessler 10+ (K-10+; Kessler et al., 2003). All tools were identified as having sound psychometric properties regarding their validity, reliability and utility in mainstream mental health services (Department of Health and Ageing, 2003a; Eagar et al., 2004). These tools are described in greater detail in Chapter 4 (see Methodology; section 4.8.2).

Of the three consumer-rated measures, only one was required to be used within each state or territory, with the governing bodies of each jurisdiction determining which measure would be implemented (see Table 2). Measures for specific populations such as children, youth and older persons were also included.¹ Collectively these measures became referred to as the National Outcomes Casemix Collection (NOCC) and were administered under the Mental Health National Outcomes and Casemix Collection protocol (MH-NOCC; Department of Health and Ageing, 2003a; 2003b).

Routine measurement of consumer outcomes is now occurring within an estimated 95% of public mental health services across Australia (Australian Mental Health Outcomes and Classification Network, 2016a). In addition, an estimated 98% of private hospitals also routinely complete and report on a subset of these measures (Private Mental Health Alliance, 2009). Summary data comparing the clinical profile of clients treated by mental

¹ The reader is referred to the report by the National Mental Health Working Group (2003) for a comprehensive treatise of all mandated measures in use throughout Australia.

health services within Australia have begun to emerge (Australian Mental Health Outcomes and Classification Network, 2005, 2013), with outcome data proving to be useful in assisting with program planning, service development, and resource allocation (Burgess et al., 2015; Coombs & Meehan, 2003; Garland et al., 2003).

Table 1: NOCC measures completed with adult users of mental health services

Measure	Rated By	Overarching Constructs
Health of the Nation Outcome Scales (HoNOS)	Clinician	Mental health and social functioning
Life Skills Profile (LSP-16)	Clinician	Functional ability
Mental Health Inventory (MHI)	Consumer	Psychological distress and wellbeing
Behaviour And Symptom Identification Scale (BASIS-32)	Consumer	Symptom and problem difficulty
Kessler 10+ (K-10+)	Consumer	Non-specific psychological distress

Note. Table adapted from Burgess, Pirkis & Coombs (2006).

Table 2: NOCC consumer-rated measures used each state/territory

State/Territory	Territory
Victoria	BASIS 32
New South Wales	K10+
Tasmania	BASIS 32
Australian Capital Territory	BASIS 32
Northern Territory	K10+
South Australia	K10+
Western Australia	MHI-38
Queensland	MHI-38

Note. BASIS 32 = Behaviour and Symptom Identification Scale; K10+ = Kessler 10; MHI-38 = Mental Health Inventory.

1.4 Challenges in the Collection of NOCC Data

With the introduction of mandatory reporting of NOCC data within Australia, a minimum compliance target of 85% was set for public health organisations (Department of Human Services, 2009). Similar targets were also set for privately owned psychiatric facilities, with a minimum compliance rate of 70% being established (Private Mental Health Alliance, 2009). When the mandatory reporting of NOCC measures was first introduced, rates were far below the 85% minimum standard, with implementation of routine outcome measurement being variable across states and territories (Pirkis et al., 2005). However, compliance has been steadily improving, with the most recently published data set for the period 2014 – 2015 indicating that the average national compliance rates for completion of the HoNOS within adult inpatient services ranged between 92% - 94%, depending on collection occasion (i.e., admission, review or discharge; Australian Mental Health Outcomes and Classification Network, 2016b). This represents a notable increase in compliance rates since April 2005, during which a rate of 79% compliance was recorded (Australian Mental Health Outcomes and Classification Network, 2005), with only 60% of Australian public mental health services participating in routine outcome measurement at that time. There are a number of factors that have likely contributed to this increase in compliance rates, including ongoing monitoring and support provided by the Australian Mental Health Outcomes and Classification Network, as well as incorporating compliance with outcome measurement into the key performance criteria of accreditation and funding for mental health services (Burgess et al., 2015).

The use of ROM within mental health services has now become common practice internationally (Trauer 2010); however, the implementation of such frameworks is highly

variable across jurisdictions. Collection of ROMs is currently occurring within Australia, New Zealand, Canada, Germany, Italy, and the United States (Trauer, 2010). Amongst these, Australia and New Zealand have been heralded as world leaders in this field (Eagar, Trauer & Mellsop, 2005; Pirkis et al., 2005; Slade, 2002b), with both countries possessing a coherently developed approach to treatment-level routine outcome assessment (Jacobs, 2009). Moreover, Australia and New Zealand both have well-developed centralised systems that are supported by government protocols to facilitate data collection nationally (Eagar, Trauer, & Mellsop, 2005; Pirkis et al., 2005; Slade, 2002). Within Australia, it is also possible for members of the public to access aggregated data obtained from all states and territories via a Web Decision Support Tool (wDST), which can be accessed on the Australian Mental Health Outcomes and Classification Network's website (<http://wdst.amhocn.org/>). This system provides public access to aggregated data submitted by each state and territory, and enables the data set to be freely interrogated with regard to a variety of high level descriptors (e.g., age, gender, legal status; Burgess et al., 2015; see also chapter five for additional details of the wDST).

While increasing compliance rates have been lauded during reviews of NOCC data (Australian Mental Health Outcomes and Classification Network, 2016a; Australian Institute of Health and Welfare, 2005; Brown & Pirkis, 2009; Eagar, Burgess & Buckingham, 2003), it has also been noted that compliance with data collection protocols does not necessarily equate with either useful or good quality data (e.g., Australian Mental Health Outcomes and Classification Network, 2008). Indeed, there are a range of challenges regarding the administration and use of these measures; each of which have the potential to affect the validity and reliability of the data produced.

As with all mandated health reporting mechanisms, staff adherence to, or understanding of, assessment processes and recording requirements is vital. Failure to complete measures in an accurate and timely manner may result in the data being omitted, constituting a failure on the part of the organisation to fulfil its mandatory reporting requirement (National Mental Health Working Group, 2003).

Secondly, reliability of ratings is also critical (Hill & Lambert, 2004). Inaccurate data have implications for individual patients, health care organisations and the dataset as a whole. For the individual patient, inaccurate assessment of their clinical and functional needs may impede their access to, transfer, or discharge from services; it may also suggest a spurious course of illness/recovery; and will remain part of that consumer's permanent clinical record (Burlingame, Lambert, Reisinger, Neff & Mosier, 1995; Saltman, Myers, Kendrick & Fischer, 1998; Vermillion & Pfeiffer, 1993). From an organisational perspective, inaccurate data may not reliably capture the level of acuity or needs of their consumer group, with resulting comparisons made against other services being flawed. If inaccurate and unreliable data were to be a common feature across services, the integrity of casemix information may be compromised and would not permit accurate evaluation of the nation's attainment of the goals set out in the National Mental Health Plans (National Mental Health Working Group, 2003).

Even when all reasonable steps have been taken to ensure staff are compliant and competent with reporting requirements, and that such ratings are made in an accurate and reliable manner, the potential remains for mechanistic or data entry errors to arise (e.g., Dambro & Weiss, 1988; Ludwick, 2009; Wahi, Parks, Skeate, & Goldin, 2008). It is considered that such errors may appear at key points in the rating process, with the process

of transcribing data from rating sheet into the electronic database being one such potential source of error (Wahi et al., 2008). Previous audits examining data transcription errors within health services located in Victoria have observed an error rate of approximately 1% (e.g., Healthcare Management Advisors, 1999).

1.5 Measuring Outcomes of Mentally Disordered Offenders

While the NOCC collection of measures has demonstrated good utility within mainstream services (e.g., McKay & McDonald, 2008; Pirkis, Burgess, Kirk, Dodson & Coombs, 2005; see also Chapter 2 and Chapter 4 for additional details), concern has been expressed regarding the use of these measures within specialist fields, particularly forensic, dual diagnosis, and indigenous mental health (Maddison, Marlee, Webb, Berry & Whitelock, 2016; National Mental Health Working Group, 2003; Shinkfield & Brennan, 2010; Owens, 2010).

During the inception of the NOCC, the need to investigate measures that might be used with patients of specialist services was explicitly acknowledged. Furthermore, it was noted that the applicability of outcome measures designed for use by general adult mental health services, should be evaluated in a *forensic* context to ensure that they effectively capture the needs of mentally disordered offenders (National Mental Health Working Group, 2003).

The term “mentally disordered offender” (MDO) is used within a variety of psychiatric and legal contexts; however, the definition of this term is often inconsistent. For the purpose of the present study, the following operational definition has been applied:

“A person who comes into contact with the criminal justice system because they have committed, or are suspected of committing, a criminal offence, and: who may be acutely or chronically mentally ill” (NACRO, 2006, pp 2).

Moreover, within Australia, the *National Statement of Principles for Forensic Mental Health* (Mental Health Standing Committee, 2006) relates specifically to offenders who are referred for psychiatric assessment or treatment, as well as people that are found not fit to enter a plea, or found not guilty by reason of mental impairment. Additionally, these principals also apply to people in mainstream mental health services who are a significant danger to others and who require the involvement of a specialist forensic mental health service (Commonwealth of Australia, 2006).

The laws that apply to the provision of treatment for MDOs within Australia vary from state to state. Legislation governing the provision of involuntary treatment to MDOs and civil (non-forensic) patients frequently differs with respect to grounds for discharge and length of admission. Depending upon the legal requirements within a jurisdiction, MDOs may remain in a treatment facility even in the absence of psychiatric symptoms, with treatment being focused on issues of risk and other forensic needs (e.g., Andreasson et al., 2014; Turner & Salter, 2008). As such, forensic mental health patients can remain in psychiatric care even following amelioration of acute mental health difficulties, since legislation that detains them requires that they only be discharged to the community when they no longer represent a ‘serious endangerment’ to the community (e.g., *Crimes (Mental Impairment and Unfitness to be Tried) Act*, 1997).

At an elementary level, there are many similarities between the social and clinical needs experienced by patients of forensic and civil mental health services (Shaw, 2002). However, it is recognised that MDOs form a heterogeneous group that may fall into any diagnostic category (Cohen & Eastman, 1997). This has significant implications for measurement of outcomes for this population. As Dickens and colleagues (Dickens, Sugarman & Walker, 2007) note, users of forensic psychiatric services present not only with the mental health difficulties and functional impairments seen in general psychiatry, but they can also demonstrate a history of criminal behaviour, violent or sexual offending, personality and behavioural disturbances, self-harm, and/or co-morbid substance use (Coid, Kahtan, Gault, Cook, & Jarman, 2001; Davoren et al., 2015; Ogloff, Lemphers, & Dwyer, 2004; Ogloff, Talevski, Lemphers, Simmons & Wood, 2015). In addition, consideration frequently needs to be given to level of security, risks, and/or risk management that is required for this client group (Davoren et al., 2015; Kennedy, O'Neill, Flynn & Gill, 2010; Shaw 2002). As such, outcomes for MDOs encompass a wide variety of problem areas, including those beyond narrowly defined mental health outcomes (Cohen & Eastman, 2000). It is also considered that offending behaviour within this population can arise from factors that are not causally related to mental health difficulties. Rather, such behaviour may stem from criminological factors that are similar to non-mental health forensic populations (Dickens, Sugarman & Walker, 2007; Cohen & Eastman, 2000).

Treatment within forensic mental health services therefore seeks not only to provide symptomatic relief from mental illness, but also the amelioration of additional risks that these clients present to themselves and others (Andreasson et al., 2014; Davoren et al., 2015; Mullen, 2006). As such, while the focus of care within civil mental health services is on patients' mental wellbeing, the treatment of mental illness is a necessary but not

sufficient focus in forensic mental health. Due to extra treatment and legislative demands, the average length of inpatient care received by forensic mental health patients is significantly longer than that received by their non-forensic peers (Turner & Salter, 2008). Indeed, this was demonstrated empirically in a recent cross-sectional study of forensic and non-forensic service users, whereby patients detained under a forensic mental health order were found to remain in hospital significantly longer than those detained under civil mental health legislation (Sharma et al., 2015). Given these important points of difference between forensic and civil mental health services, it has been hypothesised that the current outcome measures may be limited in their ability to monitor the broader range of needs inherent in a forensic population. Moreover, it may not be appropriate to compare forensic and non-forensic services using the same tools, nor expect the same level of performance – without first identifying those clinical and risk elements that contextualise and influence outcome (Australian Mental Health Outcomes and Classification Network, 2008). Despite this, there are currently no outcome measures in the NOCC suite that were designed or validated for use with a forensic psychiatric population (Department of Human Services, 2008).

In recent years, the differences between consumers of forensic and non-forensic mental health services have been acknowledged (Australian Health Ministers, 2008). Given the heterogeneity that exists amongst forensic mental health clients (Shinkfield & Ogloff, under review), outcome measurement for this population should be broad enough to capture the disparate clinical and risk related needs that relate to a client's progress towards discharge. This position was made explicit in a report issued by the Victorian Government entitled "*Because Mental Health Matters*", which placed an increased focus on addressing the needs of patients of specialist services (Department of Human Services, 2008). This report also acknowledged the burgeoning demand for forensic psychiatric

services within Australia and noted that a significant proportion of people within the criminal justice system experienced psychiatric difficulties. Support for this assertion in Victoria prisons emerged in 2006, when Ogloff and colleagues demonstrated that 28% of newly remanded adult offenders in Australia had some form of mental illness; with over 50% of prisoners reporting previous assessment and/or treatment for mental health difficulties (Ogloff, Davis, Rivers, & Ross, 2006). This was consistent with the findings of an earlier study conducted within New South Wales, which found 48% of newly received prisoners and 38% of sentenced inmates had suffered a psychotic, affective or anxiety disorder in the twelve months prior to their incarceration (Butler & Allnutt, 2003).

With the publication of “*Because Mental Health Matters*” (Department of Human Services, 2008), the government’s plan for mental health services, a number of goals were outlined to address the deficits identified. These objectives included improving the planning of clinical pathways for forensic clients, strengthening coordination between services, and the development of *common assessment tools suitable for measuring the range of needs possessed by a forensic psychiatric population* (Department of Human Services, 2008; emphasis added). However, in the most recent review of MH-NOCC (National Mental Health Information Development Expert Advisory Panel, 2013), it was reported by the National Mental Health Information Development Expert Advisory Panel that in the ensuing five years no such measure or measures had been identified, nor had an evaluation been undertaken of the existing tools used within a forensic context. It was asserted that a clear gap remains in the measures employed for forensic services with respect to outcomes relating to risk, security and legal issues. The present study therefore seeks to address this gap in our knowledge by investigating the use of routine outcome measures within forensic population, specifically the use of such tools with MDOs.

1.6 Overview of Chapters

This thesis begins with a review of the outcome measurement literature, particularly as it pertains to forensic mental health in the Australian context (Current Chapter). This is followed by a review of existing measures, which could potentially be employed for the task of routine outcome measurement in a forensic mental health environment (Chapter 2). This literature review concludes with a discussion of the aims of this thesis (Chapter 3) and research methods employed (Chapter 4). A series of three empirical research papers are then presented, in which the aims of the thesis were investigated (Chapters 5 – 8). Finally, this thesis ends with an integrated general discussion, linking the findings of all experiments and thereby addressing the overall aims of the thesis (Chapter 9).

1.7 Timeframe and Completion of Thesis

This thesis was completed on a part-time basis, by publication, over a period of 6 years (2010 – 2017). Throughout this time, various components of the study were completed, prepared as manuscripts, and submitted for publication in academic journals. Each of the articles comprising the body of the thesis were written with reference to the contemporaneous literature available at the time of publication. However, over the course of completing this thesis, additional studies have appeared in the literature that were not available for inclusion in the original publications that comprise the body of this thesis. Therefore, where relevant, studies that became available after the manuscripts within this thesis were published have included within the integrated discussion chapter of this thesis. Moreover, supplementary information has also been included at the conclusion of chapters containing published articles, where relevant.

It is also acknowledged that the literature which emerged over the course of the project has shifted the landscape of forensic mental health outcome measurement somewhat since this project was commenced. For example, new measures have been developed that were not available at the time the study was established, which were therefore not able to be included in the project. Likewise, additional support has been generated for several of the measures reviewed in this thesis, which was not available at the time each of the articles were accepted for publication. However, as the literature that was reviewed at the outset of the thesis served as the basis for each of the empirical studies, it was not possible to retrospectively include new tools or findings into the present research. These factors are discussed further in the integrated discussion section of this thesis and consideration is given to how the evolving literature pertaining to forensic mental health outcome measurement impacts upon the findings of the present body of work.

Chapter Two: Outcome Measurement in Forensic Mental Health

2.1 Overview of Chapter Two

This chapter introduces a review of the existing outcome measurement literature, with a focus on identifying tools developed for use within forensic mental health environment.

2.2 Forensic Outcome Measures

In the previous chapter it was asserted that outcome measurement in forensic mental health suffers from a lack of tools designed to effectively capture the disparate needs of forensic mental health patients. At the time of developing the NOCC collection of tools for use across mental health services across Australia, it was noted that there was a paucity of tools developed and validated for use with this population. However, given the increased focus on this population in recent decades and indeed on outcome measurement more broadly, it was anticipated that the number of tools designed for this purpose would have increased over the sixteen years since the NOCC suite was developed.

2.3 Preamble to Published Paper: “A Review and Analysis of Routine Outcome Measures for Forensic Mental Health Services”

The first publication from this thesis reviews outcome measures that could potentially be used in a forensic mental health setting as measures of *functioning*, *recovery*, *risk*, and *placement pathways*. Analysis of the instruments identified was conducted to

evaluate their suitability for this task against a series of specified criteria. Finally, recommendations were offered for future research and development of outcome measurement tools for use with this population.

The review of forensic outcome measures was undertaken in 2011 and was subsequently accepted for publication in 2014. As such, the review reflects the state of the forensic outcome measurement literature up to and including the year 2011. The findings of this review subsequently shaped the direction of the project, as well as decisions regarding the tools selected for further investigation, which formed the basis for the empirical studies contained within this thesis.

The following article was published in the *International Journal of Forensic Mental Health*. This is a peer-reviewed journal of the International Association of Forensic Mental Health Services (ISSN 1499-9013 [Print], 1932-9903 [Online]), which has been published since 2002 and now is published four times per year. In 2014, the *International Journal of Forensic Mental Health* had an impact factor of 1.054.

2.4 Authorship Indication Form: Chapter Two



Swinburne Research

Authorship Indication Form

For PhD (including associated papers) candidates

NOTE

This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each associated 'paper'. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each associated paper to be included in your thesis.

DECLARATION

We hereby declare our contribution to the publication of the 'paper' entitled:

Measuring the progress and outcome of forensic mental health patients

First Author

Name: Gregg Shinkfield

Signature: _____

Percentage of contribution: 85%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Reviewed literature, conducted analysis, prepared and revised manuscript.*

Second Author

Name: Professor James Ogloff

Signature: _____

Percentage of contribution: 15%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Assisted with conceptualisation of study and manuscript, revised manuscript.*

Principal Coordinating Supervisor: Professor James Ogloff

Signature: _____

Date: _____

2.5 Declaration by Co-authors

The undersigned hereby certify that:

- a) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
- b) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- c) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- d) there are no other authors of the publication according to these criteria;
- e) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- f) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s):

Centre for Forensic Behavioural Science,
Swinburne University of Technology and Forensicare
505 Hoddle Street, Clifton Hill
Victoria, Australia

Professor J. Ogloff:		Date
----------------------	--	------

2.6 Published Paper One: “A Review and Analysis of Routine Outcome Measures for Forensic Mental Health Services”

A Review and Analysis of Routine Outcome Measures for Forensic Mental Health Services

Gregg Shinkfield and James Ogloff

Victorian Institute of Forensic Mental Health (Forensicare), Victoria, Australia; Centre for Forensic Behavioural Science, Swinburne University of Technology, Victoria, Australia

Considerable progress has been made in recent years towards implementing routine outcome measures within mental health services. However, the applicability of these tools for forensic-mental health populations has been questioned. A review and analysis was conducted to identify tools that could validly be applied in a forensic context, to provide a measure of *functioning*, *recovery*, *risk*, and *placement pathways*. Nineteen instruments were initially identified and evaluated against a hierarchy of criteria. While no tool assessed all domains of interest, six tools were ultimately considered to have potential utility as outcome measures for users of forensic mental health services.

Keywords: outcome measure, forensic, mental health

Measuring the progress of mental health service consumers¹ has been an issue of international concern since the early 1980s (Stedman, Yellowlees, Mellsoy, Clarke, & Drake, 1997; Trauer, 2010). Policy makers across many jurisdictions have acknowledged that regular evaluation of consumer outcomes data is necessary to ensure effective delivery of services (Health Research Council of New Zealand, 2003; Quality Network for Forensic Mental Health Services, 2009). In response to this need for standardized client data, significant advances have been made within a number of countries towards implementing routine outcome measures (ROMs). While there is little consensus between jurisdictions as to how outcomes are evaluated, collection of ROMs is currently occurring within Australia, New Zealand, Canada, Germany, Italy, and the United States (Trauer, 2010). Amongst these, Australia and New Zealand have been heralded as world leaders in this field, with each possessing well-developed centralized systems that are supported by government protocols to facilitate data collection nationally (Eagar, Trauer, & Mellsoy, 2005; Pirkis et al., 2005; Slade, 2002). Standardized suites of ROMs have been

mandated for use by all mental health facilities in both jurisdictions, and are referred to as the National Outcomes Case-mix Collection (NOCC) in Australia and the Program for the Integration of Mental Health Data (PRIMHD) in New Zealand.² Both suites of ROMs have been designed to collect a broad range of demographic and casemix data and require the completion of a 12-item clinician-rated measure of mental health and social functioning known as the Health of the Nation Outcome Survey (HoNOS; Wing et al., 1998). In addition, the NOCC suite also requires that ambulatory services complete the Life Skills Profile – 16 (LSP-16; Rosen, Hadzi-Pavlovic, & Parker, 1989) to obtain a measure of psychosocial functioning, and all services offer consumers the option to complete one of two self-report measures of symptomatology and functioning. The consumer rated tools used for this purpose are the Behaviour and Symptom Identification Scale (BASIS-32; Eisen, Dill, & Grob, 1994) and Kessler 10+ (K-10+; Kessler et al., 2003).

Within Australia, routine measurement of consumer outcomes now occurs in an estimated 85% of public mental

Address correspondence to Gregg Shinkfield, Victorian Institute of Forensic Mental Health, Locked Bag 10, Fairfield, Victoria, 3078, Australia. E-mail: gregg.shinkfield@forensicare.vic.gov.au

¹The terms consumer, client and patient have been used interchangeably within this review to refer to a person who has engaged with a mental health service for assessment or treatment.

²The reader is referred to reports by the National Mental Health Working Group (2003; http://amhocn.org/static/files/assets/5ddbb17d/NOCC_Specs_V1.5.pdf) and Beveridge, Papps, Bower, & Smith (2012; <http://www.tepou.co.nz/download/asset/502>) for a review of all mandated measures currently used in Australia and New Zealand.

health services (Brown & Pirkis, 2009) and 98% of private hospitals (Private Mental Health Alliance, 2009). Consumer outcomes data are employed for a range of tasks, including service evaluation, program planning, and resource allocation, as well as to assist with developing treatment plans for individual consumers (Coombs & Meehan, 2003; Garland, Kruse, & Aarons, 2003; Slade et al., 2006; Trauer, 2003; Walters et al., 1996). Summary data comparing the clinical profile of clients treated by mental health services have begun to emerge (Australian Mental Health Outcomes & Classification Network, 2005) with these data being used to influence the development of state and federal health policy (Brown & Pirkis, 2009; Coombs & Meehan, 2003). However, while the ROM tools used within Australia and New Zealand have demonstrated good utility within general mental health services (e.g., McKay & McDonald, 2008; Pirkis, Burgess, Kirk, Dodson & Coombs, 2005), concern has been expressed regarding the use of these measures within specialist fields; particularly forensic, dual diagnosis, and indigenous psychiatry (Department of Health and Ageing, 2003; National Mental Health Working Group, 2003). The present review focuses on the use of ROM within forensic populations, specifically with Mentally Disordered Offenders.

The term "Mentally Disordered Offender" (MDO) is used within a variety of psychiatric and legal contexts; however, the definition of this term often varies across studies and regions. For the purpose of the present review, the following definition has been applied: a person who comes into contact with the criminal justice system because they have committed, or are suspected of committing, a criminal offense, and who may be acutely or chronically mentally ill (Nacro, 2006). At an elementary level, there are many similarities between the social and clinical needs experienced by patients in forensic mental health setting and their counterparts in general mental health (Shaw, 2002). However, several authors have noted that users of forensic mental health services present not only with the mental health difficulties and functional impairments seen in general psychiatry, but they also demonstrate a history of criminal behavior, including violent or sexual offending, with a high prevalence of co-morbid personality disorder, behavioral disturbance, self-harm and/or substance use (Dickens, Sugarman, & Walker, 2007; Ogloff, Lemphers, & Dwyer, 2004). In addition, consideration frequently needs to be given to level of security, level of risk, and risk management that is required for this client group (Kennedy, O'Neill, Flynn & Gill, 2010; Shaw, 2002). As such, treatment within forensic mental health services seeks not only to provide symptomatic relief from mental illness, but also the amelioration of additional risks that these clients present to themselves and others (Mullen, 2006). Legislation governing the provision of involuntary treatment to MDOs and patients in general mental health services frequently differ across jurisdictions with respect to grounds for discharge

and length of admission. Depending upon the legal requirements within each jurisdiction, a MDO may be required to remain in a treatment facility even in the absence of psychiatric symptoms, with treatment being focused on issues of risk and other forensic needs. Due to these extra treatment demands, the average length of inpatient care received by forensic consumers is significantly longer than that provided to their non-forensic peers (Turner & Salter, 2008).

Given these important points of difference between forensic and general mental health service users, it has been hypothesized that the ROMs currently used in Australia may be limited in their ability to assess the broader range of needs inherent in a forensic population. Indeed, it may not be appropriate to compare forensic and general mental health services using the same tools, nor expect the same level of performance without first identifying those clinical and risk elements that contextualize and influence outcome (Australian Mental Health Outcomes & Classification Network, 2008). Australia is not unique in that it is currently experiencing burgeoning demand for forensic mental health services, with a significant proportion of people who enter the criminal justice system having psychiatric difficulties. Recent studies have reported that 28% of newly remanded adult offenders in Australia had some form of mental illness; with over 50% of prisoners having received previous mental health assessment or treatment (Ogloff, Davis, Rivers, & Ross, 2006). Despite this, there are currently no outcome measures in the NOCC suite that were designed or validated for use with a forensic psychiatric population (Department of Human Services, 2008).

Due to the increased focus on ROM in recent years, several reviews have been published regarding tools that could be used within forensic mental health services. However, to date, these reviews have generally concluded that few tools were available that would be effective in fulfilling this role (e.g., Vess, 2001). It has also been reported that the development of tools designed to assess the broad range of needs and outcomes for a forensic mental health population has received far less attention than tools designed specifically to measure *security* and *risk* related factors (Thomas et al., 2008). With an overall lack of progress reported in this area, some reviewers have expanded their search to focus on outcome measures used either within a forensic-research context (Chambers et al., 2009; Fitzpatrick et al., 2010), or in specific treatment contexts such as those addressing substance use disorders (Deady, 2009). In addition, several surveys of the literature have focused on outcome measurement tools used solely within in general mental health. These reviews have been conducted within Australia (Andrews, Peters, & Teesson, 1994; Pirkis et al., 2005) and internationally (National Institute for Mental Health in England, 2008).

Of the tools that currently exist, which were designed and validated for use in a forensic environment, there appears to be little consensus regarding which measures are most appropriate for use (Chambers et al., 2009). Concerns

regarding the applicability, reliability and sensitivity of several measures have also been raised (Dickens, Sugarman, & Walker, 2007; Vess, 2001). However, the authors are aware that since the publication of the reviews cited above, several new tools have emerged in the literature. Moreover, studies describing the psychometric properties, validity and reliability of previously reviewed tools have also been published. Thus, a survey of the literature was undertaken, with the aim of updating the current knowledge base regarding ROM tools designed to measure the functioning of individuals receiving treatment in forensic mental health settings. In doing so, this review also aimed to identify measures that that could be applied in an Australian context to assist in the assessment, treatment planning, monitoring and prediction of outcomes for forensic psychiatric patients nationally. It was anticipated that any measures identified might be field tested in the future, with a view to being included in a suite of tools for use in such settings. The present review was established and funded by the Forensic Mental Health Information Development Expert Advisory Panel (FMHIDEAP), a transnational committee established to provide direction for the future development of forensic mental health information in Australia.

During the development of a protocol for the present review, it was noted the concept of ‘functioning’ was a broad term that encompassed a range of distinct domains. For example, service providers frequently discuss clinical functioning, cognitive functioning, vocational functioning, and psychosocial functioning. Other aspects of functioning in the forensic context could also include; risk, recidivism, recovery, personality traits, intellectual capacity, and

substance misuse. Related to each of these factors is the notion of a consumer’s *placement pathway*, or the treatment environment that would be best address that individual’s treatment needs (e.g., mental health, offending, and psychosocial factors), while also providing the level of security required to maintain safety. The concept of ‘functioning’ in this context was subsequently delineated to include the broad domains of *functioning (clinical/psychosocial)*, *recovery*, *risk*, and *placement pathway*, as outlined in Table 1.

As well as specifying the domains listed above, a further set of criteria were proposed by FMHIDEAP to guide the selection of tools. These criteria were deemed necessary to ensure that each tool had undergone adequate scientific scrutiny and could feasibly be applied in the present jurisdiction. A substantive framework already exists within Australia regarding the criteria that tools should meet in order to be included in this suite of measures (e.g., Burgess, Pirkis, Coombs, & Rosen, 2010; Department of Health and Ageing, 2003). It was therefore specified by FMHIDEAP that each of the measures identified should:

- (1) Explicitly measure domains related to functioning (clinical/psychosocial), recovery, risk, and placement pathway (Table 1);
- (2) Be brief and easy to use (≤ 50 items);
- (3) Yield quantitative data;
- (4) Have been scientifically scrutinized and used in two or more peer reviewed studies;
- (5) Be applicable to the local jurisdiction;
- (6) Be applicable for both inpatient and outpatient environments; and

TABLE 1
Comparison of Treatment Needs for General Mental Health and Forensic Mental Health Consumers

Domain	Non-Forensic Clients	Forensic Clients
Functioning (clinical / psychosocial)	Psychiatric Symptoms Psychosocial - Relationships (including social withdrawal) - Personality - Activities of Daily Living Cognitive Insight (Mental Health) Physical Health Vocational (including activities)	Psychiatric Symptoms Psychosocial - Relationships (including social withdrawal) - Personality - Activities of Daily Living Cognitive Insight (Mental Health, Offending) Physical Health Vocational (including activities)
Recovery	Client Perspective of recovery Service Perspective of recovery	Client Perspective of recovery Service Perspective of recovery
Risk	to SELF to OTHERS of SUBSTANCE USE	to SELF to OTHERS of SUBSTANCE USE of RE-OFFENDING - Violence - Sexual Violence - General Recidivism
Placement Pathway	Purpose of treatment (mental health)	Level of security required (i.e., low, medium or high) Whether current security placement is appropriate Legislative requirements Purpose of treatment (mental health, offending)

- (7) Demonstrate sound psychometric properties (e.g., of internal consistency, validity, reliability and sensitivity to change).

These criteria were used to analyze and exclude instruments by employing a hierarchical criterion-based analysis. In doing so, an instrument excluded on the basis of one criterion was not assessed against any subsequent criteria. An identical approach was used by Burgess and colleagues in a recent review of Recovery Focused measures (Burgess et al., 2010).

METHOD

In the first instance, the MEDLINE and PSYCINFO electronic databases were searched for pertinent literature published between the years 1980 to 2011.³ Any potentially relevant peer-reviewed journal articles were retrieved, and their reference lists scanned for further measurement related articles. Databases were also scanned for literature citing each of the articles identified and forward searching was used to obtain the most recently published work. The names of any potentially relevant measures identified were subsequently included as additional search terms and the above sources were further interrogated. In addition, other sources of information were obtained via the Internet. This included both local and international sources, with a focus on reports produced by government agencies, academic facilities and a range of professional bodies. Finally, experts in the fields of outcome measurement and forensic mental health were contacted regarding pre-publication or unpublished work.

Hierarchical Analysis

A hierarchical criterion-based analysis was used to evaluate each of the instruments identified. This approach was such that any instrument judged by the authors as failing to meet any of the criteria outlined below would be excluded from subsequent analysis. The following criteria were employed for this evaluation process:

³Search terms used: ["outcome measure*" or "routine outcome measure*" or ("outcome" & "measure*") or ("need*" & "assess*") or "risk assessment" or "instruments" or "measurement" or "achievement measure*" or "aptitude measure*" or "attitude measure*" or "criterion referenced tests" or "group testing" or "individual testing" or "instrument*" or "inventor*" or "occupational measure*" or "perform* tests" or "personality measure*" or "post test*" or "pre test*" or "profile*" or "psych* evaluation" or "psych* assessment" or "psychometrics" or "questionnaires" or "rating" or "screening" or "sociometric" or "standardized tests" or "statistical measure*" or "surveys" or "symptom checklists" or "test*" or "recovery" or "recovery measure*"] and ["forens*" or "justice" or "criminal" or "offend*"] (Limiters: Publication year: January 1980 – May 2011).

- (1) Tools should explicitly measure domains related to functioning.
 - The tools identified were required to measure most, or all, of the domains relating to *Functioning*, *Recovery*, *Risk*, and *Placement Pathway*, as outlined in Table 1.
- (2) Tools should be brief and easy to use.
 - It should be possible for a single clinician, regardless of discipline, to complete the tool
 - Tools should contain no greater than 50 items and should not require extensive training or expertise in the administration of psychometric instruments.
 - Where an instrument is rated on the basis of a structured or semi-structured interview, the average time of interview should be 30 minutes or less.
- (3) Tools should yield quantitative data.
 - To permit rapid data entry and the storage of large quantities of outcome data, it was deemed necessary that these tools should generate quantifiable and easily coded information (i.e., did not contain descriptive or other qualitative data).
 - It was further considered that quantitative data would permit the most efficient means of tracking the progress of individual consumers and would readily lend itself to investigation by statistical analysis.
- (4) Tools should have been scientifically scrutinized.
 - Tools were required to have featured in two or more peer-reviewed studies.
- (5) Tools should be applicable for both inpatient and outpatient environments.
 - Tools were required to demonstrate efficacy across a range of settings and could be used with both inpatient and outpatient consumer groups.
- (6) Tools should demonstrate sound psychometric properties.
 - Tools were required to have demonstrated an adequate degree of internal consistency, validity, reliability and sensitivity to change.
- (7) Tools should be applicable to the local jurisdiction.
 - Tools were required to have either been validated for use in an Australian context, or have been constructed with language and concepts readily accessible and applicable to an Australian population.

RESULTS

Overview of Instruments Identified

Article retrieval occurred between December 2010 and May 2011. At its conclusion, a total of 19 instruments were

TABLE 2
Forensic Measures Identified

Abbreviation	Instrument Name	Key Reference(s)
ASP	Atascadero Skills Profile	Vess, 2001
BSI	Behavioural Status Index (B.E.S.T. Index)	Woods & Reed, 1999
BJMHS	Brief Jail Mental Health Screen	Osher, Scott, Steadman, & Robbins, 2004
BPRS	Brief Psychiatric Rating Scale	Overall & Gorham, 1962; Ventura et al., 1993
CANFOR	Camberwell Assessment of Needs: Forensic Version	Thomas et al., 2003
CARDINAL	Cardinal Needs Scale	Marshall, Hogg, Gath, & Lockwood, 1995
CJ-CEST	Criminal Justice – Client Evaluation of Self and Treatment	Simpson, 2005
DOST	Defendant and Offender Screening Tool	Ferguson & Negy, 2006
DUNDRUM	The Dundrum Quartet	Kennedy, O'Neill, Flynn, & Gill, 2010
FIOS	Forensic Inpatient Observation Scale	Timmerman, Vastenburger, Emmelkamp, 2001
HCR-20 v2	Historical Clinical Risk – 20 (Version 2)	Webster, Douglas, Eaves, & Hart, 1997
HoNOS-Secure	Health of the Nation Outcome Scales: for users of secure/forensic services	Sugarman & Walker, 2007
IMR	Illness Management and Recovery Scales	Mueser et al., 2004
LS/CMI	Level of Service/Case Management Inventory	Andrews, Bonta, & Wormith, 2004
LSI-R:SV	Level of Service Inventory (Revised: Screening Version)	Andrews & Bonta, 1998
MHRM	Mental Health Recovery Measure	Young & Bullock, 2003; Bullock et al., 2002
PANSS	Positive and Negative Syndrome Scales	Kay, Opler, & Fiszbein, 2000
PIC-R	Problem Identification Checklist – Revised Version	Ostapiuk, Stringer & Craig, 2000; Nagi, Ostapiuk, Craig, & Hacker, 2009
START	Short Term Assessment of Risk and Treatability	Webster, Martin, Brink, Nicholls & Desmarais, 2009

identified as having been developed or validated for use within forensic mental health settings (Table 2). These 19 instruments each appeared to be potentially useful for evaluating some aspect of *Functioning*, *Recovery*, *Risk* or *Placement Pathway*. A summary profile of these instruments has been provided in Table 3.

Hierarchal Analysis

Criterion 1: Explicitly Measures Domains Related to Functioning

Each tool was first examined with respect to its ability to assess *Functioning*, *Recovery*, *Risk* and *Placement Pathway*. Table 3 summarizes the domains covered by each of the 19 tools. Close examination of these data revealed that no single tool assessed all of the domains specified. Notably, the domains that were least represented were those of *personality* and *client's perspective of recovery*. The lack of tools that focused on these domains was not entirely unexpected. In a review of recovery focused measures (Burgess et al., 2010) it was found that of the 22 tools identified, only one had been formally validated within a forensic context and one additional tool was reported to have been designed for *potential* use with an offender population. With respect to the measurement of personality, of those tools that included items which explicitly assessed this domain (e.g., Historical Clinical Risk – 20 version 2 [HCR-20] and Problem Identification Checklist – Revised Version [PIC-R]), personality was presented as a historical factor and therefore not considered amenable to change (e.g., HCR-20 item H9 *personality disorder*). As such, the concept of personality *per se* did not appear to lend itself well to outcome assessment, particularly over a relatively

short period (e.g., three months). While, it may be possible to infer aspects of an individual's personality through examining the dynamic items of these tools, (e.g., HCR-20 Clinical items *Negative Attitudes* [C2] and *Impulsivity* [C4]; or START items *Relationships* [2], *Rule Adherence* [15], *Conduct* [16], or *Coping* [19]) the use of items in this manner is not explicitly described within their respective professional manuals.

In contrast to *personality* and *recovery*, the domain most widely represented was that of *risk to others*. This finding was consistent with previous reviews (e.g., Thomas et al., 2008), which observed an over representation of tools focusing on risk in comparison to measures assessing other areas of need.

While no single tool measured all of the domains proposed; five tools were found to cover a broad range of domains, these were: Atascadero Skills Profile (ASP), Camberwell Assessment of Needs: Forensic Version (CANFOR), The Dundrum Quartet (DUNDRUM [scales 1–4]), Health of the Nation Outcome Scales: for users of secure/forensic services (HoNOS-Secure) and Short-Term Assessment of Risk and Treatability (START). These measures were notable in that they included at least one item from the domains *risk*, *recovery* and *placement pathway*, and assessed multiple items from the *functional* domains.

Several tools were also deemed suitable for further consideration as they appeared to assess a smaller, yet adequate, subset of the domains proposed. As such, it may be possible to use these tools in concert with other instruments to provide a full assessment of an individual's needs. Specifically these were the Behavioural Status Index (BSI), Cardinal Needs Scale (CARDINAL), Criminal Justice – Client Evaluation of Self and Treatment (CJ-CEST),

TABLE 3
Domains Assessed by Each of the 19 Tools Identified

Domain	Sub-Domain	ASP	BSI	BIMHS	BPRS	CANFOR	CARDINAL	CJ-CEST	DOST	DUNDRUM	FIOS	HCR-20 (v2)	HoNOS-Secure	IMR	LS/CMI	LSI-R:SV	MHRM	PANSS	PIC-R	START
1 Functioning (clinical / psychosocial)	Psychiatric Symptoms (including psychosis)	✓	✓		✓	✓	✓	✓	✓	✓ ^b	✓	(✓)	✓	✓	✓		(✓)	✓	✓	✓
	Relationships (including withdrawal)	✓	✓		✓	✓	✓	✓	✓	✓ ^c	✓	✓	✓	✓	✓	✓	(✓)	✓	✓	✓
	Personality Activities of Daily Living	✓	✓		✓	✓	✓			✓ ^d	✓	✓ ³	✓	(✓)	(✓)		(✓)		(✓)	✓
	Cognitive Insight	(✓)	✓		(✓)	✓	✓	✓	✓	✓ ^d	✓	✓	✓	✓			✓	✓	(✓)	✓
	Physical Health	✓	✓	✓	✓	✓	✓			✓ ^e	✓	✓	✓	✓	✓	✓	✓	✓	(✓)	✓
2 Recovery	Vocational (including activities)	✓		✓	✓	✓	✓			✓ ^c	✓		✓	✓	✓	✓	✓	✓		✓
	Client Perspective	✓			(✓)	(✓)	(✓)		✓	✓ ^d				✓	✓	✓	✓			✓
	Service Perspective	✓	✓		(✓)	(✓)	(✓)	✓	✓	✓ ^{a/b}			✓	✓	✓		✓	✓		✓
3 Risk	to SELF	✓ ¹	✓		✓ ²	✓ ²	✓ ²			✓ ^{a/d}	✓	✓	✓	✓	✓	✓	✓	✓		✓
	to OTHERS	✓ ¹	✓		✓ ²	✓ ²	✓ ²			✓ ^c			(✓)	✓	✓	✓	✓	✓		✓
4 Placement Pathway	of RE-OFFENDING	✓ ¹	✓		✓ ²	✓ ²	✓ ²	✓	✓	✓ ^e			✓	✓	✓	✓				✓
	of SUBSTANCE USE	✓ ¹			✓ ²	✓ ²	✓ ²			✓ ^e			✓	✓	✓	✓				✓
4 Placement Pathway	Level of security required / appropriate									✓ ^{a/b}			✓		✓	✓				✓

Note. ✓ = Domain included, (✓) = Domain only partially or indirectly included.

¹ASP has been listed here as it focuses on an individual's ability of to manage these risk domains.

²CANFOR includes user/staff assessment of the domains of safety to self/others, alcohol/drug, Arson/Sexual Offending.

³Personality Disorder and Psychopathy are coded as historical factors in items H7 and H9 of HCR-20(v2).

^aDUNDRUM 1 (Triage security), ^bDUNDRUM 2 (Triage urgency), ^cDUNDRUM 3 (Program completion), ^dDUNDRUM 4 (Recovery items).

Forensic Inpatient Observation Scale (FIOS), Illness Management and Recovery Scales (IMR), Level of Service/Case Management Inventory (LS/CMI), and the Problem Identification Checklist – Revised Version (PIC-R).

There were six tools, however, that were not considered to encompass a broad-enough range of domains to warrant further analysis. While the Brief Psychiatric Rating Scale (BPRS) and Positive and Negative Syndrome Scale (PANSS) were seemingly useful tools for assessing change in psychiatric illness, these instruments did not address the domains of recovery, risk or placement pathway. This was also found to be true for a range of other psychiatric symptom rating scales commonly used in non-forensic mental health settings (e.g., Brief Symptom Inventory [BSI], Derogatis, 1993). Likewise, the Defendant and Offender Screening Tool (DOST) provided little information outside of the areas of risk, symptoms and cognitive functioning. The Level of Service Inventory—Revised: Screening Version (LSI-R:SV) was considered to be focused too specifically on offending related risk factors and provided little information about functioning or a consumer's recovery. The Brief Jail Mental Health Screen (BJMHS) proved to be too brief to assess adequately any of the required domains. Finally, although the HCR-20 (version 2)⁴ was noted to be a well-established tool for the assessment of risk of violence to others, the specificity of this tool precluded its use for describing other dynamic domains of risk such as harm to self, substance use or general recidivism. Moreover, although the HCR-20 contained one item pertaining to symptomatology (i.e., *Active Symptoms of Major Mental Illness* [C3]), other aspects of functioning not directly related to violence were absent (e.g., self-care, cognitive functioning and physical health). This tool was also found to lack items related to recovery or placement pathway. As such, the relatively narrow focus of the HCR-20 (v2) limited its utility as a broad measure of outcome in forensic mental health settings. Finally, the HCR-20 (v2) requires formal training and experience to ensure valid completion, requires a comprehensive review of collateral material and a comprehensive clinical interview to complete, and it is most typically used by psychologists and psychiatrists rather than other mental health professionals.

The results of analyses and considerations similar to that undertaken for the HCR-20 (v2) also held true for a number of other prominent risk assessment measures, such as the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice & Cormier, 1998), Violence Risk Scale (VRS; Wong & Gordon, 1996), and Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003), amongst others. While each of these tools were found to be well validated and have proven

useful for monitoring the constructs they were intended to measure, such tools were deemed not to assess the broad range of functioning and mental health outcomes pertinent to forensic mental health clients. As such, these tools were not considered to fit within scope of the research question.

Consequently the BPRS, DOST, HCR-20, LSI-R:SV and BJMHS were excluded from further analysis.

Criterion 2: Is Brief and Easy to Use (≤50 Items)

Of the remaining 13 instruments, the BSI, CARDINAL and CJ-CEST could be definitely excluded on the basis of length (see Appendix A). While the PIC-R fell slightly outside of the inclusion criteria (57 items), it was considered close enough to be retained for further analysis.

As well as those instruments that were excluded on the basis of length, the LS/CMI was also eliminated at this point as it did not meet the criteria 'easy to use.' While this tool has demonstrated efficacy in the prediction and management of general recidivism, the administration manual cautions that test users are required to either have familiarity with psychometric testing or should undertake specific training before using this tool (Andrews et al., 2004). As such, this may limit the range of clinicians who could administer the LS/CMI in a valid and reliable manner.

The criterion 'easy to use' was also carefully considered in relation to the START. While this tool consists of only 20 items (with two additional case-specific items permitted), the START utilizes a structured professional judgment approach and co-rating by members of a multidisciplinary team is encouraged (Webster et al., 2009). As such, this raised concerns as to whether or not this tool could be completed easily by individual clinicians, from a variety of mental health disciplines, as a mandatory measure of patient outcome. However, the authors of the START suggest that no specific qualifications are required for use of this tool. Furthermore, they specifically note that the START was originally developed for use by members of the nursing profession, whose professional training does not typically include the use of psychometric instruments (Webster et al., 2006). Support for these assertions has been found in a series of recently published articles, which have attested to the ease of use and applicability of this measure in both forensic and general psychiatric settings (Crocker et al., 2011; Doyle, Lewis, & Brisbane, 2008; Webster et al., 2009). Clinicians from a range of professional backgrounds (i.e., nursing, social work and psychology) were surveyed regarding the experience of using the START. Of these, 79% reported no difficulties in using this tool, 65% reported feeling *fairly* or *very* confident in their ratings, and 92% reported that the information required to complete the START was readably available or easily obtained (Crocker et al., 2011). This builds on earlier work by Doyle and colleagues (Doyle et al., 2008), who similarly reported 82% of respondents indicated that they

⁴The authors are aware that the HCR-20 has recently been updated to version 3 (HCR-20^{V3}; Douglas, Hart, Webster, & Belfrage, 2013), however this was not available for consideration during the present review period.

experienced no difficulties using this tool, 95% reported feeling *fairly*, *moderately* or *very* confident in their ratings, and 95% reported that the information required to complete the START was readably available or easily obtained. The time taken to administer the START has been documented as ranging between 20 – 30 minutes; however, longer times were reported while staff were unfamiliar with the tool (Crocker et al., 2011; Doyle et al. 2008). In light of these findings, the START was retained for further evaluation in the present review.

Therefore, after considering the issues of brevity and ease of use, four instruments were excluded (i.e., BSI, CARDINAL, CJ-CEST and LS/CMI) leaving nine for further analysis.

Criterion 3: Yields Quantitative Data

Each of the remaining nine instruments were noted to produce quantitative data (see Appendix A); therefore, no instruments were excluded based on this criterion.

Criterion 4: Has Been Scientifically Scrutinized and Used in Two or More Peer Reviewed Studies

With the exception of the PIC-R, each of the remaining nine instruments had been made available for scrutiny in the professional literature, either in the form of an assessment manual or journal publication. In addition, all nine instruments had accrued at least one validation study conducted by the tool's author or development team (relevant publications denoted with [†] in reference list). Four tools were noted to have been the subject of validation studies conducted by independent research teams, namely: CANFOR (Emmanuel & Campbell, 2009; Long et al., 2008; Romeva et al., 2010), HoNOS-Secure (Pillay, Oliver, Butler, & Kennedy, 2008; Segal et al., 2010), MHRM (Andresen, Caputi & Oades, 2010; Bullock et al., 2009), and START (Braithwaite, Charette, Crocker, & Reyes, 2010; Doyle et al., 2008; Kroppan et al, 2011; Nonstad et al, 2010; Wilson, Desmarais, Nicholls, & Brink, 2010). However, three of the nine instruments were deemed to have not been sufficiently scrutinized in the professional literature. No evidence could be obtained of the ASP or FIOS appearing in peer-reviewed studies other than those in which the tool was first published. In addition, while the authors of the PIC-R have indicated that in the 10 years since this tool was first published, it has featured in a number of studies evaluating its use as a dynamic measure of change (personal communication, L. Craig, June 2011); unfortunately these studies remain largely unpublished and unavailable for peer review in the professional literature. Should empirical support emerge for any of these tools in the future, it is considered that each may warrant further investigation as a measure of patient outcome in a forensic context.

Careful consideration was also given to the MHRM, as it had been noted in a previous review of recovery focused

measures (Burgess et al., 2010) that empirical support for this tool has largely been presented in the form of unpublished reports and conference papers. However, it was encouraging to learn that since the publication of the review by Burgess and colleagues the MHRM has featured in two studies as a measure of consumer recovery (Andresen et al., 2010; Bullock et al., 2009). Therefore, the MHRM was retained in the present review for further analysis.

Criterion 5: Can Be Applied to the Local Jurisdiction

Close examination of the wording and concepts described within each of the six remaining tools suggests that they could all be readily adapted for use in a range of jurisdictions with only minor alterations to the text (i.e., by substituting names of programs or services from the local jurisdiction). However, it was noted that each of the tools had been written using English language and had been constructed based upon western constructs of mental health, service provision and social structure. As such, adapting these tools may present greater challenges for those jurisdictions in which different language or social constructs exist.

While none of these six instruments were developed within the jurisdiction of the present study (i.e., Australia), three were found to have been featured in literature that described their use in an Australian context. Specifically, the CANFOR and HoNOS-Secure have both been the subject of published research projects conducted within at least one Australian forensic service (e.g., CANFOR: Thomas et al, 2008; HoNOS-Secure: Segal et al., 2010). In addition, a modified version of the START has been included as part of a broader assessment package developed by an Australian community-forensic service (Carroll, 2008). No difficulties have been reported in using these tools within this context. Of the three remaining instruments (e.g., DUNDRUM, IMR and MHRM), each was described by their respective authors as having been designed and tested with a diverse ethnic sample of consumers across a range of psychiatric environments (e.g., Bullock, 2005; Kennedy et al., 2010; Mueser et al., 2004). All six instruments were therefore retained for further analysis on the basis of *applicability across jurisdictions*.

Criterion 6: Is Applicable for Both Inpatient and Outpatient Environments

Each of the remaining six instruments were reported by their respective authors to be applicable with both inpatient and outpatient client groups.

Criterion 7: Demonstrates Sound Psychometric Properties

Appendix B summarizes the psychometric properties of the remaining six instruments. As noted, consideration was

given to the robustness of the following properties: *Internal Consistency*, *Validity*, *Reliability* and *Sensitivity to change*.

While no specific threshold was set for each of the psychometric properties examined, of the six tools evaluated, it was not deemed necessary to exclude any instrument based on this criteria. That is, no studies were located that suggested any of these tools were invalid, unreliable or insensitive to change. Each of the six tools investigated were found to have demonstrated adequate validity, and most had been evaluated with several clinical populations and against multiple criteria (e.g., construct validity, concurrent validity, and predictive validity). However, only three tools demonstrated evidence of satisfactory performance across all psychometric domains, namely the HoNOS-Secure, MHRM, and START (refer to Appendix B for details).

Conversely, no information could be obtained regarding the *sensitivity to detect change* for three tools (e.g., CANFOR, DUNDRUM & IMR), and the item structure of the CANFOR had not been tested for *internal consistency*. While these tools were reportedly valid and reliable, the absence of information regarding sensitivity for detecting change was considered a potential weakness and further research to address this is recommended.

SUMMARY

Table 4 summarizes the instruments deemed to meet the threshold criteria at each of the seven levels of the hierarchy. Ultimately, the 19 instruments identified at the outset of this review, were reduced to six: Camberwell Assessment of Needs: Forensic Version, DUNDRUM Quartet, Health of the Nation Outcome Scale for Users of Secure / Forensic Services, Illness Management and Recovery Scales, Mental Health Recovery Measure, and the Short-Term Assessment of Risk and Treatability.

DISCUSSION

The present review was designed to investigate the question, 'What tools are currently available to measure the *functioning*, *risk*, *recovery* and *placement pathways* of individuals receiving mental health treatment in forensic settings?' At the conclusion of this study, 19 instruments had been identified and each was evaluated against a set of seven criteria regarding content, length, psychometric properties and applicability. Analysis subsequently revealed that six of these tools met the requirements specified: CANFOR, DUNDRUM, HoNOS-Secure, IMR, MHRM and the START.

While no individual tool assessed all four domains of *functioning* (clinical / psychosocial), *recovery*, *risk*, and *placement pathway*, each had strengths and weaknesses relative to the others. To this end, all six tools demonstrated

TABLE 4
Summary of Forensic Tools Meeting Criteria at Each Level of the Hierarchy

All instruments	ASP	FIOS
	BSI	HCR-20
	BJMHS	HoNOS-Secure
	BPRS / PANSS	IMR
	CANFOR	LS/CMI
	CARDINAL	LSI-R:SV
	CJ-CEST	MHRM
	DOST	PIC-R
	DUNDRUM	START
1. Measures domains related to functioning	ASP	HoNOS-Secure
	BSI	IMR
	CANFOR	LS/CMI
	CARDINAL	MHRM
	CJ-CEST	PIC-R
	DUNDRUM	START
2. Brief and easy to use	FIOS	
	ASP	IMR
	CANFOR	MHRM
	DUNDRUM	PIC-R
	HoNOS-Secure	START
3. Quantitative data	FIOS	
	ASP	IMR
	CANFOR	MHRM
	DUNDRUM	PIC-R
	HoNOS-Secure	START
4. Scientifically scrutinised	FIOS	
	ASP	IMR
	CANFOR	MHRM
	DUNDRUM	PIC-R
	HoNOS-Secure	START
5. Local Jurisdiction	CANFOR	IMR
	DUNDRUM	MHRM
	HoNOS-Secure	START
6. Inpatient and outpatient	CANFOR	IMR
	DUNDRUM	MHRM
	HoNOS-Secure	START
7. Psychometric properties	CANFOR	IMR
	DUNDRUM	MHRM
	HoNOS-Secure	START

utility as repeatable measures of patient *functioning*. However, only the IMR and MHRM monitored a consumer's progress towards *recovery* goals; yet, neither offered any information regarding risk or placement pathway. The HoNOS-Secure, START and DUNDRUM measured *functioning*, *risk* and *placement pathway*, but did not address the issue of recovery. Finally, the CANFOR assessed *functioning*, *risk* and was deemed to partially monitor aspects of *recovery*; however, it offered little information regarding placement pathway. In the absence of a single tool that could comprehensively assess all four domains, it was considered possible that several of the instruments identified might be combined to form a *suite* of forensic outcome measures. This is consistent with the recommendation offered by Andrews, Peters, and Teesson (1994), following their scoping work from which the NOCC suite of measures was developed.

It should also be noted, however, that of these six tools, four did not fit precisely within all of the criteria specified at each level of analysis. In the case of the CANFOR, DUN-DRUM and IMR, each lacked information regarding their internal consistency or sensitivity to change. However, as there was no evidence to suggest that these tools performed poorly on either of these psychometric domains (indeed each had demonstrated sound validity and reliability); all three were retained on the proviso that further psychometric evaluation be conducted. Consequently, if the criteria specified at the outset of the study were applied strictly, only the HoNOS-Secure and START could be said to fully meet all seven of the standards required. As such, these two measures might serve as a useful starting point for future investigations into the development of a suite of forensic outcome measures; however, the remaining four tools may also warrant further consideration.

Before turning to the concluding remarks of this study, a brief comment is offered regarding the range of tools included in this review. Specifically, it is recognized that several prominent risk assessment measures did not feature in this report (e.g., VRAG, VRS, and RSVP). Each of these tools were found to focus specifically on assessing of domains of risk, rather than evaluating the broad range of functioning and mental health outcomes pertinent to forensic mental health clients. As such, they did not easily fit within the scope of the present study. This point was illustrated by the early elimination of the HCR-20 and LS/CMII risk assessment tools from the hierarchical analysis. Although these tools are likely to be considered an essential component of forensic assessments in many services, their relatively narrow focus on specific forms of risk may not produce meaningful results for all forensic clients and key aspects of consumer outcome could be overlooked.

This does not suggest, however, that risk assessment tools such as the HCR-20, VRAG and RSVP cannot be usefully employed in the ongoing evaluation of forensic clients. Indeed, the HCR-20 has been successfully adopted as a routine measure of patient progress in forensic services across the United Kingdom (Quality Network for Forensic Mental Health Services, 2009). Yet, the lead author of the HCR-20 has cautioned against mandating the use of any *specific* risk assessment tool for all service users. Rather, a more flexible model is proposed whereby clinicians should be permitted to select whichever measure of risk is most appropriately employed in each case (personal communication, C. Webster, June 2011).

It may therefore seem incongruous that the START was identified as one of the six tools recommended for further investigation by the present study. However, as a risk assessment tool, the START was designed to take a broad view of risk and examines multiple domains that are relevant to most forensic mental health services, including harm to others, self-harm/suicide, vulnerability, absconding, etc. (Doyle et al., 2008). The START also prompts

clinicians to assess the strengths, vulnerabilities and treatability of an individual by considering 20 items pertaining to mental state, behavior, and functioning (Webster et al., 2009). The START is also easily administered by a diverse range of mental health professionals. As such, the START was considered applicable across a wider range of forensic services than most risk assessment tools. Yet, throughout the present analysis, the START frequently required a greater degree of consideration regarding the seven criteria than each of the other 18 tools identified. Although the START was ultimately considered to meet each of these criteria, there remain a number of potential difficulties that could arise from its use as a ROM. Specifically, concerns were identified in relation to mandating the use of a Structured Professional Judgment tool for clinicians who may be unfamiliar with such assessment techniques. In addition, difficulties may arise from reducing the rich idiographic risk information generated by the START into numerical data. However, literature which directly addresses these concerns has begun to emerge in recent years (Crocker et al., 2011; Doyle et al., 2008), providing the necessary support for the inclusion of this tool.

Limitations

As is often the case with research, there were several limitations inherent in the present review. In the first instance, this review was confined to articles written or translated into English. As such, the possibility remains that tools developed in non-English speaking countries might exist, which if translated and validated for this environment, could prove useful in an Australian context. The present review also relied heavily on electronic means of identifying and obtaining relevant information. Despite broadening the search to include Internet-based reports and other *grey literature* (Hopewell, McDonald, Clarke, & Egger, 2007), it is possible that resources pertinent to this review were not discovered. Finally, while significant effort was made to obtain copies of all articles identified through electronic databases and reference lists, it was not always possible to obtain full text documents. As such, some references were unavoidably omitted. These limitations are not particular to the present review and it is considered that any review of this sort should be viewed as an evolving process. New information should therefore be integrated into these findings as it comes to light.

Future Directions

Several directions for future research have emerged from the present review. Firstly, while each of the six instruments identified demonstrated a reasonable match to the selection criteria, further investigation of the psychometric properties of these tools may be warranted in some cases. It was also highlighted that few tools currently exist that

provide adequate assessment of a consumer's perspective on their own recovery or attainment of treatment needs within a forensic environment. This represents a significant deficit in the recovery literature and is an area that would benefit from further research. Given the limitations of the tools identified by this review, it may be possible to develop a new instrument based on the strengths of each, or through combining several tools into an assessment suite. Alternatively, existing tools might be expanded or modified to include domains of interest to forensic mental health settings. For example, the adolescent version of the START (Nicholls, Viljoen, Cruise, Desmarais, & Webster, 2010) currently includes the domain *General Offending*, which it may be possible to include in future versions of the START. Ultimately though, by expanding the research into the field of outcome measurement for MDOs, it is anticipated that the technology used for such assessments will continue to be refined over time.

CONCLUSION

The findings of this review have confirmed that a number of assessment tools have been designed or validated for use in forensic mental health settings. This represents a significant increase in the number of instruments that have been developed for this purpose over the past decade. Of the tools identified by this review, six were considered feasible for use as routine outcome measures across Australian forensic mental health services, and indeed internationally. However, additional investigations are likely to be required before one or more of these tools can be employed for this task. This review represents the foundation upon which future work could be based.

REFERENCES⁵

- [†]Andresen, R., Caputi, P., & Oades, L. G. (2010). Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Research*, 177, 309–317.
- Andrews, D. A. & Bonta, J. (1995). *The Level of Service Inventory-Revised*. Toronto, ON, Canada: Multi-Health Systems.
- *Andrews, D. A. & Bonta, J. (1998). *The Level of Service Inventory-Revised: Screening version*. Toronto, ON, Canada: Multi-Health Systems.
- Andrews, D. A. & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati, OH: Anderson Publishing Co.
- *Andrews, D. A., Bonta, J., & Wormith, S. J. (2004). *The Level of Service / Case Management Inventory*. Toronto, ON, Canada: Multi-Health Systems.
- Andrews, G., Peters, L., & Teesson, M. (1994). *Measurement of consumer outcome in mental health: A report to the National Mental Health Information Strategy Committee*. Sydney, Australia: Clinical Research Unit for Anxiety Disorders.
- Australian Mental Health Outcomes & Classification Network. (2005). *Adult national outcomes and casemix collection standard reports*. 1st edition, version 1.1. Brisbane, Queensland, Australia.
- Australian Mental Health Outcomes & Classification Network. (2008). National mental health benchmarking project forensic: review of key performance indicators. Retrieved from http://amhocn.org/static/files/assets/7fe17da1/2008_FORENSIC_KPI_Review_DRAFT_FINAL.pdf
- Beveridge, J., Papps, R., Bower, A., & Smith, M. (2012). Mental health outcomes information collection protocol – HoNOS family (v2.1). Retrieved from <http://www.tepou.co.nz/download/asset/502>
- [†]Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *The International Journal of Forensic Mental Health*, 9, 271–281.
- Brown, P. & Pirkis, J. (2009). Mental health quality and outcome measurement and improvement in Australia. *Current Opinion in Psychiatry*, 22, 610–618.
- [†]Bullock, W. A. (2005). The Mental Health Recovery Measure. In T. Campbell-Orde, Chamberlin, J., Carpenter, J., & Leff, H. S. (Ed.), *Measuring the promise of recovery: A compendium of recovery measures, Volume II* (pp. 36–41). Cambridge, MA: The Evaluation Center.
- [†]Bullock, W. A., O'Rourke, M., Farrer, E., Breedlove, A., Smith, M. K., & Claggett, A. (2005). *Evaluation of the Illness Management and Recovery (IMR) Program*. Paper presented at the 113th annual meeting of the American Psychological Association Meeting, Washington, DC.
- [†]Bullock, W. A., Sage, J., Hupp, D., O'Rourke, M., & Smith, M. K. (2009). From illness management and recovery (IMR) to wellness management and recovery (WMR): implementation and evaluation of Ohio's WMR program. *New Research in Mental Health*, 18, 312–321.
- *[†]Bullock, W. A., Wuttke, G. H., Klein, M., Bechtoldt, H., & Martin, J. (2002). Promoting mental health recovery in an urban county jail. *New Research in Mental Health*, 15, 305–314.
- [†]Bullock, W. A. & Young, S. L. (2003). *The Mental Health Recovery Measure (MHRM)*. Paper presented at the 111th annual meeting of the American Psychological Association Meeting, Toronto, Canada.
- Burgess, P., Pirkis, J., Coombs, T., & Rosen, A. (2010). Review of recovery measures (version 1.01). Retrieved from http://amhocn.org/static/files/assets/80e8bfc/Review_of_Recovery_Measures.pdf
- Campbell-Orde, T., Chamberlin, J., Carpenter, J., & Leff, H. S. (2005). *Measuring the promise: a compendium of recovery measures*. Cambridge, MA: Human Services Research Institute.
- Carroll, A. (2008). Risk assessment and management in practice: the Forensic Risk Assessment and Management Exercise. *Australasian Psychiatry*, 16, 412–417.
- Chambers, J. C., Yiend, J., Barrett, B., Burns, T., Doll, H., & Fazel, S. (2009). Outcome measures used in forensic mental health research: a structured review. *Criminal Behaviour and Mental Health*, 19, 9–27.
- Coombs, T., & Meehan, T. (2003). Mental health outcomes in Australia: issues for mental health nurses. *International Journal of Mental Health Nursing*, 12, 163–164.
- Corrado, R. R., Cohen, I. M., Hart, S. D., & Roesch, R. (2000). Diagnosing mental disorders in offenders: conceptual and methodological issues. *Criminal Behaviour and Mental Health*, 10, 29–39.
- Crocker, A. G., Braithwaite, E., Laferriere, D., Gagnon, D. C., V., & Jenkins, T. (2011). START changing practice: implementing a risk assessment and management tool in a civil psychiatric setting. *International Journal of Forensic Mental Health*, 10, 13–28.
- Deady, M. (2009). A review of screening, assessment and outcome measures for drug and alcohol settings. Retrieved from http://www.nada.org.au/downloads/Info_Management_Project/Review_of_Measures_09.pdf
- Department of Health and Ageing. (2003). Mental health national outcomes and casemix collection: technical specification of State and Territory reporting requirements for the outcomes and casemix components

⁵*Denotes key references for ROM tools reviewed in this article.

[†]Denotes references of validation studies for the ROM tools reviewed.

- of 'agreed data', Version 1.50. Canberra, Australia: Commonwealth Department of Health and Ageing.
- Department of Human Services. (2008). Because mental health matters: a new focus for mental health and wellbeing in Victoria, consultation paper. Melbourne, Australia: Victorian Government Department of Human Services.
- Derogatis, L. R. (1993). *Brief Symptom Inventory (BSI): Administration, scoring and procedures manual* (3rd ed.). Minneapolis, MN: NCS Pearson, Inc.
- [†]Dickens, G., Sugarman, P., Picchioni, M., & Long, C. G. (2010). HoNOS-Secure: tracking risk and recovery for men in secure care. *British Journal of Forensic Practice*, 12(4), 36–46.
- [†]Dickens, G., Sugarman, P., & Walker, L. (2007). HoNOS-secure: a reliable outcome measure for users of secure and forensic mental health services. *Journal of Forensic Psychiatry & Psychology*, 18, 507–514.
- Douglas, K. S., Blanchard, A. J. E., Guy, L. S., Reeves, K. A., & Weir, J. (2010). HCR-20 violence risk assessment scheme: overview and annotated bibliography. Retrieved from <http://kdouglas.files.wordpress.com/2007/10/hcr-20-annotated-biblio-sept-2010.pdf>
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk of violence – User guide*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S. & Reeves, K. (2009). The HCR-20 violence risk assessment scheme: overview and review of the research. In R. Otto & K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 147–185). Oxford, UK: Routledge/Taylor & Francis.
- [†]Doyle, M., Lewis, G., & Brisbane, M. (2008). Implementing the Short-Term Assessment of Risk and Treatability in a forensic mental health service. *Psychiatric Bulletin*, 32, 406–408.
- [†]Dwyer, S., Davoren, M., Abidin, Z., Doyle, E., McDonnell, K., & Kennedy, H. G. (2011). The DUNDRUM Quartet: validation of structured professional judgement instruments DUNDRUM-3 assessment of programme completion and DUNDRUM-4 assessment of recovery in forensic mental health services. *BMC Research Notes* 4, 1–12.
- Eagar, K., Trauer, T., & Mellis, G. (2005). Performance of routine outcome measures in adult mental health care. *Australian and New Zealand Journal of Psychiatry*, 39, 713–718.
- Eisen, S. V., Dill, D. L., & Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry*, 45, 242–247.
- Emmanuel, M. K., & Campbell, M. H. (2009). Comparing staff and client perceptions of needs in a British single-sex medium secure unit. *Psychological Reports*, 104, 1033–1034.
- *Ferguson, C. J. & Negy, C. (2006). Development and preliminary validation of a defendant and offender screening tool for psychopathology in inmate populations. *Criminal Justice and Behavior*, 33, 325–346.
- 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 Technology Assessment*, 14.
- [†]Flynn, G., O'Neill, C., & Kennedy, H. G. (2011). DUNDRUM-2: Validation of a structured professional judgment Instrument assessing priority for admission from the waiting list for a Forensic Mental Health Hospital. *BMC Research Notes*, 4. Retrieved from <http://www.biomedcentral.com/1756-0500/4/230>
- [†]Flynn, G., O'Neill, C., McInerney, C., & Kennedy, H. G. (2011). The DUNDRUM-1 structured professional judgement for triage to appropriate levels of therapeutic security: retrospective-cohort validation study. *BMC Psychiatry*, 11. Retrieved from <http://www.biomedcentral.com/1471-244X/11/43>
- Garland, A., Kruse, M., & Aarons, G. (2003). Clinicians and Outcome measurement: what's the use? *Journal of Behavioural & Health Services Research*, 30, 393–405.
- Garner, B. R., Knight, K., Flynn, P. M., Morey, J. T., & Simpson, D. D. (2007). Measuring offender attributes and engagement in treatment using the Client Evaluation of Self and Treatment. *Criminal Justice and Behavior*, 34, 1113–1130.
- Greenwood, A., & Burt, G. (2001). Validity of the brief psychiatric rating scale within a forensic inpatient hospital. *Expert evidence*, 8, 15–30.
- 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*. Vancouver, BC, Canada: The British Columbia Institute Against Family Violence.
- Health Research Council of New Zealand. (2003). *Partnerships for evidence-based policy and practice*. Auckland, New Zealand: Health Research Council of New Zealand.
- Hopewell, S., McDonald, S., Clarke, M. J., & Egger, M. (2007). Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews*.
- *Kay, S. R., Opler, L. A., & Fiszbein, A. (2000). *The Positive and Negative Syndrome Scale (PANSS) manual*. Toronto, ON, Canada: Multi-Health Systems Inc.
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, 60, 184–189.
- *Kennedy, H. G., O'Neill, C., Flynn, G., & Gill, P. (2010). *Dangerousness understanding, recovery and urgency manual (the DUNDRUM quartet) v1.0.22*. Dublin National Forensic Mental Health Service.
- [†]Kroppan, E., Nettet, M. B., Nonstad, K., Pedersen, T. W., Almvik, R., & Palmstierna, T. (2011). Implementation of the Short-Term Assessment of Risk and Treatability (START) in a forensic high secure unit. *The International Journal of Forensic Mental Health*, 10, 7–12.
- [†]Leese, M., Thornicroft, G., Shaw, J., Thomas, S., Mohan, R., Harty, M. A., & Dolan, M. (2006). Ethnic differences among patients in high-security psychiatric hospitals in England. *The British Journal of Psychiatry*, 188, 380–385.
- [†]Long, C. G., Dickens, G., Sugarman, P., Craig, L., Mochty, U., & Hollin, C. R. (2010). Tracking risk profiles and outcome in a medium secure service for women: use of the HoNOS-Secure. *International Journal of Forensic Mental Health*, 9, 215–225.
- [†]Long, C. G., Webster, P., Waine, J., Motala, J., & Hollin, C. R. (2008). Usefulness of the CANFOR-S for measuring needs among mentally disordered offenders resident in medium or low secure hospital services in the UK: a pilot evaluation. *Criminal Behaviour and Mental Health*, 18, 39–48.
- *Marshall, M., Hogg, L. I., Gath, D. H., & Lockwood, A. (1995). The Cardinal Needs Schedule - a modified version of the MRC Needs for Care Assessment Schedule. *Psychological Medicine*, 25, 605–617.
- McKay, R., & McDonald, R. (2008). Expensive detour or a way forward? The experience of routine outcome measurement in an aged care psychiatry service. *Australasian Psychiatry*, 16, 428–432.
- *[†]Mueser, K. T., Gingerich, S., Salyers, M. P., McGuire, A. B., Reyes, R. U., & Cunningham, H. (2004). *The Illness Management and Recovery (IMR) scales (client and clinician versions)*. Concord, NH: Dartmouth Psychiatric Research Center.
- Mullen, P. E. (2006). Schizophrenia and violence: from correlations to preventive strategies. *Advances in Psychiatric Treatment*, 12, 239–248.
- Nacro. (2006). Liaison and diversion for mentally disordered offenders. Retrieved from <http://www.nacro.org.uk/data/files/liaisondiversion-formdos-may06-936.pdf>
- *[†]Nagi, C., Ostapiuk, E. B., Craig, L. A., Hacker, D., & Beech, A. R. (2009). Using the revised Problem Identification Checklist to predict inpatient and community violence: a pilot study. *British Journal of Forensic Practice*, 11(3), 4–13.
- National Institute for Mental Health in England. (2008). Mental health outcomes compendium. Retrieved from http://cpaa.co.uk/files/DH_093677.pdf
- National Mental Health Working Group. (2003). Mental health national outcomes and casemix collection: technical specification of state and

- territory reporting requirements for the outcomes and casemix components of 'agreed data', Version 1.50. Canberra, Australia: Commonwealth Department of Health and Ageing.
- [†]Nicholls, T. L., Brink, J., Desmarais, S., Webster, C. D., & Martin, M.-L. (2006). The Short-Term Assessment of Risk and Treatability (START), a prospective validation study in a forensic sample. *Assessment*, 13, 313–327.
- Nicholls, T.L., Viljoen, J., Cruise, K., Desmarais, S.L. & Webster, C.D. (2010). *Short-Term Assessment of Risk and Treatability: Adolescent Version (START:AV)* (Abbreviated manual). Port Coquitlam, Canada: BC Mental Health and Addictions Services.
- [†]Nonstad, K., Nasset, M. B., Kroppan, E., Pedersen, T. W., Nøttestad, J., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norwegian high secure hospital. *The International Journal of Forensic Mental Health*, 4, 294–299.
- Ogloff, J., Lemphers, A., & Dwyer, C. (2004). Dual diagnosis in an Australian forensic psychiatric hospital: prevalence and implications for services. *Behavioral Sciences and the Law*, 22, 543–562.
- Ogloff, J. R. P., Davis, M. R., Rivers, G., & Ross, S. (2006). *The identification of mental disorders in the criminal justice system*. Canberra, Australia: Australian Institute of Criminology.
- Ohio Department of Mental Health. (2004). *Procedural manual* (6th ed.). Columbus, OH: Ohio Department of Mental Health.
- *Osher, F., Scott, J. E., Steadman, H. J., & Robbins, P. C. (2004). *Validating a brief jail mental health screen: Final report for National Institute of Justice*. Washington, DC: National Institute of Justice.
- *Ostapiuk, E., Stringer, I., & Craig, L. (2000). *Some modifications and revisions to the Problem Identification Checklist*. Unpublished manuscript. Available from Forensic Psychology Practice Ltd, 98 Sheffield Road, Boldmere, Sutton and Coldfield, Birmingham, UK.
- *Overall, J. E. & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports*, 10, 799–812.
- [†]Pillay, S. M., Oliver, B., Butler, L., & Kennedy, H. G. (2008). Risk stratification and the care pathway. *Irish Journal of Psychological Medicine*, 25, 123–127.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005). *Routine measurement of outcomes in Australia's public sector mental health services*. Melbourne, Australia: The University of Melbourne.
- Pirkis, J., Burgess, P., Kirk, P., Dodson, S., & Coombs, T. (2005). *Review of standardised measures used in the National Outcomes and Casemix Collection (NOCC)*. Sydney, Australia: NSW Institute of Psychiatry.
- Private Mental Health Alliance. (2009). *Annual statistical report from the PMHA's centralised data management service regarding the services provided by participating private hospitals with psychiatric beds and private psychiatric day hospitals: 1 July 2007 to 30 June 2008*. Canberra, Australia: Strategic Planning Group for Private Psychiatric Services.
- Quality Network for Forensic Mental Health Services. (2009). *Outcome and activity measures working group: Summary document*. London, UK: Royal College of Psychiatrists.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. (1998). *Violent offenders: appraising and managing risk*. Washington, DC: American Psychological Association.
- [†]Romeva, G. E., Rubio, L. G. G., Güerre, S. O., Miravet, M. J. R., Cáceres, A. G., & Thomas, S. D. M. (2010). Validación clínica de la escala Camberwell Assessment of Need-Forensic version, CANFOR para la evaluación de necesidades de personas con trastorno mental en el ámbito penitenciario. [Clinical validation of the CANFOR scale (Camberwell Assessment of Need-Forensic version) for the needs assessment of people with mental health problems in the forensic services]. *Actas Españolas de Psiquiatría*, 38, 129–137.
- Rosen, A., Hadzi-Pavlovic, D., & Parker, G. (1989). The Life Skills Profile: A measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin*, 15, 325–337.
- [†]Salyers, M. P., Godfrey, J. L., Mueser, K. T., & Labriola, S. (2007). Measuring illness management outcomes: a psychometric study of clinician and consumer rating scales for illness self-management and recovery. *Community Mental Health Journal*, 43, 459–480.
- [†]Segal, A., Daffern, M., Thomas, S., & Ferguson, M. (2010). Needs and risks of patients in a statewide inpatient forensic mental health population. *International Journal of Mental Health Nursing*, 19, 223–230.
- Shaw, J. (2002). Needs assessment for mentally disordered offenders is different. *The Journal of Forensic Psychiatry & Psychology*, 13, 14–17.
- *Simpson, D. D. (2005). *TCU treatment assessment forms*. Fort Worth, TX: Texas Christian University, Institute of Behavioral Research.
- Slade, M. (2002). The use of patient-level outcomes to inform treatment. *Epidemiologia e Psichiatria Sociale*, 11, 20–27.
- Slade, M., McCrone, P., Kuipers, E., Leese, M., Cahill, S., Parabiaghi, A., Priebe, S., & Thornicroft, G. (2006). Use of standardised outcome measures in adult mental health services: randomised controlled trial. *British Journal of Psychiatry*, 189, 330–336.
- Stedman, T., Yellowlees, P., Mellis, G., Clarke, R., & Drake, S. (1997). *Measuring consumer outcomes in mental health: Field testing of selected measures of consumer outcome in mental health*. Canberra, Australia: Department of Health and Family Services.
- [†]Sugarman, P., & Everest, H. (1999). An audit of five regional secure psychiatric services: South Thames NHS Multi-Authority Clinical Audit Group (available from the authors).
- *Sugarman, P., & Walker, L. (2007). Health of the nation scale for users of secure/forensic services (version 2b). Retrieved from <http://www.rcpsych.ac.uk/researchandtrainingunit/honos/secure.aspx>
- [†]Sugarman, P., Walker, L., & Dickens, G. (2009). Managing outcome performance in mental health using HoNOS: experience at St Andrew's Healthcare. *Psychiatric Bulletin*, 33, 285–288.
- *Thomas, S., Harty, M. A., Parrott, J., McCrone, P., Slade, M., & Thornicroft, G. (2003). *The Forensic CAN (CANFOR). A needs assessment for forensic mental health service users*. London, UK: Gaskell.
- [†]Thomas, S., McCrone, P., & Fahy, T. (2009). How do psychiatric patients on prison healthcare centres differ from inpatients in secure psychiatric inpatient units? *Psychology, Crime & Law* 15, 729–742.
- [†]Thomas, S. D. M., Slade, M., McCrone, P., Harty, M. A., Parrott, J., Thornicroft, G., & Leese, M. (2008). The reliability and validity of the forensic Camberwell Assessment of Need (CANFOR): a needs assessment for forensic mental health service users. *International Journal of Methods in Psychiatric Research*, 17, 111–120.
- *[†]Timmerman, I. G. H., Vastenburg, N. C., & Emmelkamp, P. M. G. (2001). The Forensic Inpatient Observation Scale (FIOS): development, reliability and validity. *Criminal Behaviour and Mental Health*, 11, 144–162.
- Trauer, T. (2003). Analysis of outcome measurement data from the four Victorian 'round one' agencies. Retrieved from <http://www.health.vic.gov.au/mentalhealth/outcomes/exec-summary.pdf>
- Trauer, T. (Ed.). (2010). *Outcome measurement in mental health: theory and practice*. Cambridge, UK: University Press.
- Turner, T., & Salter, M. (2008). Forensic psychiatry and general psychiatry: re-examining the relationship. *Psychiatric Bulletin*, 32, 2–6.
- *Ventura, J., Lukoff, D., Nuechterlein, K. H., Liberman, R. P., Green, M., & Shaner, A. (1993). Brief Psychiatric Rating Scale (BPRS): Expanded version 4.0 scales, anchor points, and administration manual. *International Journal of Methods in Psychiatric Research*, 3, 227–243.
- *[†]Vess, J. (2001). Development and implementation of a functional skills measure for forensic psychiatric inpatients. *Journal of Forensic Psychiatry & Psychology*, 12, 592–609.
- Walters, G., Kirkby, K., Marks, I., Whiteford, H., Andrews, G., & Swinson, R. (1996). Outcome measurement: sharing experiences in Australia. *Australasian Psychiatry*, 4, 316–318.
- *Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing the risk for violence* (Version 2). Vancouver, Canada: Simon Fraser University.

*Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START)* (Version 1.1). Port Coquitlam, BC, Canada: Forensic Psychiatric Services Commission and St. Joseph's Healthcare.

†Webster, C. D., Nicholls, T., Martin, M.-L., Desmarais, S. L., & Brink, J. (2006). Short-Term Assessment of Risk and Treatability (START): the case for a structured professional judgment scheme. *Behavioral Sciences and the Law*, 24, 747–766.

†Wilson, C., Desmarais, S. L., Nicholls, T. L., & Brink, J. (2010). The role of client strengths in assessments of violence risk using the Short-Term Assessment of Risk and Treatability (START). *The International Journal of Forensic Mental Health*, 9(4), 282–293.

Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *British Journal of Psychiatry*, 172, 11–18.

Wong, S. & Gordon, A. (1996). *Violence Risk Scale – experimental version 1 (VRS)*. Saskatchewan, Canada: Regional Psychiatric Centre, Solicitor General of Canada.

*Woods, P. & Reed, V. (1999). The Behavioural Status Index (BSI): some preliminary reliability studies. *International Journal of Psychiatric Nursing Research*, 5, 554–561.

*Young, S. L. & Bullock, W. A. (2003). *The Mental Health Recovery Measure*. Toledo, OH: University of Toledo, Department of Psychology.

APPENDIX A Description of Forensic Outcome Measure Instruments

Instrument	Date	Country	License / Cost	Training Required	Number of items	Description
Atascadero Skills Profile (ASP)	2001	United States	Public Domain	No	10	The <i>Atascadero Skills Profile</i> was developed for use with forensic psychiatric inpatients. The ASP provides an assessment of functioning across 10 domains considered relevant to post-discharge success, namely: self-management of psychiatric symptoms through behavior; self-management of psychiatric symptoms through medication; substance abuse prevention skills; self-management of assaultive behavior; control of self-injurious or suicidal behavior; self-care; independent living skills; control of deviant sexual impulses and behaviors; interpersonal skills; and leisure and recreation skills. The ASP is a repeatable measure, which assists in monitoring patient response to treatment over the course of hospitalization. Quantitative ratings are made via Likert scales of varying length, based on patient behavior over a 90 day period.
Behavioural Status Index (BSI/BEST-index)	1999	United Kingdom	No information available	Yes	70	The <i>Behavioural Status Index</i> is a behaviorally-focused instrument designed to assess patient needs based on the theoretical construct of <i>Social Risk</i> . Social Risk is defined by the authors as the extent to which deficits in physical or mental functioning, activity, participation, and insight predispose individuals to difficulties in social adjustment and aggressive or irresponsible / criminal behavior. The BSI is a repeatable measure, consisting of 70 items encompassing the broad domains of Risk, Insight, Communication Skills, Social Skills, and Self / Family Care. Items produce quantitative scores from 1 (worst) to 5 (best).
Brief Jail Mental Health Screen (BJMHS)	2004	United States	Public Domain	No	8	The <i>Brief Jail Mental Health Screen</i> is a brief non-repeatable tool designed for administration by correctional or health staff to identify prison entrants who may require psychiatric assessment. Consisting of eight 'yes/no' questions, the BJMHS takes 2–3 minutes to complete and requires minimal training. Six questions focus on current experience of mental disorders and two enquire about previous hospitalization and use of psychiatric medication. Scores of two or higher, or answering 'yes' to either psychiatric history question indicates need for further psychiatric evaluation.

(Continued on next page)

APPENDIX A
Description of Forensic Outcome Measure Instruments (Continued)

Instrument	Date	Country	License / Cost	Training Required	Number of items	Description
Brief Psychiatric Rating Scale (BPRS)	1962 1993	United States	Public Domain	Yes	18 & 24 item versions Interview time: 20 – 30 mins	The <i>Brief Psychiatric Rating Scale</i> was designed to assist with assessment of individuals experiencing psychiatric disorders. It provides a quantitative profile of symptom characteristics and enables the change in severity and type of symptoms to be monitored over time (Fitzpatrick et al., 2010). Ratings are based on a semi-structured interview, with anchor points for each symptom enabling the presence and severity to be assessed. Since publication, the BPRS has been modified to improve reliability and usability through development of clear anchor points and a manual based rating system (Greenwood & Burt, 2001). Although originally developed for use in mainstream psychiatry, emerging literature supports the use of the BPRS with a forensic psychiatric population (Corrado, Cohen, Hart & Roesch, 2000; Greenwood & Burt, 2001). The BPRS is a repeatable measure and has demonstrated sensitivity to detecting change over time (Ventura et al., 1993).
Camberwell Assessment of Need - Forensic Version (CANFOR)	2003	United Kingdom	Cost associated with initial purchase of manuals	Yes	25 (Staff and service user ratings)	The <i>Camberwell Assessment of Need - Forensic Version</i> is a repeatable needs assessment for use with individuals experiencing mental health problems in contact with forensic services. Containing 25 items, the CANFOR covers a broad range of needs domains, including: basic life skills, mental health difficulties, functioning, substance use, safety to self and others, interpersonal needs, and offending issues. The CANFOR captures the views of service users, carers and staff for each domain (Thomas et al., 2008). Ratings are based on the consumer's experience over the previous month, where 0 = no problem, 1 = need is present but currently being met, or 2 = need is present and currently unmet.
Cardinal Needs Schedule (CARDINAL)	1995	United Kingdom	No information available	Yes	Interview time: 60–75 minutes	The <i>Cardinal Needs Schedule</i> was designed to measure the psychiatric and social care needs of patients with psychiatric disorders living in the community. Based on standardized interviews, the CNS evaluates 15 areas of functioning, including mental and physical health, living and interpersonal skills, and behavior. A series of criteria are employed to identify problems requiring immediate action. The instrument permits systematic evaluation of both consumer and carer views.
Criminal Justice - Client Evaluation of Self and Treatment (CJ-CEST)	2005	United States	Public Domain	Minimal	80 / 115 item versions available	The <i>Criminal Justice - Client Evaluation of Self and Treatment</i> tool is a consumer focused self-report measure, composed of 15 quantitative scales across three broad domains: Treatment Motivation, Psychosocial Functioning, and Treatment Engagement. Designed as a repeatable measure, the CJ-CEST can be administered throughout the course of an intervention to capture data relevant to treatment planning and outcome measurement. Developed originally as a community-based measure for use in the treatment of substance use, the CJ-CEST has been adapted for use with offender populations by modifying the language used (e.g., Garner et al., 2007). The CJ-CEST is freely available at http://www.ibr.tcu.edu . This tool does not directly measure mental health difficulties.
Defendant and Offender	2006	United States	Public Domain	Minimal	67 items	The <i>Defendant and Offender Screening Tool</i> is a self-report measure, intended to screen individuals entering the

(Continued on next page)

APPENDIX A
Description of Forensic Outcome Measure Instruments (Continued)

Instrument	Date	Country	License / Cost	Training Required	Number of items	Description
Screening Tool (DOST)						criminal justice system for serious mental illness, cognitive impairment and aggressiveness. Five domains are assessed: <i>Social Desirability</i> ; <i>Malingering</i> ; <i>Cognitive Impairment</i> ; <i>Psychosis</i> ; <i>Aggressiveness</i> . This tool was not designed for repeated use and would require further testing if it were to be used in this manner (personal communication, C. Ferguson, May 2011).
The Dundrum Quartet (DUNDRUM)	2010	Ireland	Public Domain	Yes	4 scales (TS = 11 TU = 6 PC = 7 RI = 5)	The <i>Dundrum Quartet</i> contains four structured professional judgement scales designed to assess individuals with psychiatric needs within the forensic system. Scales 1 and 2 (<i>Triage Security</i> [TS] and <i>Triage Urgency</i> [TU]) provide pre-admission assessments to assist in decision making regarding both the placement of consumers within a service and the urgency with which they require treatment. Scales 3 and 4 (<i>Programme Completion</i> [PC] and <i>Recovery Items</i> [RI]) monitor patient change and readiness for transfer to a less secure or community setting. The Dundrum tools are quantifiable and were designed to be used in concert with other structured professional judgment risk assessment tools (i.e., HCR-20).
Forensic Inpatient Observation Scale (FIOS)	2001	Netherlands	Public Domain	No	35	The <i>Forensic Inpatient Observation Scale</i> is a 35-item tool, developed for use with forensic psychiatric inpatients. It assesses functioning and behavior across six domains: self-care, social behavior, oppositional behavior, insight regarding offence, verbal skills, and distress. This tool does not focus extensively on psychiatric symptoms, but does include antisocial personality features (e.g., oppositional behavior and attitudes to offending). The FIOS is a repeatable measure, designed to monitor patient response to treatment over the course of hospitalization. Quantitative ratings are made via 5-point Likert scales (1 = never to 5 = always), based on patient behavior over a three week period.
Historical, Clinical Risk (HCR-20)	1995	United States	Ongoing licensing costs associated with use	Yes	20	<i>Historical, Clinical Risk - 20</i> is a structured professional judgment risk assessment tool, designed to assist clinicians evaluate twenty risk factors associated with violence (10 historical, 5 current clinical, and 5 future risk management). While the 10 historical factors are static in nature, the clinical and risk management factors have shown utility for monitoring change in risk over time (Douglas & Reeves, 2009; Douglas, Blanchard, Guy, Reeves & Weir, 2010). Items are rated on a 0 – 2 scale, where 0 = absent, 1 = partially present, and 2 = definitely present.
Health of the Nation Outcome Scale for Users of Secure Services (HoNOS-Secure)	2007	United Kingdom	Public Domain	Minimal	19 items (12 clinical & 7 security items)	The <i>Health of the Nation Outcome Scale for Users of Secure Services</i> is a member of the HoNOS family of tools, adapted specifically for use in forensic settings. Developed as a repeatable, quantitative measure, the HoNOS-secure assists with tracking a consumer's overall clinical progress and need for secure care or risk management procedures over time (instrument and guides available at: http://www.rcpsych.ac.uk/researchandtrainingunit/honos/secure.aspx). Non-forensic versions of this tool are included in the NOCC suite of measures within Australia.
Illness Management and Recovery Scales (IMR)	2004	United States	Public Domain	Minimal	15 items (client and clinician versions)	The <i>Illness Management and Recovery Scales</i> were developed as a repeatable measure that quantifies a patient's progress towards management of their illness and achieving treatment goals. Consisting of both consumer and clinician versions (15 items each), the IMR utilizes a recovery focused approach to assess: personal goals, social supports, substance use, functioning, medication adherence, coping skills and participation in meaningful activities.

(Continued on next page)

APPENDIX A
Description of Forensic Outcome Measure Instruments (*Continued*)

Instrument	Date	Country	License / Cost	Training Required	Number of items	Description
Level of Service / Case Management Inventory (LS/CMI)	2004	United States	Ongoing licensing costs associated with use	Yes	Section One = 43 items Interview time 20 – 30 mins.	The <i>Level of Service/Case Management Inventory</i> was designed to measure risk and need factors associated with general recidivism. Based on the Level of Service Inventory – Revised (Andrews & Bonta, 1995), section one of the LS/CMI provides an overall quantitative assessment of risk. The remaining 10 sections are used for the purpose of case co-ordination, risk management and treatment planning. Administration of the LS/CMI requires that the user is either familiar with psychometric testing or have undertaken specific training in this tool.
Level of Service Inventory-Revised: Screening Version (LSI-R:SV)	1998	United States	Ongoing licensing costs associated with use	Minimal	8	The <i>Level of Service Inventory-Revised: Screening Version</i> was developed as a brief screening version of the Level of Service Inventory – Revised (Andrews & Bonta, 1995). Like the full version, the LSI-R:SV provides a quantitative assessment of risks and needs associated with empirically derived risk factors for general recidivism: criminal history, criminal attitudes, criminal associates, and antisocial personality pattern. In addition, the LSI-R:SV samples the domains of employment, family, and substance abuse. Although brief, it has shown utility in treatment planning and predicting antisocial behaviour or recidivism during admission and upon release (Andrews & Bonta, 2006). The LSI-R:SV was not designed to assess psychiatric difficulties.
Mental Health Recovery Measure (MHRM)	2003	United States	Public Domain	No	30	The <i>Mental Health Recovery Measure</i> is a self-report measure designed to assess the recovery process of individuals with serious mental illness. The MHRM draws on a recovery focused approach to measure the domains of <i>Overcoming stuckness</i> , <i>Self-empowerment</i> , <i>Learning</i> , <i>Basic functioning</i> , <i>Overall wellbeing</i> , and <i>New potentials</i> (Bullock, 2005). The MHRM produces a quantitative assessment of recovery and is validated for use in forensic settings.
Positive and Negative Syndrome Scale (PANSS)	2000	Canada	Cost associated with purchase of manuals and materials	Yes	30	The <i>Positive and Negative Syndrome Scale</i> is a 30-item scale used to evaluate the presence, absence and severity of Positive (7 items), Negative (7 items) and General Psychopathology symptoms (16 items) of schizophrenia. Each item is rated on a 7-point scale (1 = absent; 7 = extreme), based on information obtained via a semi-structured interview (SCI-PANSS).
Problem Identification Checklist (PIC-R)	2000	Unknown	No information available	Unknown	57	The <i>Problem Identification Checklist</i> was developed as a quantitative measure of dynamic risk for antisocial conduct (e.g., violence or absconding) in forensic psychiatric populations. Rated via semi-structured clinical interview or file review, the PIC-R assesses an individual across the domains Social Withdrawal, Psychotic Symptoms, Inappropriate/ Procriminal Behaviours, Mood Problems, Social Withdrawal, Dynamic Antisocial Factors.
Short Term Assessment of Risk and Treatability (START)	2004	Canada	Cost associated with initial purchase of manuals No ongoing licensing fees	Yes	22	The <i>Short Term Assessment of Risk and Treatability</i> was developed for use with psychiatric and forensic-psychiatric patients. It is a structured professional judgement tool designed to identify a range of risks over a short period of time (i.e., one month). Consisting of 22 items, each item is rated on a three point scale (0,1,2) as being either a strength or risk for the consumer: Social Skills, Relationships, Occupation, Recreation, Self-Care, Mental State, Emotional State, Substance Use, Impulse Control, External Triggers, Social Supports, Material Resources, Attitudes, Medication Compliance, Rule Adherence, Conduct, Insight, Plans, Coping, and Treatability. Items are used to guide ratings on 7 risk domains (violence, self-harm, suicide, unauthorised leave, substance abuse, self-neglect, being victimised). The START was designed as a repeatable measure for use with inpatient and community based clients.

APPENDIX B
Psychometric Properties of the CANFOR, DUNDRUM Quartet, HoNOS-Secure, IMR, MHRM, and START

Instrument		Psychometric properties
CANFOR	General Notes	Validity and reliability studies have been conducted with the CANFOR in medium-secure and high-secure psychiatric facilities in the United Kingdom (Thomas, 2001), Australia (Thomas, Slade, McCrone, Harty, Parrott, Thornicroft, Leese, 2008) and Spain (Romeva, Rubio, Güerre, Miravet, Cáceres, Thomas, 2010). The CANFOR has also been used in community-forensic and prison-inpatient services. This tool has shown adequate reliability and validity in all settings and has shown utility in assisting with service planning (Long, Webster, Waine, Motala & Hollin, 2008).
	Internal consistency	No information has been published regarding the internal consistency of the CANFOR.
	Validity	When investigated with 50 forensic mental health professionals and 60 forensic mental health service users, content validity was found to be satisfactory (Thomas et al., 2008). Concurrent validity was also found to be satisfactory, as evidenced by positive correlations with the Global Assessment of Functioning (APA, 1994) and a five-point GAF needs scale (Thomas et al., 2008). The consumer and clinician versions of the tool have been shown to correlate well with each other (Thomas et al., 2008; Thomas, McCrone, & Fahy, 2009).
	Reliability	Reliability studies were completed with 77 forensic service users and 65 staff in both high and medium secure psychiatric services in the UK. Inter-rater reliability was high for ratings by consumers (0.991) and staff (0.998). Similarly high reliability was found for unmet needs rated by consumers (0.985) and staff (0.972). Test-retest reliability over a two week period was found to be moderately-high for consumers (0.795) and staff (0.852). Similar levels were found for ratings of unmet needs (0.813 and 0.699, respectively; Thomas, Slade, McCrone, Harty, Parrott, Thornicroft, Leese, 2008).
DUNDRUM	Sensitivity to change	The CANFOR's sensitivity to detect change requires further investigation (Segal et al., 2010).
	General Notes	The DUNDRUM QUARTET is currently undergoing a series of evaluations to establish the psychometric properties of these scales (Kennedy, personal communication, April 2011; Dwyer et al., 2011; Flynn et al., 2011). The following information is drawn from both published and pre-publication studies supplied by the authors of this tool. The Dundrum Quartet consists of four separate scales, these have been abbreviated as follows: <i>Triage Security (TS)</i> , <i>Triage Urgency (TU)</i> , <i>Programme Completion (PC)</i> and <i>Recovery Items (RI)</i> .
	Internal consistency	The DUNDRUM scales <i>Triage Security (TS)</i> , <i>Programme Completion (PC)</i> and <i>Recovery Items (RI)</i> have each demonstrated good internal consistency: Cronbach's alpha = 0.949, 0.911 and 0.887, respectively (Flynn et al., 2011; O'Dwyer et al., submitted).
	Validity	The DUNDRUM-TS has demonstrated good predictive validity with two samples of remanded offenders. The TS was able to differentiate between those individuals who would be admitted to hospital or discharged back to prison or community (AUC = 0.893, sensitivity = 0.782, specificity = 0.922). Of those individuals admitted to hospital, the TS scale was employed to classify the level of security to which an individual would be admitted (e.g., open, low-secure, or medium/high-secure ward). The receiving operator characteristics for open-low and low-medium/high classifications yielded an AUC of 0.805 and 0.866, respectively. Item to outcome correlations were significant for all 11 items (0.270 - 0.874; Flynn et al., 2011). The DUNDRUM-PC and -RI scales demonstrated good concurrent validity, as evidenced by significant ($p = 0.001$) and moderate correlations with a range of measures designed to assess dynamic risk (HCR-20 & S-RAMM), patient needs (CANFOR) and psychiatric symptoms (PANSS) (Dwyer et al., 2011).
HoNOS-Secure	Reliability	The DUNDRUM-TS has shown good inter-rater reliability, with Kappa values greater than 0.85 ($p < 0.001$) for 7 of the 11 items, and Spearman's rank correlation coefficients greater than 0.75 ($p < 0.001$) for all items (Flynn et al., 2011). Likewise the DUNDRUM-PC and RI demonstrated Kappa values ranging from 0.44 to 0.77 (all $p < 0.001$) and Spearman rank correlation coefficients greater than 0.51 (range 0.51 to 0.95, all $p < 0.02$; Dwyer et al., submitted).
	Sensitivity to change	The DUNDRUM-TS and TU were not designed for use as a repeated measure (Kennedy et al., 2010). The sensitivity to change of the DUNDRUM PC and RI scales have not yet been tested.
	General Notes	Developed from the Working Age Adults version of the Health of the Nation Outcome Scales (Wing et al., 1998), the HoNOS-secure benefits from a substantial literature base which has established its parent tool as being a valid and reliable instrument (e.g., Pirkis, Burgess, Kirk, Dodson & Coombs, 2005). The HoNOS-secure consists of two scales, namely a Security scale and a modified version of the original HoNOS scale.
	Internal consistency	The HoNOS-secure has demonstrated acceptable internal consistency, with Cronbach's alphas of 0.73 for the security scale and 0.79 for the HoNOS scale (Dickens, Sugarman & Walker, 2007).
HoNOS-Secure	Validity	The HoNOS-Secure has demonstrated a broad degree of consensus with ratings obtained by the CANFOR. Specifically, significant positive correlations have been shown between HoNOS-Secure total scores and the staff rated CANFOR total score ($r = 0.40$, $p < 0.05$), total met needs ($r = 0.39$, $p < 0.05$) and total unmet needs ($r = 0.24$, $p < 0.05$) (Segal, Daffern, Thomas, Ferguson, 2010).

(Continued on next page)

APPENDIX B
Psychometric Properties of the CANFOR, DUNDRUM Quartet, HoNOS-Secure, IMR, MHRM, and START (Continued)

Instrument		Psychometric properties
IMR		The HoNOS-secure has also demonstrated adequate convergent validity with the Working Age Adult version of this tool (Sugarman & Everest, 1999). As such, it may be possible to infer validity of the HoNOS scale from the body of literature regarding the parent tool (Long, Dickens, Sugarman, Craig, Mochty, & Hollin, 2010). However, the authors of the HoNOS-Secure caution that further work is needed to fully establish the validity of the Security scale (Dickens, Sugarman, Picchioni, Long, 2010; Sugarman, Picchioni, Long, 2010).
	Reliability	The HoNOS-Secure has demonstrated good inter-rater reliability for both the Security and HoNOS scales. The Security scale demonstrated Kappa values greater than 0.53 for 6 out of 7 items and the HoNOS scale produced Kappa values greater than 0.65 for 8 out of 12 items. However, the authors of the HoNOS-Secure note it is necessary to further establish test – retest reliability of this tool (Sugarman, Walker, Dickens, 2009).
	Sensitivity to change	The HoNOS-Secure has demonstrated adequate ability to detect clinical change over time (Long et al., 2010; Sugarman, Walker, Dickens, 2009).
	General Notes	The IMR was recently reviewed by both Burgess et al. (2010) and Campbell-Orde et al. (2005). As no new literature has emerged since the publication of these reviews, the following results are largely analogous to those previously reported.
	Internal consistency	The client and clinician versions of the IMR have both demonstrated good internal consistency (Cronbach's alpha = 0.68 – 0.72 and 0.71 – 0.81, respectively; Mueser et al., 2004; Ohio Department of Mental Health, 2004).
	Validity	The concurrent validity of the IMR was established through demonstrating significant positive correlations with the Recovery Assessment Scale ($r = 0.54, p < 0.01$) and Colorado Symptom Inventory ($r = -0.38, p < 0.01$) (Campbell-Orde et al., 2005). The client and clinician versions of the IMR correlate well with each other (Burgess et al., 2010).
MHRM	Reliability	The test-retest reliability of the client and clinician versions of the IMR were found to be both good over a two week period: $r = 0.81 – 0.82$ and $r = 0.78 – 0.81$, respectively (Mueser et al., 2004; Ohio Department of Mental Health, 2004).
	Sensitivity to change	The IMR's sensitivity to detect change has not yet been established (Burgess et al., 2010).
	Internal consistency	The MHRM has demonstrated good internal consistency, with a Cronbach's alpha of 0.93 (Bullock, 2005; Bullock & Young, 2003).
	Validity	The MHRM has demonstrated concurrent validity with a series of other recovery focused tools, namely: the Empowerment Scale ($r = 0.67$), Conner-Davidson Resilience Scale ($r = 0.73$), the Resilience Scale ($r = 0.75$) and the Community Living Scale ($r = 0.57$) (Bullock, 2005). The MHRM has also demonstrated an ability to discriminate between groups of individuals at different levels of recovery, based on the level of participation in treatment or recovery programming (Bullock, 2005; Bullock, Wuttke, Klein, Bechtoldt, & Martin, 2002; Bullock & Young, 2003).
START	Reliability	The MHRM has demonstrated good test-retest reliability over one and two week periods ($r = 0.92, r = 0.91$, respectively) (Bullock, 2005).
	Sensitivity to change	The MHRM has demonstrated the ability to detect change for individuals at the completion of a program designed to promote recovery (Bullock, O'Rourke, Farrer, Breedlove, Smith, & Claggett, 2005; Bullock, Sage, Hupp, O'Rourke, & Smith, 2009).
	General Notes	The START was originally published without extensive validation by the authors. However, since that time, a number of validation studies have been published and a sound literature base has been established supporting the use of this tool in both mainstream and forensic psychiatric settings.
	Internal consistency	The START has demonstrated good internal consistency, with Cronbach's alphas of 0.85 (Nonstad, Nasset, Kroppan, Pedersen, Nøttestad, Almvik et al., 2010) and 0.86 (Nicholls, Brink, Desmarais, Webster & Martin, 2006) for both the strengths and vulnerabilities scales. Item homogeneity, measured using the mean interitem correlation method, produced acceptable MIC's that were greater than 0.20 (Nicholls et al., 2006).
	Validity	Construct validity was demonstrated by Desmarais, Nicholls and Brink (2008), who noted that total scores on the strengths scale increased significantly and vulnerability total scores decreased significantly as an individual's security level decreased ($F(2,286) \geq 31.38, p < 0.001$). The START has demonstrated good levels of predictive validity towards a range of challenging behaviors in a civil psychiatric hospital (AUCs of 0.64 to 0.78; Braithwaite, Charette, Crocker & Reyes, 2010) and two forensic inpatient samples (AUCs of 0.65 - 0.77 and 0.73 – 0.74; Nicholls, Brink, Desmarais, Webster & Martin, 2006; Wilson, Desmarais, Nicholls & Brink, 2010). Higher mean strength scores were obtained by consumers who remained incident free during a 90 day follow-up period ($p = 0.001$; Nicholls et al., 2006).

(Continued on next page)

APPENDIX B
Psychometric Properties of the CANFOR, DUNDRUM Quartet, HoNOS-Secure, IMR, MHRM, and START (Continued)

Instrument	Psychometric properties
	More recently, Nonstad and colleagues (2010) have demonstrated that both the strengths and vulnerabilities scales of the START significantly predicted violence over a 90 day period: strengths (AUC = 0.77), vulnerabilities (AUC = 0.77). A Spearman's correlation between global rating of risk and the occurrence of severe violence within 90 days was positive and significant ($Rho = 0.32$, $p = 0.030$).
Reliability	The START has demonstrated good inter-rater reliability when utilized by clinicians from a range of disciplines, with a Kappa value of 0.87, $p < 0.001$ (Nicholls, Brink, Desmarais, Webster & Martin, 2006).
Sensitivity to change	The authors of the START note that this tool can be administered on multiple occasions to evaluate change (Webster, Martin, Brink, Nicholls & Middleton, 2004). This assertion was supported in recent studies by both Nonstad and colleagues (2010) and Wilson and colleagues (2010), who demonstrated that client ratings changed in the expected direction over time.

2.7 Addendum to article one: Supplementary information regarding the DUNDRUM Quartet.

Since the publication of article one (“A review and analysis of routine outcome measures for forensic mental health services”), several studies have emerged in the literature which provide further evidence to support the use of this tool in forensic mental health settings. Davoren and colleagues (Davoren, Abidin, Naughton, Gibbons, Nulty, Wright & Kennedy, 2013) investigated the predictive ability of the DUNDRUM-3 (*programme completion*) and DUNDRUM-4 (*recovery*) scales to determine the likelihood of a forensic mental health patient being granted conditional discharge from a secure facility. The results of the study demonstrated that the content of the DUNDRUM-3, showed sound predictive validity with regard to decision making regarding discharge from a forensic hospital. Moreover, a more recent examination of the DUNDRUM-3 and DUNDRUM-4 scales by the same team found that these tools demonstrated sound interrater reliability and adequate concordance between clinician and patient ratings of their recovery process (Davoren Hennessy, Conway, Marrinan, Gill & Kennedy, 2015). The authors also found that clinician ratings on the DUNDRUM-4 were significant predictors of conditional discharge. In contrast, the DUNDRUM-1 (*triage security*) scale was found to more strongly predict movement of patients between higher/lower levels of therapeutic security within the hospital environment. Finally, Abidin and colleagues (Abidin, Davoren, Naughton, Gibbons, Nulty & Kennedy, 2013) demonstrated that the DUNDRUM-3 and DUNDRUM-4 were able to assess both reduced and increased risk of violence and self-harm in mentally disordered patients residing in a secure setting. In addition, each of the studies provided evidence that the four scales of the DUNDRUM quartet demonstrated internal consistency.

Given the above, it is therefore acknowledged that the findings of the article one are no longer considered complete; as the conclusions drawn within this review do account for this body of work. It is now considered that the DUNDRUM quartet has indeed demonstrated *sensitivity to detect change* and adequate *internal consistency*, and can be considered to assess aspects of a patient's *recovery process* in a forensic context. On the basis of these findings, inclusion of the DUNDRUM scales within the current project would now be supported by the literature. However, as the above information was not available at the time of developing and embarking upon the current project (see section 2.3 of this chapter), the inclusion of these scales did not occur.

The impact of ongoing developments in the field of forensic outcome measurement and this implications of this for the present study are explored further in the integrated discussion chapter of the thesis (Chapter 9).

PART B: RESEARCH METHODS

Chapter Three: The current study

3.1 Overview of Chapter Three

This chapter describes the rationale, aims and hypotheses of the present thesis. It concludes by outlining specific research questions to be explored by this body of work.

3.2 Overall Approach to Research Methods

This thesis comprises three empirical papers examining elements of outcome measurement in forensic mental health. While specific details of the research methodology pertaining to each paper are presented in chapters six to eight, the following chapter will provide an overview of the overarching research methods used in the three empirical studies that comprise the body of this work.

3.3 Rationale and Overview of Study Aims

As described in chapter one, the use of outcome measurement tools to monitor and track change for mental health patients has become an integral part of clinical practice across Australia (Shinkfield & Ogloff, 2014). More recently, the use of such tools has extended into the forensic mental health domain. However, as was demonstrated in chapter two, until recently there have been a paucity of tools available for use specifically with this population. Moreover, due to the differences between civil and forensic mental health populations, it has been unclear whether the extant tools developed for use with civil

populations were suitable for use with forensic patients. As such, the overarching aims of the present thesis were to evaluate the outcome measures currently mandated for use in Australian mental health services, and to examine how these tools perform within a forensic mental health context. Moreover, the thesis aimed to identify a collection of forensic appropriate outcome measures, and evaluate these against the currently mandated tools; to determine which might perform best in the assessment, monitoring and prediction of future outcomes for forensic psychiatric patients.

The empirical components of this thesis were conducted within Thomas Embling Hospital; the sole forensic mental health hospital in Victoria, Australia. The study setting is described in greater detail in chapter four.

3.4 Aims of the Research

3.4.1 Objective One: Review and analysis of the outcome measurement literature to identify tools designed to monitor risk and clinical/social functioning of individuals receiving forensic mental health treatment.

Of the ROM tools that were designed and validated for use in a forensic environment which currently exist, there appears to be little consensus regarding which measures are most appropriate for use (Chambers et al., 2009). Concerns regarding the applicability, reliability and sensitivity of several measures have also been raised (Dickens, Sugarman, Walker, 2007; Vess, 2001). Thus, the first aim of the present thesis was to conduct a survey of the literature, to identify and review existing measures of patient functioning, which potentially could be applied in an Australian forensic mental health context. The concept of ‘functioning’ in this context was taken to include the broad domains of

clinical/psychosocial functioning, recovery, risk, and placement pathway. This review also aimed to identify measures that that could assist in the assessment, treatment planning, monitoring and prediction of outcomes for forensic psychiatric patients nationally. The review of forensic outcome measures was initially commissioned and funded by the Forensic Mental Health Information Development Expert Advisory Panel (FMHIDEAP), which is a transnational committee (i.e., Australia and New Zealand) established to provide direction for the future development of forensic mental health information in Australia.

The methodology, results and analysis pertaining to this review were presented in their entirety in chapter two of this thesis. As such, they have not been repeated here.

3.4.2 Objective Two: Audit of compliance, precision and reliability of outcome measures used in the Thomas Embling Hospital

The second objective of the thesis was to critically evaluate the existing clinician rated NOCC tools which were mandated for use in both civil and forensic mental health, for their reliability and precision in a forensic mental health context. Ostensibly, this component of the thesis was designed to explore the question of whether or not these ROMs were suitable for this task. Moreover, questions were explored regarding whether or not these tools were being used in a manner consistent with NOCC protocols, as well as the administration manuals of each tool.

Regardless of which outcome measures might be used within a mental health setting, if such tools were not completed in a valid and reliable manner, the information generated will be of little clinical relevance (Groth-Marnat, 2003). At best, spurious information would not be useful for informing treatment or monitoring patient progress. However, at

worst, poorly conducted assessments may suggest a course of treatment that could exacerbate a patient's needs, both in relation to their offending behaviour and mental health difficulties (Bonta & Andrews, 2012). As such, objective two of the present study sought to investigate the accuracy and timeliness of ratings made by clinical staff when using the current suite of NOCC outcome measures, as well as to evaluate compliance with MH-NOCC protocols and to examine the extent to which these tools were being completed in a reliable and valid manner.

To meet these aims, an empirical study was conducted in which two sets of currently mandated ROM tools (HoNOS and LSP-16) were audited for all patients residing within the Thomas Embling Hospital on 1st July 2010. Data obtained were compared against the MH-NOCC protocols (Department of Health and Ageing, 2003a) to evaluate the timeliness, completeness of data and adherence to the MH-NOCC procedures. Descriptive statistics were generated to identify the frequency with which the data conformed to the NOCC protocols. To evaluate the precision with which ROMs had been completed within this sample, inter-observer agreement was examined and the degree of correspondence between the ratings were reported as Kappa statistics for each ROM. Whilst developing the assessment protocol for this study, it was also anticipated that this protocol may provide a useful means of monitoring mental health clinicians' use of ROM tools and for providing corrective feedback to staff when difficulties were observed.

3.4.3 Objective Three: Evaluation of forensic measures against existing NOCC measures (predictive validity, reliability, useability / utility).

Within mainstream mental health services, the focus of treatment is focused largely upon reduction of symptoms and increasing psychosocial functioning (e.g., Cohen &

Eastman, 1997; 2000). As such, measuring patient outcomes has been a relatively straightforward process, with broad measures that evaluate these factors showing efficacy for this task (Gilbody et al., 2003; Jacobs, 2009; Pirkis et al., 2005). However, in forensic mental health settings, where patients can frequently remain in secure care even in the relative absence of psychiatric symptoms, it has been proposed that these tools may be less able to monitor the broader range of needs that constitute patient progress. In this context, patient needs encompass both mental health difficulties, as well as forensic issues pertaining to the risks that such clients pose to themselves and others. Therefore, the third objective of the thesis aimed to determine whether the existing NOCC measures were effective in capturing the range of symptoms and other treatment needs experienced by forensic patients, and moreover whether these existing tools perform better than, or at least equal to, a range of alternative “forensic based” measures.

While progress for forensic patients can be operationalised in a variety of ways, within the present thesis a number of markers were identified as analogues of progress/outcome. These were: movement between units (acute vs rehabilitation wards), episodes of risk behaviour (aggression, self-harm and substance use), and freedom of movement within the hospital setting. These markers were consistent with those used by other researchers (e.g., Davoren et al., 2013; Abidin et al., 2013) and have been described in greater detail in chapter four of this thesis.

To evaluate the relative strengths of the NOCC versus forensic specific ROM tools, two empirical studies were conducted. These have been presented in detail in research papers three and four. Paper three sought to explore whether the needs of forensic patients as a population were heterogeneous, as has been purported in the literature (Cohen &

Eastman, 1997; Keulen-de Vos & Schepers, 2016), and moreover how these needs might change over the course of an admission to a forensic mental health facility. Previous research has identified that forensic mental health patients differ significantly from their mainstream counterparts with respect to the length of time they remain within an inpatient environment (Sharma et al., 2015; Turner & Salter, 2008). Given that admission length for forensic mental health patients is far greater than for civil patients, it might be anticipated that the needs of forensic patients are more likely to change over the course of their treatment (Ruffles, 2010; Turner & Salter, 2008). Research paper three also sought to investigate the question of which set of ROM tools were best able to accurately classify the needs of forensic mental health patients. That is, whether it was possible to differentiate between groups of forensic patients on the basis of their scores on each of these different tools.

Research paper four focused specifically on one ROM tool, namely the Health of the Nation Outcome Scales (HoNOS; Wing et al., 1998), and compared this directly to the forensic version of this tool known as the HoNOS-Secure (Sugarman et al., 2009). The HoNOS is a widely used tool for monitoring consumer outcomes within mental health services. However, questions about its suitability for use in forensic mental health settings led to the development of the HoNOS-Secure. To date, no direct comparisons of these versions of this tool have appeared in the empirical literature. As such, research paper four sought to evaluate the degree to which the HoNOS and the “clinical and social functioning scale” of the HoNOS-Secure correlate with each other (see Chapter 8). Moreover, it sought to evaluate whether the HoNOS or HoNOS-Secure demonstrates better predictive validity on factors pertinent to a forensic mental health population, including mental health functioning, risk and security needs. Finally, research paper four also utilised the HoNOS

to investigate whether differences were able to be identified between civil and forensic mental health populations, on the basis of the HoNOS scores obtained by each sample.

3.5 Research Questions

There were several key research questions explored by the current research:

3.5.1 Research Questions for Objective One.

- 1 What ROM tools exist that have been developed for use in a forensic mental health context?
- 2 If forensic ROM tools exist, what domains of functioning does each capture?
- 3 If forensic ROM tools exist, would they be considered to meet the criteria for inclusion in the NOCC suite of measures in Australia?

3.5.2 Research Questions for Objective Two.

- 4 How frequently do the outcome measures completed at Thomas Embling Hospital comply with the NOCC protocol (National Mental Health Working Group, 2003)?
- 5 To what extent do the NOCC suite of measures demonstrate inter-rater reliability when used by mental health clinicians in a forensic mental health setting?
- 6 Could the protocol developed for this thesis provide a useful means of monitoring mental health clinicians' use of ROMs; specifically regarding compliance with rating protocols and identifying difficulties in a clinician's use of these tools?

3.5.3 Research Questions for Objective Three.

- 7 To what extent do the “forensic specific” ROMs capture additional treatment needs that are pertinent to patients in a forensic setting (i.e., risk of violence/self-harm, or need for containment in a secure setting). Do these tools identify treatment needs that are not captured by the NOCC tools alone?
- 8 To what extent do the forensic specific outcome measures and/or the currently mandated NOCC correlate with the real life outcomes of patients. Which set of tools provides the most useful metric of patient progress in this environment?
- 9 In evaluating the strengths and weaknesses of both the forensic and non-forensic ROMs, which measure or combination of measures would prove most useful in identifying treatment needs and ‘real life outcomes’ of forensic psychiatric patients?

3.5.4 Summary Questions.

- 10 Would there be demonstrable benefit from employing an alternative set of ROMs in forensic mental health settings over those that are currently mandated for use by the Australian government?

3.6 Hypotheses

A number of hypotheses were developed in relation to these research questions:

3.6.1 Hypotheses for objective one.

- 1 It was hypothesised that new forensic focused mental health outcome measures tools would have appeared in the literature during the two decades since the NOCC suite of ROMs was developed.
- 2 That the majority of forensic specific ROM tools would focus on the evaluation and monitoring of risk, rather than mental health and general functioning.
- 3 Based on previous reviews of outcome measurement tools being evaluated for inclusion in the Australian NOCC suite of measures, it was hypothesised that a small proportion of the tools identified from this review would meet NOCC inclusion criteria.

3.6.2 Hypotheses for objective two.

- 4 That analysis of NOCC outcome measures completed within TEH would reveal reporting rates to be below the 85% compliance target set by the Department of Health and Ageing (2009).
- 5 Based on the findings of previous research, (e.g., Pirkis et al., 2005) it was hypothesised that a high degree of inter-rater reliability would be observed between clinician and auditor ratings in the research sample.
- 6 That the use of a standardised protocol may assist in the training or provision of feedback for clinicians using these tools as part of their routine clinical practice.

3.6.3 Hypotheses for objective three.

- 7 That tools developed specifically for use with a forensic population would better differentiate between forensic mental health patients at different stages of progress towards recovery and discharge than tools developed for use in civil mental health settings.
- 8 That for forensic mental health patients, clinical/social needs would be most prominent at the point of admission, with forensic/security needs becoming the primary treatment focus towards discharge.
- 9 That when the needs of forensic mental health patients were examined collectively, distinct groups of patients would be identified amongst this cohort.

Chapter Four: Method

4.1 Overview of Chapter Four

The current chapter describes the methodology employed within the thesis to investigate each of the aims outlined in chapter three. The literature that was reviewed to guide the development of this study and to support the rationale for the approach selected is also discussed. Finally, the development of data collection instruments and the statistical analyses employed are presented. This chapter is divided into three main sections, each of which pertain to the three empirical papers generated by the thesis.

4.2 Overview of Method

The research questions outlined in chapter three were addressed in three separate, but related, empirical studies. The first used existing data generated by patients residing with the Thomas Embling Hospital. These data took the form of clinical notes and ROMs that had previously been completed as part of routine clinical practice. The remaining two empirical studies utilised a prospective data collection approach to examine the utility of a range of ROMs with a forensic mental health population.

4.3 Study Setting

Each of the studies were conducted at the Thomas Embling hospital. Thomas Embling hospital is the sole forensic mental health inpatient facility within the state of Victoria, Australia. Victoria is the southern-most state in mainland Australia with a population of just over 6.0 million people, approximately 4.4 million of whom live in the

capital city, Melbourne (Australian Bureau of Statistics, 2016). Thomas Embling hospital provides secure care for up to 116 patients across seven wards. The wards are structured to encompass the spectrum of patient recovery from acute care to community reintegration. The physical setting of Thomas Embling hospital is located on 8.4 hectares and comprises a campus design with seven accommodation units, education and recreational facilities (Victorian Institute of Forensic Mental Health, 2007). The acute and subacute units are separated from the main campus via a secure airlock. As such, patients housed in these units are able to access the campus only with the assistance of staff. Access to the main campus is determined by a patient's treating team, following assessment of mental health acuity and level of risk towards self and others. Patients residing in the rehabilitation / community reintegration units are able to access the main compound freely.

All patients within the hospital are detained under involuntary treatment orders, which are broadly separated into two main categories: forensic patients, who have been found either unfit to stand trial or not guilty of an offence on the grounds of mental impairment; and security patients, who are prisoners requiring assessment or treatment for mental health disorder. A small proportion of patients are also detained under civil involuntary hospitalisation orders. This group includes a small number of civil patients who require hospitalisation, as well as security patients whose sentences have expired; but whom still require secure psychiatric hospitalisation.

4.4 Legislative Context

Detention of patients at the Thomas Embling hospital is provided for and governed by several pieces of legislation, which straddle aspects of both mental health and criminal law. These include the *Crimes (Mental Impairment and Unfitness to be Tried) Act, 1997*;

Mental Health Act (Vic), 2014; and *Crimes Act, 1958*. Of particular relevance to the present study is the *Crimes (Mental Impairment and Unfitness to be Tried) Act 1997 (CMIA)*, as this provides the grounds upon which a person may be detained as a forensic patient when found not guilty of an offence due to mental impairment. Thomas Embling hospital, and its governing organisation the Victorian Institute of Forensic Mental Health (Forensicare), is legally responsible for the management of all forensic patients in Victoria.

The defence of mental impairment is outlined in Section 20 of the CMIA. In order for a defence of mental impairment to be established, it must be proven that at the time of committing the offence in question, the person had been suffering from a mental impairment which had the effect that:

- a) the person did not know the nature and quality of the conduct, or*
- b) the person did not know that the conduct was wrong (that is, he or she could not reason with a moderate degree of sense and composure about whether the conduct, as perceived by reasonable people, was wrong)*

(Crimes (Mental Impairment and Unfitness to be Tried) Act, 1997)

While the criteria required to be found not guilty by reason of mental impairment are provided within the CMIA, mental impairment is not explicitly defined in the Act (O'Donahoo & Simmonds, 2016). Rather, the state of Victoria has adopted a variation of the British M'Naghten Rules as the standard test for criminal liability for mentally disordered defendants. As such, mental illness, intellectual disability, and conditions such as cognitive impairment may all fall within the scope of this defence (O'Donahoo & Simmonds, 2016). However, a recent review of mental impairment cases in Victoria has revealed that in practice, non-psychotic mental illnesses do not readily form the basis for a

successful defence of mental impairment (Wondemaghen, 2014). Indeed, since the introduction of the CMIA, all cases in which mental impairment has succeeded as a defence have concerned offenders who had been psychotic at the time of committing their offence (Victorian Law Reform Commission, 2003; Wondemaghen, 2014). Moreover, serious personality disorder and psychopathy are not considered to fall within the scope of this defence (Mullen & Ogloff, 2009). Despite this, there is high incidence of comorbidity between mental illness, personality disorder and cognitive impairment within this population (e.g., Hayward & Moran, 2008; Ogloff et al., 2015). As such, the patient population of Thomas Embling hospital frequently presents with multiple and complex needs across many of these domains.

4.5 Procedure

The following section describes the methodology used to achieve the objectives of this thesis. Additional information pertaining to each objective have also been presented in each of the empirical papers presented in chapters six, seven and eight.

4.6 Objective One: Review and Analysis of the Outcome Measurement Literature

Objective one of the thesis sought to review and analyse the existing outcome measurement literature, to identify tools designed for the task of monitoring the risk and clinical/social functioning of individuals receiving forensic mental health treatment. The methodology employed in conducting the review and analysis was described in its entirety in the published manuscript presented in chapter two. As such, a description of this procedure has not been repeated here.

4.7 Objective Two: Audit of Compliance, Precision and Reliability of Current Outcome Measures Used in the Thomas Embling Hospital

Objective two sought to investigate the accuracy and timeliness of ratings made by clinical staff when using the currently mandated suite of NOCC outcome measures. It also sought to evaluate staff compliance with NOCC protocols and to examine the extent to which these tools were being completed in a reliable and valid manner. To achieve this, two ROM tools were selected for evaluation, namely the HoNOS and LSP-16. These tools were selected as they capture a broad range of clinical and functional impairment, based on clinical evaluation of a standardised set of criteria outlined in their respective treatment manuals.

As objective two sought to understand the way in which these tools were currently being completed, it was determined that it would be most useful for the present study to audit outcomes data that had previously been completed by clinical staff. That is, to evaluate data collected retrospectively, rather than capturing new data in a prospective manner. It was considered that if new data were obtained, the act of collecting data might influence clinician behaviour and impact on the manner in which these tools were being used (i.e., Hawthorne-observer effect; see Monahan & Fisher, 2010).

To achieve the aims of objective two, the first three sets of outcome measures completed for each patient residing within Thomas Embling Hospital were examined. As such, outcome measures that had been completed for patients on admission and at 91-day and 182-day reviews were included for analysis (as per MH-NOCC protocol). As many patients had resided in the hospital for periods longer than 182 days, limiting data

collection to the first three measurement periods ensured that ratings made at similar points in a patient's admission were evaluated.

The audit of ROMs was undertaken by a team of eight mental health nurses at the request of the principal investigator and student researcher. This team performed the task under supervision and guidance of a senior nurse from the clinical administration team of the hospital (Grade 5 Registered Psychiatric Nurse) and the student researcher. Members of the auditing team were selected for their clinical expertise, as well as their familiarity with the ROM tools and assessment protocols.

The audit of previously completed NOCC ROMs commenced on 1st July 2010 and was completed over the course of one month. Access to patient files was facilitated by the Health Information Manager and team of ward clerks throughout Thomas Embling Hospital. To determine which patient files were to be audited, a list of all patients who were residing in Thomas Embling hospital on 1st July 2010 was provided by the Health Information Manager, along with patient identifier numbers (UR numbers) and location in which their relevant files were stored. Sets of patient files were delivered to the auditing team as required and were stored securely by the ward clerks when not in use.

4.7.1 Data collection tool – audit protocol.

To standardise the process of data collection, an audit tool and protocol manual was developed. The data collection tool developed for objective two is presented in Appendix G. In addition, a protocol manual was developed to provide detailed instructions for the auditing team to guide them in conducting the audit. This manual has been presented in Appendix F.

In addition to collecting data that were directly pertinent to the objectives of the present study, other aspects of ROM completion that were of interest to the study setting were also included. Whilst these data are not discussed in the present thesis, they were used to inform quality improvement initiatives within the hospital. Thus, the data collection protocol (Appendix G) contains several domains in addition to those relating specifically to the objectives of the present thesis of monitoring compliance with NOCC protocols and evaluating the reliability and validity of the tool.

Prior to any data being collected, the student researcher met with all members of the auditing team and provided training in the correct data collection procedure. Several example assessments were conducted in collaboration with the auditing team to ensure all members of the team understood the protocol and were able to complete the task in an accurate and consistent manner.

Details of specific data collection components for objective two have been detailed in the following sections.

4.7.2 Adherence to NOCC protocols.

In the first instance, the study protocol required auditors to record details of when and how each outcome measure had been completed. This included procedural information such as whether each measure had been completed within the expected timeframe, as well as the frequency with which items had been omitted. Patient admission dates were used to calculate the timeframes within which each set of ROMs should have been completed. These data were then compared against the NOCC protocols (Department of Health and Ageing, 2003a) to evaluate the timeliness, completeness of data and adherence to the MH-

NOCC procedures. Descriptive statistics were generated to identify the frequency with which the data conformed to NOCC protocols.

4.7.3 Precision of ratings.

To evaluate the degree of precision with which ROMs had been completed within this sample, auditors identified one set of ROMs per patient for further investigation and analysis. The method used for determining which set of outcome measures would be selected was standardised as follows. Where data from the 91-day review were available, this was selected in the first instance. If data from the 91-day review were unavailable, but the 182-day review was present, this was selected as the second preference. However, if neither a 91- or 182-day review was available, then data collected during the admission period was used. This procedure was specified on the understanding that a ROM completed during a review period would be informed by a longer period of clinical observation and greater familiarity with the patient than those completed within the first two weeks of admission.

Having identified one set of outcome measures to be investigated further, the date upon which those tools had been completed was noted, and patient records (i.e., clinical file notes) written during the two weeks preceding this date were then reviewed. A two week review period was selected, as this is the rating period specified in the HoNOS user manual (Wing et al., 1998). Based on the information contained in the patient file, a senior nurse then independently re-rated the ROMs without reference to the ratings previously provided by the original treating nurse. Scores generated by the treating clinician and the auditing nurse were then evaluated for inter-observer agreement. The degree of

correspondence between the two sets of ratings were reported as Kappa statistics for each item.

Kappa statistics were selected for this purpose as they provide a quantitative measure of the degree to which two or more observers agree on the presence or absence of a factor being evaluated. Kappa statistics are commonly interpreted in the following manner: < 0.0 less than chance agreement, 0.01–0.20 slight agreement, 0.21– 0.40 fair agreement, 0.41– 0.60 moderate agreement, 0.61–0.80 substantial agreement, 0.81–0.99 almost perfect agreement (Cohen, 1960).

4.8 Objective Three: Evaluation of Forensic Measures and NOCC Measures

Objective three sought to investigate a range of questions pertaining to the relative strengths and weakness of the currently mandated NOCC tools, in comparison to measures developed for use specifically with forensic mental health patients.

Objective three was achieved via two empirical studies (study three and four). Study three sought to investigate which ROM tools were best able to classify the needs of forensic mental health patients and to differentiate between groups of forensic patients on the basis of their clinical and risk related needs. Moreover, study three also aimed to investigate the question of whether the treatment needs of forensic patients are heterogeneous (e.g., Cohen & Eastman, 1997; Keulen-de Vos & Schepers, 2016), and whether these treatment needs change over the course of admission.

Study four sought to determine whether the existing NOCC measures were effective in capturing the range of symptoms and needs experienced by forensic patients, and

moreover whether these perform better than or equal to a range of alternative measures developed specifically for use in forensic mental health settings. Finally, study four also directly compared two versions of the HoNOS (i.e., HoNOS and HoNOS-Secure), to determine which performed best with this population.

Both objectives two and three utilised the same data set gathered in Thomas Embling Hospital. Three forensic mental health outcome measurement tools were selected for this purpose. These tools were:

- Health of the Nation Outcomes Scale for Users of Secure and Forensic services (HoNOS-Secure)
- Forensic Camberwell Assessment of Needs Short Version (CANFOR-S)
- Level of Service Inventory – Revised: Screening Version (LSI-R: SV)

The above ROMs were selected on the basis of the findings generated by study one of the thesis, whereby these tools demonstrated a sound ability within the literature to monitor a range of needs possessed by a forensic mental health population, whilst also having adequate psychometric properties (Shinkfield & Ogloff, 2014). The LSI-R:SV was included as validated short risk assessment of general recidivism (Andrews & Bonta, 1998). It was further considered that within the scope of a doctoral thesis it would not be possible to investigate each of the forensic ROM tools identified from the review of the literature (Chapter 2). As such, at the time of commencing research, the HoNOS-Secure and CANFOR tools were considered the most likely candidates to perform this role, with the recommendation provided that future research being given to investigating the START, DUNDRUM quartet, IMR and MHRM.

In addition to the forensic ROMs, data were collected from the currently mandated NOCC tools. These data had been gathered for each patient as part of standard clinical practice during their admission to the study setting. These tools were:

- Health of the Nation Outcomes Scales: working age adults version (HoNOS)
- Life Skills Profile – 16 Item version (LSP-16)
- Behaviour and Symptom Identification Scale (BASIS-32)

Each of the above tools were completed for all patients who consented to participate. Ratings were conducted as per the NOCC protocol on admission, discharge and every 91-days that a patient remained within the study setting. All ratings were undertaken by mental health clinicians (e.g., psychiatric nurses, psychologists, occupational therapists and social workers) who had received training in the use of these tools to increase reliability of ratings (Rock & Preston, 2001). A description of the training provided to staff has been presented in section 4.8.2.

On each rating occasion, the forensic and NOCC measures were rated by separate clinicians (i.e., two clinicians were used at each collection occasion, with one clinician rating the HoNOS and the other rating the HoNOS-Secure), with ratings being based on the patient's presentation over the same period of time (i.e., two weeks) upon which the NOCC tools were based.

Finally, data were also recorded regarding a patient's freedom of movement (restricted/unrestricted access to the campus), ward placement (acute/subacute unit), and number of risk incidents accrued during the two week rating period (aggression, self-harm and substance use). See section 4.8.4 for further details.

4.8.1 Measures and materials.

Based on the findings presented in research article one (see Chapter 2), five outcome measurement tools were selected for further evaluation in the present study, namely: *Health of the Nation Outcome Scales (HoNOS)*, *Life Skills Profile (LSP-16)*, *Behaviour and Symptom Identification Scale (BASIS-32)*, *Health of the Nation Outcome Scale for Users of Secure Services (HoNOS-Sec)*, *Camberwell Assessment of Need - Forensic Version (CANFOR)*. In addition, data regarding risk of general recidivism were identified via the *Level of Service Inventory-Revised: Screening Version (LSI-R:SV)*. These tools are described in the following sections.

4.8.2 NOCC outcome measures

4.8.2.1 *Health of the Nation Outcome Scales (HoNOS; Wing et al., 1998).*

The HoNOS is a 12-item clinician rated measure, designed to monitor four broad areas of clinical and social functioning for people with severe mental illness: behavioural problems, cognitive and physical impairment, symptomatic problems, and social functioning. Ratings are made by comparing item descriptions with a client's presentation over the past two weeks (Wing et al., 1998). Individual items are rated on a 0 – 4 scale, where 0 = *No Problem* and 4 = *Severe to Very Severe Problem*. In a major review of the HoNOS by Pirkis and Colleagues (Pirkis et al., 2005a; Pirkis et al., 2005c) it was found that the HoNOS had undergone extensive investigation of its psychometric properties, with this tool demonstrating sound validity, reliability and utility in mainstream mental health services. This review was later extended by the Te Pou research team (Te Pou, 2012), which provided further support for the earlier findings. The findings of the above reviews have been described in table 3 (Pirkis et al., 2005a; Pirkis et al., 2005c; Te Pou, 2012).

Table 3: Summary of HoNOS psychometric properties (adapted from Pirkis et al., 2005a; Pirkis et al., 2005c; Te Pou, 2012)

Content validity	Construct validity	Criterion validity	Predictive validity	Test-retest reliability	Inter-rater reliability	Sensitivity to change
Good	Good	Good	Good	Adequate	Adequate	Adequate

4.8.2.2 Life Skills Profile (LSP-16; Rosen, Hadzi-Pavlovic, Parker & Trauer, 2006).

The LSP-16 is a clinician-rated instrument comprising 16 items designed to measure four broad domains of social and adaptive functioning: self-care, antisocial behaviour, withdrawal, and compliance with treatment (Rosen, Hadzi-Pavlovic, Parker & Trauer, 2006). The LSP-16 was designed for use with individuals living with schizophrenia and chronic mental illness in the community (Pirkis, Burgess, Kirk, Dodson & Coombs, 2005). A consumer's functioning is rated on each of the 16 items with respect to their behaviour over the preceding three-month period. Items are rated on a 0 – 3 scale, with a variety of different anchor points used depending on the content of each item. Overall, LSP-16 items are rated for the degree to which they have been present over the past three month, between 0 = *Never* to 3 = *Always*. The LSP-16 was developed as an abbreviated version of the LSP-39 (Rosen, Hadzi-Pavlovic & Parker, 1989) and was designed to emphasise the presence of life skills rather than focus on a client's deficits.

The LSP-16 has undergone a range of psychometric evaluations, both by the original authors of the tool, as well as independent research teams. The LSP-16 demonstrates sound validity, inter-rater reliability and utility in mainstream mental health services (Pirkis et al.,

2005; Shinkfield & Ogloff, 2015; Webster, Bretherton, Goulter & Fawcett., 2013). However, two studies (Trauer et al. 1995; Parker et al. 2007) have demonstrated that the communication subscale shows a poor inter-rater reliability and internal consistency (Minichino, Francesconi & Carrión, 2017). As with the HoNOS, the LSP-16 featured in a major review conducted by Pirkis and Colleagues (Pirkis et al., 2005a). The findings of this review have been summarised in Table 4.

Table 4: Summary of LSP-26 psychometric properties (adapted from Pirkis et al., 2005a)

Content validity	Construct validity	Criterion validity	Predictive validity	Test-retest reliability	Inter-rater reliability	Sensitivity to change
Adequate	Adequate	Adequate	Adequate	Good	Good	Good

4.8.2.3 Behaviour and Symptom Identification Scale (BASIS-32; Eisen, Dill & Grob, 1994).

The BASIS-32 is a 32-item behavioural health assessment tool designed to monitor changes in a client's self-reported symptom and functional difficulties. The 32 items assess a wide range of symptoms and problems across five domains of mental health functioning and substance abuse: Relation to Self and Others, Depression and Anxiety, Daily Living and Role Functioning, Impulsive and Addictive Behaviour, and Psychosis (Eisen et al., 1994). Items are rated on a five-point scale, where 0 = No Difficulty and 4 = Extreme Difficulty. The BASIS-32 has been shown to have adequate validity and reliability, and to be sensitive to change during treatment (Pirks et al., 2005).

Table 5: Summary of BASIS-32 psychometric properties (adapted from Pirkis et al., 2005a)

Content validity	Construct validity	Criterion validity	Predictive validity	Test-retest reliability	Inter-rater reliability	Sensitivity to change
Adequate	Good	Good	Good	Good	Good	Adequate

4.8.3 Forensic outcome measures

4.8.3.1 *Health of the Nation Outcome Scale for Users of Secure Services (HoNOS-Secure; Sugarman & Walker, 2007).*

The HoNOS-Secure is a member of the HoNOS family of tools, adapted to provide a means of tracking the clinical, social and security needs of users of secure psychiatric services, prisons and forensic community services (Sugarman et al., 2009). The HoNOS-secure contains the original twelve ‘clinical and social functioning’ items of the HoNOS, modified to account for the environmental conditions typically found in a secure setting (Dickens et al., 2007)². In addition, a seven-item ‘security scale’ monitors changes in a client’s need for risk and security management procedures (Long et al., 2010). As with the HoNOS, the HoNOS-Secure ‘clinical and social functioning scale’ is rated retrospectively; based on a period of two weeks prior to the day on which tool was completed. Whereas, the ‘security scale’ is rated prospectively for the period ‘in the near future’ (Dickens et al., 2007). The HoNOS-Secure instrument and guides are freely available at: <http://www.rcpsych.ac.uk/researchandtrainingunit/honos/secure.aspx>.

² A full description of the wording modifications between the HoNOS and HoNOS-Secure can be found in Appendix L

The HoNOS-secure benefits from a substantial literature base which has established its parent tool as being a valid and reliable instrument (e.g., Pirkis et al., 2005a; Pirkis et al., 2005c; Te Pou, 2012). This tool has demonstrated acceptable internal consistency, convergent validity, inter-rater reliability and adequate sensitivity to change (Shinkfield & Ogloff, 2014; see also Appendix B of Chapter 2). Significant positive correlations have been shown between HoNOS-Secure total scores and the staff rated CANFOR total score ($r = 0.40$, $p < 0.05$), total met needs ($r = 0.39$, $p < 0.05$) and total unmet needs ($r = 0.24$, $p < 0.05$) (Segal et al., 2010). It has also demonstrated adequate convergent validity with the Working Age Adult version of this tool (Sugarman & Everest, 1999). However, the authors of the HoNOS-Secure caution that further work is needed to fully establish the validity of the Security scale (Dickens, Sugarman, Picchioni, Long, 2010).

In addition, the HoNOS-Secure has demonstrated good inter-rater reliability for both the ‘security’ and ‘clinical’ scales. The Security scale has demonstrated Kappa values greater than 0.53 for 6 out of 7 items and the HoNOS scale has produced Kappa values greater than 0.65 for 8 out of 12 items (Sugarman, Walker, Dickens, 2009).

4.8.3.2 Camberwell Assessment of Need - Forensic Version (CANFOR; Thomas et al., 2003).

The CANFOR is a needs assessment tool for use with individuals experiencing mental health problems in contact with forensic services (Thomas et al., 2003). Containing 25 items, the CANFOR covers a broad range of needs domains, including: basic life skills, mental health difficulties, functioning, substance use, safety to self and others, interpersonal needs, and offending issues. The CANFOR captures the views of patients, carers and staff for each domain (Thomas et al., 2008). Ratings are based on the

consumer's experience over the previous month, where 0 = no problem, 1 = need is present but currently being met, or 2 = need is present and currently unmet.

Validity and reliability studies have been conducted with the CANFOR in medium-secure and high-secure forensic psychiatric facilities in the United Kingdom (Thomas et al., 2009), Australia (Thomas et al., 2009) and Spain (Romeva, Rubio, Güerre, Miravet, Cáceres, Thomas, 2010). The CANFOR has also been used in community-forensic and prison inpatient services. This tool has shown adequate reliability and validity in all settings and has shown utility in assisting with service planning (Long, Webster, Waine, Motala & Hollin, 2008). However, The CANFOR's sensitivity to detect change has not been adequately investigated and requires further evaluation (Segal, Daffern, Thomas & Ferguson, 2010).

4.8.3.3 Level of Service Inventory-Revised: Screening Version (LSI-R:SV; Andrews & Bonta, 1998).

The *LSI-R:SV* was developed as a brief version of the Level of Service Inventory – Revised (Andrews & Bonta, 1995). The *LSI-R:SV* provides a well validated and reliable (e.g., Daffern, Ogloff, Ferguson, Thompson, 2005; Ferguson, Ogloff & Thomson, 2009; Lowenkamp, Lovins & Latessa, 2009) quantitative assessment of risks and needs associated with general recidivism: criminal history, criminal attitudes, criminal associates, and antisocial personality pattern. In addition, the *LSI-R:SV* samples the domains of employment, family, and substance abuse. Although brief, it has shown utility in treatment planning and predicting antisocial behaviour or recidivism during admission and upon release (Andrews & Bonta, 2006). The *LSI-R:SV* does not assess psychiatric symptoms or difficulties. Psychometric analyses of the relationship between the *LSI-R* and *LSI-R:SV* has produced correlations of .85 for incarcerated males, .68 for incarcerated females, and

.84 for probationers of both genders (Andrews & Bonta, 2006). A recent meta-analysis of studies examining the predictive validity of the Level of Service Inventory family of tools has demonstrated sound predictive accuracy when used with a range of client groups. Moreover, this tool has previously been evaluated specifically in the forensic mental health setting used by the present study (Ferguson, Ogloff & Thomson, 2009) and has demonstrated LSI-R:SV predicts recidivism at a level that is significantly above chance for any new offence (AUC = .67, $p < .001$), for nonviolent new offences (AUC = .65, $p < .001$), and for violent new offences (AUC = .60, $p < .05$).

4.8.4 Staff training and feedback mechanism

Prior to data collection, all assessing clinicians were required to attend a six hour training session to ensure familiarity with each of the forensic outcome measurement tools. The training package was developed and delivered by the student researcher (a registered psychologist with endorsement in the clinical scope of practice), in conjunction with Professor Stuart Thomas (lead author of the Camberwell Assessment of Needs: Forensic Version) and Professor Ogloff (an authorised master trainer for the Level of Service Inventory measures). The training session included the following components:

- Description of the present study, including aims and rationale
- Background and development of the HoNOS, CANFOR and LSI-R:SV
- Scoring procedure, coding rules and item descriptions of each tool
- Completion of two case vignettes for each tool (six in total)
- Consensus scoring and corrective feedback provided after each vignette.

A training manual was also developed and provided to each assessing clinician during the training session and an additional copy of this manual was provided to each hospital ward for reference. The participant manual contained a copy of all information discussed during the training, as well as the relevant sections from the assessment manuals of each tool. In total, 22 clinical staff members³ completed the training and participated as assessing clinicians for the present study.⁴

Following the initial training session, staff were provided additional support by the student researcher throughout the course of the study. Support was provided via face to face meetings, as well as regular email and telephone contact. In addition, a monthly summary and feedback email was sent to all staff participants. This email acknowledged and thanked staff for ongoing participation in the study and provided an update on the number of assessments completed. Regular contact with staff served to facilitate ongoing collection of data over the course of the study period, whilst also acknowledging the ongoing efforts of participating staff with the competing demands of regular clinical work. Finally, an electronic repository was developed on the Thomas Embling Hospital intranet site, in which all study information could be accessed by participants as required. \

³ e.g., registered psychiatric nurses, psychologists, occupational therapists, social workers and psychiatric registrars.

⁴ The participant manual developed to assist in the completion of this component of the study has not been reproduced within this thesis. This is due to both the length of the document and also due to it containing copyrighted material pertaining to each of the forensic ROM tools.

4.8.5 Patient consent

To obtain consent from potential patient participants, the student researcher attended community meetings on each hospital ward. Community meetings are a regular forum in which all patients and staff meet to discuss ward issues and share information regarding the hospital and individual wards. During these meetings, a brief description of the study was presented to the patients and any questions raised by potential participants were answered. In addition, a written explanatory statement (see appendix A) was distributed to all patients during this meeting. Following this, the student researcher had no further contact with potential participants regarding this study. Nursing staff on each ward subsequently discussed the study with interested patients and provided any assistance required to read the explanatory statement and consent form (see appendix B) to ensure they understood the nature and extent of participation.

Following discussion with their nurse, patients who wished to participate in the study by permitting their outcomes data to be assessed and analysed were assisted to complete the consent form. Completed forms were delivered to the student researcher via secure internal mail and were retained in a locked filing cabinet on university premises. As data for this study were obtained via clinician conducted assessments and from information obtained in clinical files and electronic databases, there was no ongoing burden on the patient sample to actively participate in data collection once consent has been provided.

4.8.6 Data collection tool

To facilitate data collection by the team of assessing clinicians, a data collection tool was developed for this project. The data collection tool has been presented in Appendix H.

In essence, the tool provided a structured means by which assessment data could be recorded in a manner that facilitated storage and data entry.

Once completed by an assessing clinician, the student researcher reviewed the relevant patient clinical file and recorded the following additional information:

- *Ward acuity*: Whether the patient was residing on an acute, subacute or rehabilitation/community reintegration ward at the time the ROM was collected.
- *Freedom of movement*: Whether or not the patient had been granted freedom to access the main campus of the hospital (i.e., outside of a secure ward setting) during the time that the ROM was collected.
- *Risk Incidents*: Data pertaining to any risk incidents that occurred during the two weeks preceding the collection of ROMs. Risk incidents were recorded in relation to aggression, self-harm and substance use.
- *NOCC ROM data*: The NOCC data collection sheets that were completed during the same period as the Forensic ROMs were also obtained.

4.8.7 Web decision support tools: Web reports portal

To obtain data pertaining to the average Victorian state-wide HoNOS scores for empirical paper three, data were also accessed from the Web Decision Support Tools (wDST) via the Australian Mental Health Outcomes and Classification Network's website (<http://wdst.amhocn.org/>). Within Australia, collection of ROMs by mental health services is supported by a nationwide system for reporting and analysis of outcomes data. This

system provides public access to aggregated data submitted by each state and territory, and enables the data set to be freely interrogated with regard to a variety of high level descriptors (e.g., age, gender, legal status). The wDST enables users to generate summary data regarding patient samples at a state/territory or national level. No identifiable data pertaining to an individual person or service is able to be obtained from the wDST. Rather, this functions as a means of generating high level data with which an individual or service can compare their scores against groups of people with similar demographic and casemix variables (Burgess et al., 2015).

4.9 Study Three: Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population

4.9.1 Classification of needs by NOCC and forensic based ROM tools

To investigate the first aim of objective three, data generated by the six outcome measures (i.e., HoNOS, HoNOS-Secure, LSP-16, LSI-R:SV, CANFOR, BASIS-32) were evaluated using analysis of variance (ANOVA) to identify whether significant differences were present in the scores obtained by the forensic population at different levels of ward acuity. The alpha for all tests was set at 0.05. Ward placement was used as a measure of mental health acuity, with three levels being specified: acute, sub-acute and rehabilitation/community reintegration wards. Significant effects were further examined via Scheffé post hoc comparisons to ascertain where differences occurred between the different levels of acuity.

4.9.2 Heterogeneity of needs amongst forensic mental health patients

To investigate whether variation exists within the clinical, security and risk related needs within the forensic mental health population, the ROM scores for each member of the sample were examined and their clinical and forensic/security needs were classified as being either in the high or low range (see figure one). The HoNOS and HoNOS-Secure ‘Security Scale’ were employed as a measure of clinical/functional (HoNOS) and security/forensic (HoNOS-Secure ‘security scale’) needs. These tools were selected as they demonstrated the greatest ability to differentiate between patients at each level of ward placement/acuity. This finding emerged during the initial ANOVA analysis (see article three located in Chapter 7 for details). Cut-off scores for ‘high’ versus ‘low’ clinical needs were determined by identifying the median score obtained on both measures. Whilst other methods of identifying cut-off scores are available, due to the relatively small sample size available, using the median split was considered the most appropriate means determining the point to divide low and high scores on these tools.

On the basis of the median scores, it was determined that a HoNOS value of 0 – 5 would be considered low, with a score of six or greater being indicative of a high level of clinical needs. Likewise, for security/forensic needs it was determined that a score on the HoNOS-Secure ‘security scale’ of 0 – 6 would be considered in the low range, with a value of seven or greater indicating a high level of security/forensic need. HoNOS and HoNOS-Secure scores were interrogated for each member of the sample population and the number of patients meeting criteria for each of the four categories was quantified (i.e., high or low scores for clinical and security needs; see figure one). This procedure was completed for all three levels of ward acuity, as well as for the population as a whole.

		CLINICAL/SOCIAL NEEDS	
		High	Low
RISK / FORENSIC NEEDS	High	High security needs/risk of offending, with acute psychiatric needs	High security needs/risk of offending, without prominent psychiatric needs
	Low	Acute psychiatric needs, but low-risk of offending or security needs	No prominent psychiatric needs, with low risk of offending / security needs

Figure 1: Schematic showing the four categories of high/low clinical and forensic needs, with description of each domain.

4.10 Study Four: Comparison of the HoNOS and HoNOS-Secure Within the Thomas Embling Hospital

4.10.1 Comparison of mean HoNOS scores between forensic and civil mental health patients

Mean HoNOS scores for all mental health patients within the state of Victoria were accessed via the wDST on the Australian Mental Health Outcomes and Classification Network website (<http://wdst.amhocn.org/>). The reference criteria used to generate these data were: Jurisdiction: *Victoria*, Age Group: *Adult*, Service Setting: *Inpatient*, Financial Year: *July 2010 – June 2011*. *Level of Analysis* was specified as *Collection Occasion*, to permit comparison of data collected on admission, 91-day review and discharge; as well as a global average across all collection occasions. It was observed that the forensic cohort within TEH was skewed heavily towards male patients (85.7%). It was therefore thought

that the female component of the sample might not be representative of female mental health patients generally. As such, the data obtained from the wDST were further restricted to *male patients*, and only the male portion of the forensic sample was used. Likewise, as the HoNOS was designed for use with ‘working age adults’ the sample was restricted to patients aged *18 – 65 years*. The remaining variables of diagnosis and legal status were set to ‘*All*’. Data obtained via the wDST are described in terms of sample size, mean scores, and standard deviation. Comparison of mean scores generated by the civil and forensic samples was undertaken using two-tailed t-tests. To investigate the effect size of any difference observed between the two means, Cohen’s d statistics were generated post-hoc.

4.10.2 Correlation of HoNOS and HoNOS-Secure total score and items

To investigate the degree to which the HoNOS and HoNOS-secure (clinical and social functioning scale) overlap, Pearson correlations were generated for item pairs between these two scales. This was undertaken using data generated from the forensic mental health sample. Cohen’s d statistics were generated post-hoc to further evaluate any difference observed.

4.10.3 Predictive ability of HoNOS and HoNOS-secure

To investigate whether the HoNOS-Secure performs better than or equal to the HoNOS, in terms of its predictive validity within forensic mental health settings – a series of logistic regression analyses were performed. Three dependent variables were used as markers of mental health acuity and risk: ward placement (i.e., whether the participant resided on an acute or sub-acute unit during the period of review), freedom of movement status (i.e., whether the participant had restricted or unrestricted access to the hospital

campus), and risk incidents (i.e., occurrence of aggression, self-harm or substance use). Similar markers of mental health acuity and risk have been utilised to good effect in previous research studies of a similar nature (e.g., Davoren et al., 2013; Abidin et al., 2013).

In all cases, each of the three HoNOS components (i.e., the original HoNOS and the ‘clinical and social functioning’ and ‘security’ scales of the HoNOS-secure) were employed as independent variables and were entered together as one block into the regression analysis. Standardised beta weights for each scale were examined to determine their relative contribution to the classification of patients on the dependent variables. This analysis was subsequently repeated with data obtained from a second sample of patients, collected three years after the initial sample. Finally, a post-hoc investigation was undertaken, in which the HoNOS-secure ‘security scale’ was combined with the HoNOS and a further regression analysis was conducted using the HoNOS-Secure (clinical and security scales), as well as the HoNOS with ‘security scales’ added. This was undertaken to directly compare the performance of the HoNOS/HoNOS-secure if the security scale were added to either version of this tool.

4.11 Data Collection

Data collection for empirical papers three and four occurred in two phases, with the initial phase occurring between 1 July 2010 and 1 January 2011. To evaluate the stability of findings over time, a second period of data collection occurred between 1st December 2014 and 1st May 2015.

4.12 Data Coding Protocols

Data from each component for the study were entered by the student researcher into a Microsoft Excel spreadsheet (Microsoft Office Professional Plus 2013, version 15.0.4859.1000; Microsoft Corporation). Upon completion of data collection, data were imported into the Statistical Package for Social Sciences for Windows (version 20, SPSS, Inc., Chicago, IL, USA). Prior to data analysis, all variables were manually examined and underwent basic data cleaning. A randomised sample containing 10% of all data sets were checked for accuracy of data entry. Of those data sets checked, none were found to contain transcription errors.

4.13 Ethical Approval

The studies contained within this thesis received ethics approval from the Monash University (Appendix C) and Swinburne University of Technology (Appendix D) human research ethics committees. A letter of permission was also received from the Victorian Institute of Forensic Mental Health (Forensicare) to permit collection and use of data pertaining to patients within their service (Appendix E).

Several ethical considerations specific to research using forensic mental health patients were raised and considered. Specifically, issues of informed consent and the application of privacy principles within a forensic mental health setting were relevant.

PART C: EMPIRICAL STUDIES

Chapter Five: Overview of the Empirical Papers

To this point, this thesis has reviewed the extant literature pertaining to outcome measurement in mental health, outlined the aims and research questions and described the research methods used to meet these aims. Part C of this thesis presents the empirical studies that were undertaken to answer these research questions. Each of the studies has been prepared for publication in peer reviewed journals; they will thus be presented in manuscript form. However, the pages have been re-numbered for consistency within the thesis. There are three papers, each addressing one or more of the research aims articulated in chapter three.

Paper two (presented in Chapter 6) sought to examine the accuracy with which forensic mental health clinicians were able to interpret the existing NOCC routine outcome measure items in a forensic psychiatric setting. It also sought to evaluate the degree of compliance demonstrated by clinical staff with local assessment procedures. Moreover, the study sought to examine the precision with which ratings were being conducted with these tools within a forensic mental health environment. Finally, the audit protocol employed in this study was itself appraised as a method of monitoring mental health nurses' use of routine outcome measures and providing feedback in this regard.

Paper three (presented in Chapter 7) examined whether it was possible to differentiate amongst groups of forensic patients on the basis of their scores on a sample of ROM tools. As such, the paper investigates whether the needs of forensic mental health patients were able to be better classified by forensic or non-forensic ROM tools. The second aim of the study sought to explore whether the needs of forensic patients were

indeed heterogeneous and whether their needs were subject to change over the course of admission.

The final paper (presented in Chapter 8) directly compared two versions of the Health of the Nation Outcome Scales in a forensic mental health setting (i.e., the HoNOS and HoNOS-Secure). In the first instance, differences between the HoNOS scores obtained by civil and forensic mental health populations were investigated. The second aim of paper four was to evaluate the degree to which the HoNOS and the “clinical and social functioning scale” of the HoNOS-Secure correlate with each other. That is, to what extent do these two version of the HoNOS tool overlap or demonstrate differences in the way they are interpreted in such settings. This was investigated both at the item and total score level. Finally, paper three sought to evaluate whether the HoNOS or HoNOS-Secure demonstrates better predictive validity with respect to domains such as risk and security needs possessed by forensic mental health patients.

Chapter Six: Use and Interpretation of Routine Outcome Measures in Forensic Mental Health

6.1 Overview of Chapter Six

The following chapter introduces the first empirical study of this thesis.

6.2 Preamble to Published Paper: “Use and interpretation of routine outcome measures in forensic mental health”

The second publication in this thesis aimed to both examine the precision of ratings made with these tools within a forensic mental health environment and to pilot a method of monitoring mental health nurses’ use of routine outcome measures. The audit protocol was found to be effective in evaluating both the accuracy with which nurses were able to interpret routine outcome measure items and their degree of compliance with local procedures for completing such instruments. Moreover, the results suggest that despite these routine outcome measures having been developed for use in general mental health settings, they could also be interpreted and rated with an adequate degree of reliability in a forensic mental health context. However, difficulties were observed in the applicability of several components of these tools within a forensic environment. Recommendations for future research and implications for practice are discussed.

The study upon which article two is based was conducted between 2010 and 2011. Therefore, it is acknowledged that the findings of this study reflect the state of clinical practice during that period of time.

The following article was published in the *International Journal of Mental Health Nursing*. This is a peer-reviewed journal of the Australian College of Mental Health Nurses (ISSN 1445-8330 [Print], 1447-0349 [Online]), which has been published since 1992 and now is published four times per year. In 2015, the *International Journal of Mental Health Nursing* had an impact factor of 1.95.

6.3 Authorship Indication Form: Chapter Six**Swinburne Research**

Authorship Indication Form

For PhD (including associated papers) candidates

NOTE

This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each associated 'paper'. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each associated paper to be included in your thesis.

DECLARATION

We hereby declare our contribution to the publication of the 'paper' entitled:

Use and interpretation of routine outcome measures in forensic mental health

First Author

Name: Gregg Shinkfield

Signature: _____

Percentage of contribution: 85%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Reviewed literature, obtained ethics approval, collected and coded data, conducted analysis, prepared and revised manuscript.*

Second Author

Name: Professor James Ogloff

Signature: _____

Percentage of contribution: 15%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Assisted with conceptualisation of study and manuscript, revised manuscript.*

Principal Coordinating Supervisor: Professor James Ogloff

Signature: _____

Date: _____

6.4 Declaration by Co-authors

The undersigned hereby certify that:

- a) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
- b) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- c) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- d) there are no other authors of the publication according to these criteria;
- e) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- f) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s):

Centre for Forensic Behavioural Science,
Swinburne University of Technology and Forensicare
505 Hoddle Street, Clifton Hill
Victoria, Australia

Professor J. Ogloff:		Date:
----------------------	--	-------

6.5 Published Paper Two: “Use and Interpretation of Routine Outcome Measures in Forensic Mental Health”



FEATURE ARTICLE

Use and interpretation of routine outcome measures in forensic mental health

Gregg Shinkfield^{1,2} and James Ogloff^{1,2}

¹Centre for Forensic Behavioural Science, Swinburne University and ²Victorian Institute of Forensic Mental Health (Forensicare), Melbourne, Victoria, Australia

ABSTRACT: *The present study aimed to both pilot a method of monitoring mental health nurses' use of routine outcome measures (ROM) and to examine the precision of ratings made with these tools within a forensic mental health environment. The audit protocol used in the present study was found to be effective in evaluating both the accuracy with which nurses were able to interpret ROM items and their degree of adherence with local procedures for completing such instruments. Moreover, the results suggest that despite these ROM having been developed for use in general mental health settings, they could be interpreted and rated with an adequate degree of reliability by nurses in a forensic mental health context. However, difficulties were observed in the applicability of several components of these tools within a forensic environment. Recommendations for future research and implications for practice are discussed.*

KEY WORDS: *forensic, Health of the Nation Outcome Scales, Life Skills Profile-16, mental health, outcome measure.*

The use of routine outcome measures (ROM) within mental health services is now common practice internationally (Trauer 2010). As broad measures of clinical functioning and mental health status, such tools offer a means of systematically identifying and monitoring the needs of patients, while also providing a common platform for the development of treatment plans (Pirkis *et al.* 2005a). Within Australia, the mandatory use of ROM has been an integral component of the National Mental Health Strategy since the early 1990s (Australian Health Ministers 1993). The ubiquitous application of ROM across mental health services in Australia has enabled longitudinal tracking of information regarding the needs and outcomes of individual patients, irrespective of the mental health services with which they have had contact.

Since the introduction of ROM across Australia in 2003, all public sector mental health services now routinely collect and report outcomes data (Burgess *et al.* 2012). The tools used for this task include a 12-item clinician-rated measure of mental health and social functioning known as the Health of the Nation Outcome Scales (HoNOS) (Wing *et al.* 1998). The HoNOS is completed by all mental health services, with variants used for children (Gowers *et al.* 1999) and older adults (Burns *et al.* 1999). Ambulatory (outpatient) services are also required to complete a 16-item measure of psychosocial functioning, known as the Life Skills Profile-16 (LSP-16) (Rosen *et al.* 1989). In addition, all services offer consumers the option of completing one of two self-report measures of symptomatology and functioning, namely the Behaviour and Symptom Identification Scale (Eisen *et al.* 1994) and Kessler 10+ (Kessler *et al.* 2003). The above instruments are required to be completed on admission, discharge, and every 91 days for those consumers who remain in contact with a service (Department of Health and Ageing 2003). Together, these tools form part of the

Correspondence: Gregg Shinkfield, Victorian Institute of Forensic Mental Health, Locked Bag 10, Fairfield, Vic 3078, Australia. Email: gregg.shinkfield@forensicare.vic.gov.au

Gregg Shinkfield, BSc, MSc(Hons), PGDipClinPsych.

James Ogloff, BA, MA, JD, PhD.

Accepted July 2014.

National Outcomes Casemix Collection suite of measures (NOCC) (Pirkis *et al.* 2005a).

A sizeable body of literature suggests that the HoNOS and LSP-16 both perform well within general mental health services with respect to their sensitivity, specificity, and predictive validity, when used accurately and consistently by clinicians (Pirkis *et al.* 2005b; Webster *et al.* 2013). However, the utility of such tools depends largely on the degree to which they are completed in a reliable manner, as poor adherence to assessment protocols negatively impacts upon the validity of the data obtained. To evaluate adherence with mandatory requirements and reporting protocols, governing bodies typically monitor the volume and percentage of completeness of ROM conducted by a service (e.g. Burgess & Coombs 2011; Department of Health and Ageing 2003). However, while such monitoring processes might assist in identifying service level difficulties with the completion of these tools, this does not ensure that the data collected are valid, meaningful, or useful.

The ROM tools currently used in Australia were developed and evaluated in general mental health settings. As such, there has been concern expressed regarding the use of these measures within specialist fields, such as forensic mental health, dual diagnosis (i.e. substance misuse and mental illness), and indigenous mental health (Department of Health and Ageing 2003). When considering the utility of measurement instruments, ensuring that items within the tool have been constructed in a manner that promotes ease of interpretation and reliability of ratings is imperative. Moreover, test items should be applicable and 'make sense' within the environment in which they are used (a concept referred to as 'face validity') (McColl *et al.* 2006). Evaluating the degree to which test users are able to agree on the presence or absence of individual test items (interrater reliability) is often referred to under the concept of 'precision' (Viera & Garrett 2005). However, little research has been conducted regarding the interpretation of either individual items, or ROM as a whole, in specialist health-care settings. In particular, the present study focuses on the use and interpretation of these tools within a forensic mental health environment.

In the present jurisdiction, forensic mental health is defined as the provision of assessment and treatment to individuals who both experience mental health difficulties and whose behaviour has led, or could lead, to offending (Mullen 2000). Legislation governing the provision of forensic mental health services frequently differs across jurisdictions, particularly with respect to grounds for discharge and length of admission of patients. Depending upon the legal requirements within each jurisdiction, a

forensic mental health patient might be required to remain in a treatment facility, even in the absence of psychiatric symptoms, with treatment being focused on issues of risk reduction and other forensic needs. Due to these extra treatment demands, the average length of inpatient care received by forensic consumers is often significantly longer than that provided to their non-forensic peers (Turner & Salter 2008).

While there are many similarities between the social and clinical needs experienced by patients in forensic mental health settings and their counterparts in general mental health (Shaw 2002), several authors have noted that users of forensic mental health services present not only with the mental health difficulties and functional impairments seen in general settings, but also demonstrate a history of criminal behaviour, violent or sexual offending, a high prevalence of comorbid personality disorder, behavioural disturbance, self-harm, and/or substance use (Dickens *et al.* 2007; Ogloff *et al.* 2004). In addition, consideration frequently needs to be given to level of security, level of risk, and risk management, which is required for this client group (Kennedy *et al.* 2010; Shaw 2002). Given the differences identified between consumers of general mental health and forensic mental health environments, it is possible that the ability to accurately complete these tools in such contexts might be limited, with the resulting data being unreliable.

The present study was developed to address two broad aims. Firstly, to pilot an alternative method of monitoring mental health nurses' use of ROM tools; specifically to evaluate adherence with rating protocols and to identify difficulties experienced by nurses in using these tools. Secondly, to examine the level of concordance (i.e. interrater reliability) of ratings made with the NOCC suite of measures when used by mental health nurses in a forensic mental health setting. In doing so, the extent to which the items within these tools are able to be interpreted in a consistent manner within this setting could be investigated.

MATERIALS AND METHODS

Setting and source population

The present study was conducted at Thomas Embling Hospital, the sole forensic mental health inpatient facility within the state of Victoria, Australia. The hospital provides secure care for up to 116 patients across seven wards. The wards are structured to encompass the spectrum of patient recovery from acute care to community reintegration. All patients within the hospital are detained under involuntary treatment orders, broadly separated

into two main categories: (i) forensic patients, who have been found either unfit to stand trial or not guilty of an offence on the grounds of mental impairment; and (ii) security patients, who are prisoners requiring assessment or treatment for mental illness. A small proportion of patients are also detained under civil involuntary hospitalization orders.

Data collection and analysis

To investigate the aims of this study, two ROM tools (HoNOS and LSP-16), which had previously been completed by clinical staff, were examined for all patients residing within Thomas Embling Hospital on 1 July 2010. Consumer-rated measures were excluded from the present study, as these capture a client's subjective view of their treatment needs, and as such, were not amenable to evaluation of their interrater reliability. An audit protocol was developed by the lead author to guide the collection of data by eight mental health nurses, under the supervision and guidance of a senior nurse from the clinical administration team. Members of the auditing team were selected for their clinical expertise, as well as familiarity with the ROM tools and assessment protocols. To standardize data collection, auditors examined the first three sets of outcome measures that had been completed for each patient residing at the hospital during the period of the study. As such, outcome measures that were completed for patients on admission and at the 91- and 182-day reviews were included for analysis. While many patients had resided in the hospital for periods longer than 182 days, limiting data collection to the first three measurement periods was done to ensure that ratings made at similar points in a patient's admission were evaluated. This study received ethics approval from the Monash University Human Research Ethics Committee.

Adherence to NOCC protocols

In the first instance, the study protocol required auditors to record details of when and how each outcome measure was completed. This included procedural information, such as whether each measure had been completed within the expected timeframe, as well as the frequency with which items had been omitted. Patient admission dates were used to calculate the timeframes within which each set of ROM should have been completed. These data were compared against the NOCC protocols (Department of Health and Ageing 2003) to evaluate the timeliness, completeness of data, and adherence to the NOCC procedures. Descriptive statistics were generated to identify the frequency with which the data conformed to the NOCC protocols.

Precision of ratings

Precision, as it pertains to the level of agreement between observers (interrater reliability or interobserver agreement), is often reported using Cohen's kappa statistic (Cohen 1960; Viera & Garrett 2005). Kappa provides a quantitative measure of the degree to which two or more raters agree on the presence or absence of a factor being evaluated. Kappa statistics are commonly interpreted in the following manner: <0.0, less than chance agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement (Cohen 1960).

To evaluate the degree of precision with which ROM had been completed within this sample, auditors identified one set of ROM per patient for further investigation and analysis. Determining which set of outcome measures was to be selected was standardized as follows. Where data from a 91-day review were available, this was selected in the first instance. If 91-day review data were unavailable, but the 182-day review was present, this was selected as the second preference. However, if neither a 91- or 182-day review was available, then data collected during the admission period was used. This procedure was specified on the basis that a ROM completed during a review period would be informed by a longer period of clinical observation and greater familiarity with the patient than those completed within the first 2 weeks of admission.

Having identified the date upon which a selected ROM had been completed, patient records (i.e. clinical file notes) written during the 2 weeks preceding this date were then reviewed. A 2-week review period was selected, as this is the rating period specified in the HoNOS user manual (Wing *et al.* 1998). Based on this file information, a senior nurse then independently rerated the ROM without reference to the ratings that had been previously provided by the original treating nurse. Scores generated by the treating nurse and the auditing nurse were then evaluated for interobserver agreement, and the degree of correspondence between the ratings were reported as kappa statistics for each item.

RESULTS

Outcome measures audited

The files of all patients residing in Thomas Embling Hospital during the period of the study were obtained for auditing ($n = 112$). Of these, 107 contained valid ROM records from the first 182 days of admission. This yielded an overall sampling rate of 95.5% of the patient population (Table 1).

TABLE 1: *Descriptive data of patient files audited*

Files audited	Total (n)	Valid (n)	%	
	112	107	95.5	
No. days admitted	Min	Max	Mean	SD
	19	6800	1434	(1372)
Admission length	n	%		
<3 months	23	21.5		
3–7 months	7	6.5		
>7 months	77	72		

SD, standard deviation.

TABLE 2: *Availability of completed outcome measures forms in clinical files*

	Expected (n)	Observed (n)	%
Admission	107	100	93.5
91-day review	84	69	81.0
182-day review	77	59	76.6
Total	268	228	84.7

The mean length of stay for patients in the sample was 1434 days (range: 19–6800 days). Twenty-one percent ($n = 23$) had resided within the hospital for less than 91 days; 6.5% ($n = 7$) for a period of 90–180 days; 72% ($n = 77$) had been resident for longer than 180 days. Based on these findings, it was extrapolated that 268 sets of ROM should have been completed for this sample (i.e. admission = 107, 91-day review = 84, and 182-day review = 77) (Table 2). The audit revealed that 228 sets of measures (84.7% of anticipated total) had been completed and were available in clinical files. This comprised 100 (93.5%) admission, 69 (82.1%) 91-day review, and 59 (76.6%) 182-day review sets of outcome measures. Moreover, these ROM were distributed across the following tools: HoNOS = 228 and LSP-16 = 127.

Of the 107 patient files examined in this study, the mean length of time required to complete all aspects of the audit was 26.70 min (range = 10–60 min; standard deviation (SD) = 10.98).

Completion of ROM items and reporting requirements

HoNOS missing items

Based on the admission lengths of patients within the study sample, a total of 268 HoNOS evaluations were anticipated. However, a total of 228 (80.5%) were available in patient files (admission = 100, 91-day review = 69, 182-day review = 59; 85%, 80%, and 76% of expected, respectively). Within the 228 sets of HoNOS evaluations

examined, 11 (4.8%) were found to contain missing items or incomplete data. Of these, six omissions (6%) occurred during the admission period, and the remaining five (7.2%) occurred within 91-day reviews. There were no missing items observed within the 182-day review. Of those assessments in which items had been omitted, it was found that a mean of 2.6 items (SD = 2.6) were left incomplete during admission, with one item (SD = 0) missing at the 91-day review.

The most commonly omitted items were ‘problems with living conditions’ (item 11, 5.6%) and ‘problems with occupation and activities’ (item 12, 4.7%). In addition, ‘problems with activities of daily living’ (item 10, 1%) and ‘other mental and behavioural problems’ (item 8, 1%) were also omitted to a lesser extent. No other items had been omitted.

LSP-16 missing items

While the NOCC protocol mandates that the HoNOS be completed by all services at each collection occasion, the LSP-16 is required only by services providing outpatient care. Although Thomas Embling Hospital provides inpatient residential care, and is therefore not required to collect LSP-16 data, a local protocol had been established for use of this tool during review and discharge assessments. As such, LSP-16 data generated for the 91- and 182-day review periods were available for auditing. Based on the admission lengths of patients within the study sample, a total of 161 LSP-16 evaluations were anticipated (91-day review = 84, 182-day review = 77). However, a total of 127 (78.8%) were available in patient files (91-day review = 68, 182-day review = 59, 80% and 76% of expected, respectively). Of these, five (3.9%) contained missing items or incomplete data. The most commonly omitted items were ‘Does this person generally make and/or keep up friendships?’ (item 8), ‘Does this person generally look after and take her or his own prescribed medication?’ (item 10), ‘Does this person behave irresponsibly?’ (item 15), and ‘What sort of work is this person generally capable of?’ (Item 16). Each of these items was omitted with the same frequency (1.9%).

Precision of ratings (HoNOS and LSP-16)

The kappa values generated for each of the HoNOS items (H1–H12) are reported in Table 3. With the exception of items H4 (cognitive problems) and H8 (other mental and behavioural problems), the interrater agreement for all other items was observed to be in the moderate-to-substantial range (Cohen 1960). The highest level of agreement occurred for items H5 (physical illness/disability), H6 (hallucinations and delusions), H3

TABLE 3: Interrater agreement (HoNOS items)

HoNOS no.	Item description	κ -value
H1	Overactive/aggressive/disruptive behaviour	0.7412
H2	Non-accidental self-injury	0.6834
H3	Problem drinking or drug taking	0.7773
H4	Cognitive problems	0.2543
H5	Physical illness/disability problems	0.8077
H6	Hallucinations and delusions	0.7796
H7	Depressed mood	0.5254
H8	Other mental and behavioural problems	0.4118
H9	Problems with relationships	0.6585
H10	Problems with activities of daily living	0.7711
H11	Problems with living conditions	0.7060
H12	Problems with occupation and activities	0.6534

HoNOS, Health of the Nation Outcome Scales.

TABLE 4: Interrater agreement (LSP-16 items)

LSP-16 no.	Item description	κ -value
L1	Initiating and responding to conversation	0.8793
L2	Withdrawal from social contact	0.7412
L3	Warmth to others	0.7807
L4	Personal grooming	0.7046
L5	Clean clothing	0.7768
L6	Neglect of physical health	0.4825
L7	Violence to others	0.8546
L8	Make/keep friendships	0.7701
L9	Maintenance of adequate diet	0.4881
L10	Compliance with prescribed medication	0.8572
L11	Willingness to take medication	0.8079
L12	Cooperation with health services	0.8343
L13	Problems with others in household	0.7319
L14	Offensive behaviour	0.7844
L15	Irresponsible behaviour	0.7622
L16	Work capability	0.6587

Life Skills Profile-16.

(problem drinking or drug taking), and H1 (overactive/aggressive/disruptive behaviour). This suggests that, with the exception of H4 and H8, clinicians had at least a moderate degree of agreement determining item ratings using this tool.

The kappa values generated for each of the LSP-16 items (L1–L16) are reported in Table 4. The interrater agreement for all LSP-16 items was observed to be between moderate and almost perfect (Cohen 1960). The highest level of agreement occurred for items L1 (initiating and responding to conversation), L10 (compliance with prescribed medication), L7 (violence to others), and L12 (cooperation with health services), which were rated as having almost perfect agreement between raters. The lowest level of agreement was observed on items

L6 (neglect of physical health) and L9 (maintenance of adequate diet); however, this might reflect difficulties in rating these items based on file information alone. Taken together, this suggests that clinicians had at least a moderate degree of agreement when determining item ratings using this tool within a forensic mental health environment, with many ratings attaining high levels of interrater reliability.

DISCUSSION

In the present study, we sought to pilot a method of monitoring mental health nurses' use of ROM tools and to examine the accuracy of ratings generated by the NOCC suite of tools within a forensic mental health environment. Overall, the audit protocol developed for this study was found to be a useful means of evaluating the reliability of nurse-rated ROM assessments. Rather than evaluating the percentage of ROM that have been completed at a service level, as is typically employed to infer adherence with these tools, the present methodology investigated adherence to NOCC protocols (i.e. timeliness of ratings and completeness of data) at not only a service level or clinical unit level (e.g. ward), but also in relation to individual clinicians' handling of these tools. Moreover, it was possible to use this process to evaluate the degree of accuracy with which assessments were conducted, either by individual nurses or by groups of clinicians, with reference to the rating criteria for each tool. This represents a useful process for assisting new staff to complete these tools in an accurate manner, as well as providing a means of periodically evaluating clinicians' ratings to ensure they remain accurate and do not drift over time (Velligan *et al.* 2011). It was also noted that the HoNOS and LSP-16 were able to be reliably rated from file information, when compared to ratings made by clinicians working with patients *in vivo*. The high degree of concordance between the two sets of assessments suggests that the information required to evaluate these tools was largely available in clinical notes, typically recorded during standard nursing practice. While we do not suggest that these tools should be completed without direct clinical observation of a patient, and indeed this would be contrary to the protocols specified in the respective user manual of each tool (Rosen *et al.* 1989; Wing *et al.* 1998), for the purpose of research, training, or supervision, an independent clinician who has not been involved directly with the patient could complete this task. As the average time required to complete the audit was approximately 30 min per patient, this represents an achievable investment of time to ensure the accuracy of ROM data within a service.

To the authors' knowledge, this is the first reported study to investigate HoNOS and LSP-16 ratings in a forensic mental health setting. The findings suggest that despite these tools having been developed for use with a general mental health population, the items within each tool could be interpreted within a forensic mental health environment in a consistent manner. Both the HoNOS and LSP-16 demonstrated at least moderate degrees of agreement between nursing staff, with many ratings attaining high levels of interrater reliability. This finding is consistent with a number of other studies that have demonstrated fair-to-substantial levels of interrater agreement across a variety of inpatient (e.g. Jacobs 2009) and community settings (e.g. Idaiani 2011). As such, it could be suggested that the HoNOS and LSP-16 can perform as well in a forensic mental health environment as they do in general mental health settings. However, the present data also indicated that there are several items within these tools that might be less valid for application in a forensic context.

In particular, it was observed that the HoNOS items most frequently omitted during the early phase of admission were those relating to 'problems with living conditions' (item 11, 5.6%) and 'problems with occupation and activities' (item 12, 4.7%). Both of these items require the clinician to assess the patient's environment and the availability of occupational activities within that environment, particularly with regards to how these factors meet the needs of the individual patient. Within the general adult version of the HoNOS, when evaluating a person residing on an acute hospital ward, these items instruct clinicians to rate the patient's usual accommodation, such as a residential setting or accommodation in the community. However, the population of the present study comprised patients within a forensic environment, for whom prison or another secure environment was often their most recent accommodation and likely discharge destination. As such, these items were frequently not easily interpreted in this context. Moreover, information about the patient's pre-admission environmental conditions might not be readily available to clinicians. It could be suggested that other items within the HoNOS and LSP-16 might also not fully reflect the extent of a client's problematic behaviour, due to limitations in applying item criteria within a secure environment. For example, HoNOS item 3, 'problem drinking or drug taking', focuses on a patient's use of substances during the preceding 2-week period. For most patients within a secure setting, access to substances might be limited by environmental constraints; however, the underlying problem could be demonstrated via cravings for substances, medication-seeking

behaviour, or other markers that are not assessed via the HoNOS.

An alternate version of the HoNOS currently exists, known as the HoNOS-Secure (Sugarman & Walker 2007), which has been adapted for users of secure and forensic services. The item content of the HoNOS-Secure reflects these environmental constraints and seeks to assess the impact of a secure environment upon the patient's functioning. However, the HoNOS-Secure is not currently included in the Australian NOCC suite of measures, yet the findings of the present study suggests that evaluation of this tool in comparison to the general adult version of the HoNOS might be warranted (see Shinkfield & Ogloff, 2014 for a review of the HoNOS-Secure and a broad discussion of other measures relevant to forensic populations).

With respect to adherence with NOCC protocols, it was noted that as the length of a patient's admission increased, the less likely an outcome measure was to have been completed as required (reducing from 93.5% on admission to 76.6% by the 182-day review). This finding was consistent with the overall pattern observed nationally in the collection of ROM, with lower levels of completion observed during review and discharge assessments (Burgess & Coombs 2011). However, despite higher completion rates of ROM on admission, it was found that admission assessments also demonstrated a higher percentage of omitted items than in those completed during review periods. It might be hypothesized that due to clinicians being less familiar with patients on admission, they might struggle to provide informed ratings on several items. In contrast, ratings made at the 91- and 182-day reviews were likely informed by a greater degree of familiarity with the patient and knowledge of their mental health and overall functioning.

Limitations

As is often the case in research, several limitations within the present study should be acknowledged. Most significantly, the methodology for determining the precision of ratings relied on retrospective assessments based on file information. While the results suggest that this did not present a significant impediment in this study, and indeed provides support that ratings can be obtained reliably in this manner, members of the auditing team noted that aspects of the information required for rating several ROM items were not routinely recorded in clinical files. For example, items that demonstrated the lowest degree of interrater agreement included: maintenance of adequate diet (LSP-16 item 6), neglect of physical health (LSP-16 item 9), and cognitive problems (HoNOS

item 4). These factors were frequently not commented on in patient files, unless these they had been identified as a specific area of concern or treatment need. Although it is likely that information regarding these domains would be recorded if a patient were to display such difficulties, in the absence of specific concerns this information was not routinely noted. Moreover, several of these items assess factors that might present as long-term difficulties (e.g. cognitive problems), and as such, impairments in these areas were not regularly described in daily observations unless a change in functioning had occurred. Therefore, the lower level of interrater reliability observed for these items might have resulted from methodological limitations and a lack of relevant information in the patient's clinical record, rather than an inherent problem with the items themselves.

It is also acknowledged that the data upon which this study was based were collected 4 years ago and provides only a cross-sectional view of ROM use within one clinical setting. As such, it might be possible that the findings of this study do not reflect any progress or change in clinical practice that has occurred since that time. Moreover, in the absence of data from other forensic mental health services, it is possible that these findings may not generalize across services.

CONCLUSION

The findings of the present study provide support for the assertion that the items within the HoNOS and LSP-16 are amenable to interpretation in a consistent and reliable manner in a forensic mental health environment. Moreover, the findings showed that it was possible to complete the measures reliably via file review. As such, the protocol employed within this study might prove useful in assisting with research, as well as training and ongoing monitoring of nursing and other clinical staff in their use of these tools by senior nurses or managers. However, due to several inherent differences between forensic and general mental health settings, a number of limitations were identified with the use of the HoNOS and LSP-16 in a forensic mental health environment. Specifically, limitations arose with respect to items that are influenced directly by the environment, such as problems with living conditions, problems with occupation and activities, and problem drinking or drug taking. Moreover, it was noted that these measures do not provide information regarding treatment needs that are specific to a forensic environment, such as risk of harm to others, offending behaviour, and level of security required (see Shinkfield & Ogloff, 2014 for further discussion). Therefore, despite the finding that

clinicians can utilize item criteria in a precise and reliable manner, questions were raised about the validity and utility of the general adult version of the HoNOS in a forensic mental health setting. Further evaluation of these factors appears warranted, and investigation of whether the HoNOS-Secure or another tool of this sort could be effectively substituted in the place of the HoNOS is recommended.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the significant contribution of Monica Summers (senior auditor), as well as the nursing staff of Thomas Embling Hospital and the team of ward clerks who assisted with the collection of data for this project.

REFERENCES

- Australian Health Ministers (1993). *National Mental Health Plan*. Canberra: Commonwealth Department of Health and Family Services.
- Burgess, P. & Coombs, T. (2011). *National Outcomes & Casemix Collection: Volume and Percentage Valid Clinical Ratings*. Brisbane, Queensland: Australian Mental Health Outcomes and Classification Network.
- Burgess, P., Coombs, T., Dickson, R. & Pirkis, J. (2012). Achievements in mental health outcome measurement in Australia: Reflections on progress made by the Australian Mental Health Outcomes and Classification Network (AMHOCN). *International Journal of Mental Health Systems*, 6 (4), 1–12.
- Burns, A., Beevor, A., Lelliott, P. *et al.* (1999). Health of the Nation Outcome Scales for elderly people (HoNOS 65+). Glossary for HoNOS65+ score sheet. *The British Journal of Psychiatry*, 174, 435–438.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Department of Health and Ageing (2003). *Mental Health National Outcomes and Casemix Collection: Technical Specification of State and Territory Reporting Requirements for the Outcomes and Casemix Components of 'Agreed Data'*. Version 1.50. Canberra: Commonwealth Department of Health and Ageing.
- Dickens, G., Sugarman, P. & Walker, L. (2007). HoNOS-secure: A reliable outcome measure for users of secure and forensic mental health services. *Journal of Forensic Psychiatry & Psychology*, 18, 507–514.
- Eisen, S. V., Dill, D. L. & Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry*, 45 (3), 242–247.

- Gowers, S. G., Harrington, R. C., Whitton, A. *et al.* (1999). Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA): Glossary for HoNOSCA score sheet. *The British Journal of Psychiatry*, 174, 428–431.
- Idaiani, S. (2011). Inter-rater reliability of Health of Nations Outcome Scale (HoNOS) among mental health nurses in Aceh. *Health Science Journal of Indonesia*, 2 (2), 53–57.
- Jacobs, R. (2009). *Investigating Patient Outcome Measures in Mental Health*. CHE research paper 48. York: University of York.
- Kennedy, H. G., O'Neill, C., Flynn, G. & Gill, P. (2010). *Dangerousness, Understanding, Recovery and Urgency Manual (the DUNDRUM Quartet) VI.0.22*. Dublin: Trinity College Dublin.
- Kessler, R. C., Barker, P. R., Colpe, L. J. *et al.* (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, 60 (2), 184–189.
- McColl, E., Jacoby, A. & Thomas, L. (2006). Observing exposures and outcomes concurrently: Reporting surveys or cross-sectional studies. In: T. A. Lang & M. Secic (Eds). *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers*, 2nd edn. (pp. 239–251). Philadelphia: American College of Physicians.
- Mullen, P. (2000). Forensic mental health. *The British Journal of Psychiatry*, 176, 307–311.
- Ogloff, J., Lemphers, A. & Dwyer, C. (2004). Dual diagnosis in an Australian forensic psychiatric hospital: Prevalence and implications for services. *Behavioral Sciences and the Law*, 22, 543–562.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D. & Dickson, R. (2005a). *Routine Measurement of Outcomes in Australia's Public Sector Mental Health Services*. Melbourne: The University of Melbourne.
- Pirkis, J., Burgess, P., Kirk, P., Dodson, S. & Coombs, T. (2005b). *Review of Standardised Measures Used in the National Outcomes and Casemix Collection (NOCC)*. Sydney: NSW Institute of Psychiatry.
- Rosen, A., Hadzi-Pavlovic, D. & Parker, G. (1989). The Life Skills Profile: A measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin*, 15 (2), 325–337.
- Shaw, J. (2002). Needs assessment for mentally disordered offenders is different. *The Journal of Forensic Psychiatry & Psychology*, 13, 14–17.
- Shinkfield, G. & Ogloff, J. (2014). A review and analysis of routine outcome measures for forensic mental health services. *International Journal of Forensic Mental Health*, 13, 1–20.
- Sugarman, P. & Walker, L. (2007). *Health of the Nation Outcome Scales for Users of Secure/Forensic Services (version 2b)*. [Cited 30 March 2014]. Available from: URL: <http://www.rcpsych.ac.uk/researchandtrainingunit/honos/secure.aspx>
- Trauer, T. (Ed.) (2010). *Outcome Measurement in Mental Health: Theory and Practice*. Cambridge: University Press.
- Turner, T. & Salter, M. (2008). Forensic psychiatry and general psychiatry: Re-examining the relationship. *Psychiatric Bulletin*, 32, 2–6.
- Velligan, D. I., Lopez, L., Castillo, D. A., Manaugh, B., Milam, A. C. & Miller, A. L. (2011). Interrater reliability of using brief standardized outcome measures in a community mental health setting. *Psychiatric Services*, 62 (5), 558–560.
- Viera, A. J. & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37, 360–363.
- Webster, J., Bretherton, F., Goulter, N. & Fawcett, L. (2013). Does an educational intervention improve the usefulness of the Health of the Nation Outcome Scales in an acute mental health setting? *International Journal of Mental Health Nursing*, 22, 322–328.
- Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S. & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *British Journal of Psychiatry*, 172, 11–18.

Chapter Seven: Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population

7.1 Overview of Chapter Seven

This chapter introduces the second empirical study of this thesis.

7.2 Preamble to Submitted Paper: “Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population”

In the first instance, research paper three aimed to investigate whether the needs of forensic mental health patients were better able to be classified by ROM tools developed for use within a forensic or non-forensic environment. Moreover, in doing so, the study sought to explore whether it might be possible to differentiate between groups of forensic patients on the basis of their scores on each of these different tools. It was hypothesised that tools developed specifically for use with a forensic population would provide a better metric by which to differentiate forensic mental health patients at different stages of progress towards recovery and discharge than those tools that were developed for use in civil mental health settings.

The second aim of the study sought to explore whether the needs of forensic patients were indeed heterogeneous and whether these needs were subject to change over the course of admission. Given that admission length for forensic patients is far greater than for civil patients (Davoren et al., 2015; Turner & Salter, 2008), it might be anticipated that the needs of forensic clients are more likely to change over the course of their treatment.

Therefore, it was hypothesised that when the needs of forensic mental health patients were examined, distinct groups would be identified. It was further hypothesised that clinical/social needs would be most prominent at the point of admission, with forensic/security needs becoming the primary focus towards discharge.

Results from study three add support to the notion that forensic mental health patients are a heterogeneous group. Whilst there was consistency amongst the needs of patients within the acute, subacute and rehabilitation/community reintegration wards of the hospital, across the population as a whole a variety of high/low levels of clinical and forensic/security needs were identified. Moreover, it was demonstrated that it was possible to use a number of ROM tools to track the needs of this client group. In particular the HoNOS, HoNOS-Secure and LSP-16 were found to be most effective for this task. However, as patients progress towards discharge and community reintegration, employing broader needs assessment tools may be more effective than focusing on narrower outcome measures of clinical and forensic/security domains. In this way, by employing outcome measures that capture these broad range of needs, such tools may assist treating teams focus on the different needs of patients at various points in their journey towards recovery.

The study upon which article three is based was conducted between 2010 and 2011. Therefore, it is acknowledged that the findings of this study reflect the state of clinical practice during that period of time.

The following article has been prepared and submitted for publication in the International Journal of Forensic Mental Health. This is a peer-reviewed journal of the International Association of Forensic Mental Health Services (ISSN 1499-9013 [Print],

1932-9903 [Online]), which has been published since 2002 and now is published four times per year. In 2016, the *International Journal of Forensic Mental Health* had an impact factor of 1.25.

7.3 Authorship Indication Form for Chapter Seven**Swinburne Research**

Authorship Indication Form

For PhD (including associated papers) candidates

NOTE

This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each associated 'paper'. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each associated paper to be included in your thesis.

DECLARATION

We hereby declare our contribution to the publication of the 'paper' entitled:

Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population

First Author

Name: Gregg Shinkfield

Signature: _____

Percentage of contribution: 85%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Reviewed literature, obtained ethics approval, collected and coded data, conducted analysis, prepared and revised manuscript.*

Second Author

Name: Professor James Ogloff

Signature: _____

Percentage of contribution: 15%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Assisted with conceptualisation of study and manuscript, revised manuscript.*

Principal Coordinating Supervisor: Professor James Ogloff

Signature: _____

Date: _____

7.4 Declaration by Co-authors

The undersigned hereby certify that:

- a) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
- b) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- c) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- d) there are no other authors of the publication according to these criteria;
- e) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- f) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s):

Centre for Forensic Behavioural Science,
School of Psychology & Psychiatry, Monash University
505 Hoddle Street, Clifton Hill
Victoria, Australia

Professor J. Ogloff		Date:
---------------------	--	-------

**7.5 Submitted Paper Three: “Monitoring Risk, Security Needs, Clinical and Social
Functioning Within a Forensic Mental Health Population”**

Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population

Gregg Shinkfield, BSc, MSc(Hons), PGDipClinPsych¹

Professor James Ogloff, AM, BA, MA, JD, PhD¹

¹ Centre for Forensic Behavioural Science, Swinburne University and Victorian Institute of Forensic Mental Health (Forensicare).

Correspondence concerning this article should be addressed to Gregg Shinkfield, Victorian Institute of Forensic Mental Health, Locked Bag 10, Fairfield, Victoria, 3078, Australia. Electronic mail may be sent to gregg.shinkfield@forensicare.vic.gov.au

Submission for International Journal of Forensic Mental Health

Word Count: 4900

Abstract

Forensic mental health patients present with a variety of clinical, social, and forensic/security needs. As such, monitoring and tracking these disparate needs can be a complex task. In recent years, progress has been made towards identifying the most efficacious tools for assessing the needs of forensic mental health consumers. To extend upon existing research, the present study evaluated three forensic (HoNOS-Secure, CANFOR, LSI-R:SV) and three non-forensic (HoNOS, LSP-16, BASIS-32) tools for their clinical utility in assessing the needs of a forensic mental health population. Moreover, the extent to which the needs of this population are heterogeneous was also investigated, as well as the degree to which these needs differ over the course of admission. Results demonstrated that the HoNOS, HoNOS-Secure and LSP-16 were most effective at differentiating between clients at different levels of acuity. However, as a patient moved towards discharge/community reintegration, tools that focused on a broader range of life skills became increasingly pertinent. It was also observed that the clinical and forensic/security needs of forensic patients differed across the population. Implications for clinical practice, as well as the integration of outcome measurement in service delivery are discussed.

Keywords: outcome measure; forensic mental health; needs assessment

Monitoring Risk, Security Needs, Clinical and Social Functioning within a Forensic Mental Health Population

Forensic mental health patients are a heterogeneous group (Cohen & Eastman, 1997; Keulende Vos & Schepers, 2016). They present not only with the mental health difficulties and functional impairments seen in general psychiatry, but they can also demonstrate a variety of other needs including criminal behaviour, violent or sexual offending, personality and behavioural disturbances, self-harm, and/or high rates of co-morbid substance use (Coid et al., 2001; Ogloff et al., 2015; Ogloff, Lemphers, & Dwyer, 2004). In addition, consideration frequently needs to be given to issues of security, dangerousness, and risk management (Kennedy et al., 2010; Shaw, 2002). Even within a single forensic mental health population, the needs of patients can vary significantly. For some patients, mental health issues are the primary concern, with mental health difficulties contributing directly to offending behaviour. However, for others, mental health issues may not be the key factor underpinning their offending; with their criminogenic needs being more akin to non-mentally disordered offenders (Andrews & Bonta, 2006). However, such individuals may also come into contact with a forensic mental health service due to comorbid mental health issues, which impact on their daily and long term functioning.

Being able to monitor and track the disparate mental health and forensic/security needs of a forensic mental health population is a complex, yet necessary, task for informing treatment planning and monitoring progress towards recovery and discharge (Pirkis et al., 2005a). Within civil mental health services, the use of routine outcome measurement (ROM) tools is now embedded in clinical practice and service delivery across several international jurisdictions (Shinkfield & Ogloff, 2014; Trauer, 2010). Indeed, a range of tools have been developed for this task. These tools are generally well validated and demonstrate good clinical utility in civil mental health services (Pirkis et al., 2005b). However, within forensic mental health settings,

there has been a significant delay in identifying the most efficacious tools for monitoring outcomes of forensic mental health patients. Moreover, given the paucity of literature focusing on ROMs within forensic mental health populations, there has been a tendency for services and government agencies to adopt tools that were developed and validated for use with civil populations; with little research exploring whether or not they are indeed suitable for a forensic population (Shinkfield & Ogloff, 2014). For the most part, existing outcome measurement tools that were developed specifically for use with forensic populations have typically focused on issues of risk; with less attention being given to the broader clinical and psychosocial needs of forensic mental health patients (Keulen-de Vos & Schepers, 2016; Thomas et al., 2008).

Within Australia, as in other parts of the world, there is burgeoning research regarding the question of ‘what tools are most suitable for tracking the mental health and risk related needs of forensic patients’ (Shinkfield & Ogloff, 2015; 2016). Indeed the need to address this deficit was specifically acknowledged by the Australian government as being an important issue that needs to be addressed in the field of ROM within this jurisdiction (NMHIDEAP, 2013). Specifically, it was asserted that a clear gap remains in the measures employed for forensic services with respect to outcomes relating to risk, security and legal issues. The present study therefore sought to add to our knowledge of outcome measurement in forensic mental health, by evaluating three forensic and three non-forensic ROM tools for their clinical utility with a forensic mental health population. Moreover, we also sought to evaluate the extent to which the needs of forensic mental health patients are heterogeneous and whether they differ over the course of their admission. Extending this question further, if differences amongst this clinical group were identified, we sought to investigate whether it might be feasible to use existing ROM tools to differentiate between groups of forensic patients on the basis of their clinical, social and security related needs.

Aims and hypotheses

In the first instance, the present study aimed to investigate whether the needs of forensic mental health patients would be better classified by forensic or non-forensic ROM tools. That is, which set of tools would differentiate most effectively between groups of forensic patients on the basis of their outcome scores. It was hypothesised that tools developed specifically for use with forensic populations would provide a better metric by which to differentiate forensic mental health patients at different stages of progress towards recovery and discharge than those tools that were developed for use in civil mental health settings (hypothesis one).

The second aim of the study sought to explore whether the needs of forensic patients were indeed heterogeneous and whether these needs were subject to change over the course of admission. Previous research has identified that forensic mental health patients differ significantly from their mainstream counterparts with respect to the length of time they remain within an inpatient environment (Davoren et al., 2015; Turner & Salter, 2008). Given that admission length for forensic patients is far greater than for civil patients, it might be anticipated that the needs of forensic clients are more likely to change over the course of their treatment. Therefore, it was hypothesised that when the needs of forensic mental health patients were examined as a whole, distinct groups would be identified (hypothesis two). It was further hypothesised that clinical/social needs would be most prominent at the point of admission, with forensic/security needs becoming the primary focus towards discharge (hypothesis three).

Method

The study was conducted at the Thomas Embling Hospital (TEH), the sole forensic mental health inpatient facility within the state of Victoria, Australia. The hospital provides secure care for up to 116 patients across seven wards. The wards are structured to encompass the spectrum of patient recovery from acute care to community reintegration. All patients within

the hospital are detained under involuntary treatment orders, broadly separated into two main categories: forensic patients, who have been found either unfit to stand trial or not guilty of an offence on the grounds of mental impairment; and security patients, who are prisoners requiring assessment or treatment for mental health disorder. A small proportion of patients are also detained under civil involuntary hospitalisation orders.

Measures and materials

The present study is part of a larger research initiative, which has been reported on in a previous issue of this journal (**Insert reference citing author's previous work here**). This body of research was designed to identify and evaluate tools that could be used as routine outcome measures in forensic mental health. As such, the selection of tools for evaluation in the present study was based on previous research conducted by the authors. For a comprehensive discussion of the identification and selection of the tools used in the present study, the reader is referred to (**Insert reference citing author's previous work here**). Of the multitude of ROM tools available in the extant literature, five were ultimately selected for evaluation. The first three tools were developed for use with civil mental health populations and were also mandated for use by mental health services across Australia, namely: *Health of the Nation Outcome Scales* (HoNOS; Wing et al., 1998), *Life Skills Profile* (LSP-16; Rosen et al., 2006), *Behaviour and Symptom Identification Scale* (BASIS-32; Eisen, Dill & Grob, 1994). In addition, two ROM tools that had been developed specifically for use with a forensic mental health population were selected: *Health of the Nation Outcome Scale for Users of Secure Services* (HoNOS-Secure; Sugarman & Walker, 2007), and *Camberwell Assessment of Need - Forensic Version* (CANFOR; Thomas et al., 2003). Finally, a brief assessment tool that provides an estimate of risk for general recidivism was also selected, namely the *Level of Service Inventory-Revised: Screening Version* (LSI-R:SV; Andrews & Bonta, 1998).

Conceptually, these tools were able to be grouped on the basis of the broad needs areas they were designed to assess. As such these tools were considered in the following groupings: ‘clinical scales’, ‘security scales’, ‘needs scales’, and ‘risk scales’ (see Table 1). Each of these tools are described below:

Health of the Nation Outcome Scales (HoNOS; Wing et al., 1998)

The HoNOS is a 12-item clinician rated measure, designed to monitor four broad areas of clinical and social functioning for people with severe mental illness: behavioural problems, cognitive and physical impairment, symptomatic problems, and social functioning. Ratings are made on the basis of a client’s presentation over the past two weeks (Wing et al., 1998).

Health of the Nation Outcome Scale for Users of Secure Services (HoNOS-Secure; Sugarman & Walker, 2007)

The HoNOS-Secure is a member of the HoNOS family of tools, which has been adapted to provide a means of tracking the clinical, social and security needs of users of secure psychiatric services, prisons and forensic community services (Sugarman et al., 2009). The HoNOS-secure contains the original twelve ‘clinical and social functioning’ items of the HoNOS, which were modified to account for the environmental conditions typically found in a secure setting (Dickens et al., 2007). In addition, a seven-item ‘security scale’ monitors changes in a client’s need for risk and security management procedures (Long et al., 2010). As with the HoNOS, the HoNOS-Secure ‘clinical and social functioning scale’ is rated retrospectively; based on the previous two week period. Whereas, the ‘security scale’ is rated prospectively for the period ‘in the near future’ (Dickens et al., 2007).

Life Skills Profile (LSP-16; Rosen et al., 2006)

The LSP-16 is a clinician-rated instrument comprising of 16 items designed to measure four

broad domains of social and adaptive functioning: self-care, antisocial behaviour, withdrawal, and compliance with treatment (Rosen et al., 2006). The LSP-16 was designed for use with individuals living with schizophrenia and chronic mental illness in the community (Pirkis et al., 2005b). A patient's functioning is rated on each of the 16 items with respect to their behaviour over the preceding three month period. The LSP-16 was developed as an abbreviated version of the LSP-39 (Rosen et al., 1989) and was designed to emphasise the presence of life skills rather than focus on a client's deficits (Pirkis et al., 2005b).

Camberwell Assessment of Need - Forensic Version (CANFOR; Thomas et al., 2003)

The CANFOR is a needs assessment tool that was designed for use with individuals in contact with forensic services who are experiencing mental health problems (Thomas et al., 2003). Containing 25 items, the CANFOR covers a broad range of needs areas, including: basic life skills, mental health difficulties, functioning, substance use, safety to self and others, interpersonal needs, and offending issues. The CANFOR captures the views of service users, carers and staff for each domain (Thomas et al., 2008). Ratings are based on the patient's experience over the previous month, where 0 = no problem, 1 = need is present but currently being met, and 2 = need is present and currently unmet.

Behaviour and Symptom Identification Scale (BASIS-32; Eisen, Dill & Grob, 1994)

The BASIS-32 is a 32-item behavioural health assessment tool designed to monitor changes in a client's self-reported symptoms and functional difficulties. The 32 items assess a wide range of symptoms and problems across five domains of mental health and social functioning: Relation to Self and Others, Depression and Anxiety, Daily Living and Role Functioning, Impulsive and Addictive Behaviour, and Psychosis (Eisen, Dill & Grob, 1994).

Level of Service Inventory-Revised: Screening Version (LSI-R:SV; Andrews & Bonta, 1998)

The LSI-R:SV was developed as a brief version of the Level of Service Inventory – Revised (Andrews & Bonta, 1995). The LSI-R:SV provides a quantitative assessment of risks and needs associated with general recidivism: criminal history, criminal attitudes, criminal associates, and antisocial personality pattern. In addition, the LSI-R:SV samples the domains of employment, family, and substance abuse. Although brief, it has shown utility in treatment planning and predicting antisocial behaviour or recidivism during admission and upon release (Andrews & Bonta, 2006). The LSI-R:SV does not assess mental health difficulties.

Additional Outcomes Data

In addition to the formal outcome measurement tools described above, data were also recorded regarding a patient's ward placement (i.e., whether residing on an acute, subacute or rehabilitation/community reintegration unit). Within the hospital setting, transfer between units was based primarily on a patient's recovery from mental health difficulties. As such, ward placement was considered to reflect the acuity of a patient's mental health difficulties and progression towards discharge.

This study received ethics approval from the Swinburne University of Technology and Monash University human research ethics committees, as well as from the Forensicare Research Committee.

Data collection and analysis

Clinical staff completed the HoNOS, HoNOS-Secure, LSP-16, LSI-R:SV and the clinician rated version of the CANFOR for all patients within the study sample. Ratings were completed over a six month period for patients at the point of admission, discharge and every 91-days that

the patient remained within the hospital. All ratings were undertaken by mental health clinicians (e.g., psychiatric nurses, psychologists, occupational therapists and social workers) who had received training in the use of these tools to increase reliability of ratings (Rock & Preston, 2001). Ratings were made in accordance with the administration manuals for each tool (e.g., Eisen, Dill & Grob, 1994; Rosen et al., 2006; Sugarman & Walker, 2007; Thomas et al., 2003; Wing et al., 1998). In addition, patients were invited to complete the BASIS-32, as well as the patient rated version of the CANFOR. Demographic data and ward placement were also recorded at each collection occasion.

To investigate the first aim of the study, data generated by the six measures were evaluated using analysis of variance (ANOVA) to identify whether significant differences were present in the scores obtained by the forensic population at different levels of acuity. The alpha for all tests was set at 0.05. Ward placement was used as a measure of mental health acuity, with three levels being specified: acute, sub-acute and rehabilitation/community reintegration wards. Significant effects were further examined via Scheffé post hoc comparisons to ascertain where differences occurred between the different levels of acuity.

To investigate the second aim of the study, regarding whether or not variation exists within the clinical, security and risk related needs within the forensic mental health population, the ROM scores for each member of the sample were examined and their clinical and forensic/security needs were classified as being either in the high or low range (see figure one). The HoNOS and HoNOS-Secure ‘Security Scale’ were employed to provide a measure of clinical (HoNOS) and security/forensic (HoNOS-Secure ‘security scale’) needs. These tools were selected as they demonstrated the greatest ability to differentiate between patients at each of the three levels of ward placement/acuity. This finding emerged during the initial ANOVA analysis described above (see Table 3). Cut-off scores for ‘high’ versus ‘low’ clinical needs were determined by

identifying the median score obtained on both measures. Whilst other methods of identifying cut-off scores are available, due to the relatively small sample size available, using the median split was considered the most appropriate means determining the point to divide low and high scores on these tools.

On the basis of the median scores, it was determined that a HoNOS value of 0 – 5 would be considered low, with a score of six or greater being indicative of a high level of clinical needs. Likewise, for security/forensic needs it was determined that a score on the HoNOS-Secure ‘security scale’ of 0 – 6 would be considered in the low range, with a value of seven or greater indicating a high level of security/forensic need. HoNOS and HoNOS-Secure scores were interrogated for each member of the sample population and the number of patients meeting criteria for each of the four categories was quantified (i.e., high or low scores for clinical and security needs; see figure one). This procedure was completed for all three levels of ward acuity, as well as for the population as a whole.

Results

Sample characteristics

A total of 202 assessments were completed, of which 89 (44%) were conducted for patients residing in an acute unit, 78 (39%) for patients in a sub-acute unit and 35 (17%) for rehabilitation/community integration. Most patients were male ($n = 217$, 85.7%). At the time of data collection, the total patient population of TEH was 116, of which 100 were male (84.7%), with the hospital being divided into 60 acute beds, 40 sub-acute beds and 16 community reintegration beds (51%, 35% and 14% respectively). As such, it was considered that the sample obtained provided a good representation of the hospital population.

Evaluation of outcome measures with a forensic population

With respect to the first aim of the study, each set of ROM tools were evaluated using ANOVA to identify any differences in scores between patients at each level of ward acuity. The first set of outcome measures considered were those presented within the ‘Clinical Scales’ group. As demonstrated in Table 2, ANOVA results for the HoNOS and HoNOS-Secure ‘clinical scale’ indicated that a significant difference emerged in the scores obtained by patients across the three levels of ward acuity. While the BASIS-32 approached significance, there remained a degree of overlap in scores of patients across all three wards. Closer examination of the HoNOS and HoNOS-Secure ‘clinical scale’ via Scheffé post-hoc analysis revealed that the mean scores obtained for patients at all levels of ward placement differed from each other to a statistically significant extent. However, there was no differentiation between BASIS-32 scores at any level of ward acuity.

With regards to those outcome measures presented within the ‘Needs Scales’ group, several findings emerged from the data presented in Table 2. Firstly, ANOVA results for the CANFOR subscale ‘Patient Ratings of Met Needs’ suggested that there was no difference in the scores obtained by patients across the three levels of ward acuity. However, the remaining subscales of the CANFOR (i.e., ‘Clinician Rating of Met Needs’ and ‘Clinician/Patient Ratings of Unmet Needs’), as well as the LSP-16, indicated significant differences across ward acuity. Examination of Scheffé post-hoc analysis for these scales (see Table 3) revealed that while the CANFOR unmet needs subscales (both patient and clinician ratings) were able to effectively differentiate between patients residing on an acute unit and the other two levels of ward acuity, there were no statistically significant differences noted between patients on sub-acute and rehabilitation wards. However, a more distinct difference emerged with respect to mean scores obtained by the LSP-16, with the post-hoc analysis indicating that mean scores for all three levels of acuity differed to a statistically significant extent.

With respect to the remaining outcome measures within the ‘security scales’ and ‘risk scales’ (i.e., HoNOS-Secure ‘security scale’ and LSI-R:SV), results presented in Table 2 indicated that significant differences emerged in the mean scores obtained across the three levels of ward acuity. However, closer inspection of these differences via Scheffé post-hoc analysis revealed that while mean scores of the HoNOS-Secure (security scale) differed across all three levels of acuity, differences were only observed between acute and subacute/rehabilitation for LSI-R:SV. That is, no difference in mean scores were found between LSI-R:SV scores for patients on the subacute and rehabilitation wards.

Clinical and Forensic/Security needs of forensic mental health patients

As demonstrated in Table 4, it was found that there was indeed variation in the clinical and forensic/security needs amongst the present sample of forensic mental health patients. Examining the population as a whole, it was noted that the majority of the sample (53.5%) were identified as possessing high needs in both the clinical and security/forensic domains. Moreover, a fifth (20.8%) of the population possessed high forensic/security needs, but low clinical needs. Table 4 also suggested that a small proportion of patients have high clinical needs, but low security needs (6.4%%); and the remainder (19.3%) are considered to have low needs in both the clinical/security domains.

To understand this range of needs in greater detail, Table 4 also presents the proportion of patients with high/low clinical and forensic/security needs for each of the three levels of ward acuity. Data indicated that the majority of patients within the acute wards demonstrated a high level of need in both the clinical and forensic domain. In contrast, those patients in the rehabilitation/community integration wards were considered low on both domains. Finally, there was much greater variation in clinical and forensic/security needs amongst those patients residing on the subacute units, with a fairly even distribution of patients across each of the four

needs categories.

Discussion

The present study was designed to address two key questions. Firstly, to determine whether it would be possible to identify and differentiate amongst groups of forensic patients on the basis of clinical and forensic needs. Secondly, to identify whether tools designed for use with a forensic or non-forensic populations were most readily able to perform this task. The final aim of the study was to examine the range of clinical/forensic needs amongst forensic populations and to describe how these needs change over the course of admission to a secure forensic hospital.

In the first instance, six ROM tools were selected for ANOVA evaluation with post-hoc Scheffé analysis. It was observed that with the exception of the BASIS-32 and the 'Met Needs' subscale of the patient rated version of the CANFOR, the majority of mean scores obtained on these measures differed significantly for patients at different levels of ward acuity across the forensic hospital. However, only the HoNOS, HoNOS-Secure 'clinical and security scales,' and the LSP-16 were able to significantly differentiate among patients across all three levels of acuity. The mean scores obtained by clients in the acute, subacute and rehabilitation/community integration wards did not overlap and were statistically distinct for the HoNOS, HoNOS-Secure and LSP-16. The remaining tools (i.e., CANFOR and LSI-R:SV) demonstrated differences between the mean scores obtained by patients on the acute unit and the subacute/rehabilitation wards; however, neither tool was sensitive enough to detect differences between patients residing on the subacute and rehabilitation wards.

These results provided partial supported for hypothesis one, as one of the three tools that

performed best as a measure of change for forensic clients had been developed specifically for use with forensic populations. These findings also provided some support for hypothesis two, as it suggested that distinct groups of forensic patients do indeed exist within the sample population, with the groups possessing different levels of clinical/forensic need.

The results obtained from the LSI-R:SV perhaps warrant further consideration, as this was the only tool designed to directly evaluate the construct of risk of general recidivism. Whilst the LSI-R:SV was able to differentiate between clients residing on an acute / non-acute ward, it was less effective at differentiating between clients in the subacute/rehabilitation wards. This finding was not entirely unexpected for two main reasons. Firstly, the LSI-R:SV is comprised of a mixture of static and dynamic items, with the static items (e.g., ‘two or more previous convictions’ and ‘arrested under the age of 16’) not being subject to change over time. Given that the LSI-R:SV contains eight items, only the six dynamic items are able to vary and demonstrate change. Moreover, for a client who meets criteria for both of the static items, their overall score can never reduce lower than two out of eight. This floor effect appears to have occurred within the present sample, with the mean score for patients within the rehabilitation/community reintegration wards being 2.5. As such, it might be argued that this reduces the utility of the LSI-R:SV to demonstrate change in risk over time within secure environments. By contrast, in previous research conducted with patients from the Thomas Embling Hospital, the LSI-R:SV proved good predictive validity for re-offending by patients upon discharge from the hospital (Ferguson, Ogloff, & Thomson, 2009).

Secondly, it is also noted that the patients residing in forensic hospitals as a whole are typically not generalist offenders. Rather, as a population, they are generally detained after committing a serious offence whilst mentally unwell. In this sense, their risk of reoffending is likely to be

captured more effectively by tools that focus on violent recidivism rather than general offending *per se*. In the field of outcome measurement, the ability to track and evaluate factors pertinent to a client group's recovery is paramount (**Insert reference citing author's previous work here, 2016**). As such, the LSI-R:SV appeared limited in its capacity to perform this role.

On the basis of data obtained from each of the ROM tools in the present study, it appears that the HoNOS, HoNOS-Secure or LSP-16 were the most effective at performing the task of monitoring changes pertinent to a clients' recovery and progress towards discharge in the forensic mental health population in the current sample.

The second component of this study was to explore the question of whether the needs of the forensic mental health sample were heterogeneous, and whether these needs changed over the course of their admission. Through identifying the presence of high/low clinical and security/forensic needs, hypothesis two was supported in that distinct groups of forensic patients were identified. Taken as a whole, it was observed that the majority of patients within the sample possessed high levels of need in both the clinical and forensic/security domains. This was to be expected, given that admission criteria into the secure hospital were such that only individuals with both forensic and mental health needs were able to be admitted. However, data also revealed that a significant proportion of clients had high levels of need in only one domain (i.e., either clinical or forensic/security needs), and indeed approximately a fifth of the entire population was considered to have low levels of need across both domains.

Turning specifically to results obtained by patients on the acute units, while the majority were identified as having high clinical/security need, approximately one-eighth (12.5%) were considered to have high clinical but low forensic/security needs, and a further 6.8% were considered to have low needs across both domains. While this finding may seem somewhat surprising, it should be noted that legislation governing Custodial Supervision Orders within

the state of Victoria (Australia) provides the ability for a person to be detained in custody within an approved mental health service after a finding of not guilty because of mental impairment (*Crimes (Mental Impairment and Unfitness to be Tried) Act*, 1997). As such, there are occasions where a person may have been unwell at the time of their offence, but while their mental state stabilises once treatment has been provided, they may continue to require the security of the acute unit in order to manage their ongoing physical risk. Given the limited number of beds within the hospital, it is possible that a person may reside on an acute unit in the absence of symptoms whilst awaiting transfer to a less acute ward. What this highlights, is the complex nature of patient composition within a forensic setting and the importance of being able to use tools that are sensitive enough to detect a range of needs across both clinical and security domains in order to effectively target interventions for all residents.

The present findings also raised questions regarding the needs of patients residing in the rehabilitation/community reintegration unit. It was observed from the data that the majority (94.2%) of patients on this ward were considered to have low levels of both clinical and forensic/security needs on the basis of their ROM scores. As such, it may be questioned why these patients continue to reside in such a highly secure and restrictive environment if their clinical and forensic/security needs are evaluated to be low. Indeed a body of research has emerged over the past decade which suggests that forensic patients may be subjected to ongoing secure care for longer periods than may be necessary to satisfy the criteria that they ‘no longer presents a risk of serious endangerment’ to themselves or others (Ruffles, 2010; Victorian Law Reform Commission, 2014). However, while this question is outside the scope of the present study, it may warrant further investigation by future research. Alternatively, the results obtained may also be viewed as a limitation of these tools regarding their sensitivity and ability to accurately detect needs in this subgroup of clients. It might be argued that for this subgroup of patients, focusing treatment on broader needs such as those identified by the LSP-

16 and CANFOR may be of greater value than the narrower domains captured by clinical or forensic/security measures such as the HoNOS / HoNOS-Secure.

At this point in a patient's journey towards recovery and discharge, the function of a mental health service is to assist the patient develop and master the skills required for independent living upon discharge into the community. However, despite this, ongoing use of clinical and forensic/security ROMs may be useful to assist in identifying any changes in a client's presentation that may suggest an exacerbation in mental health difficulties or increased need to focus on security issues.

Taken together, the above findings provided support for hypothesis three, suggesting that the needs of this population change over the course of their admission towards. As such, ROM tools used with this population should also demonstrate capacity to measure the range of needs required to assist patients progress towards community reintegration and discharge.

Conclusion

The findings of the present study add support to the notion that forensic mental health patients are a heterogeneous group. Whilst there was consistency amongst the needs of patients within the acute, subacute and rehabilitation/community reintegration wards of the hospital, across the population as a whole a variety of high/low levels of clinical and forensic/security needs were identified. Moreover, it was demonstrated that it was possible to use a number of ROM tools to track the needs of this client group. In particular the HoNOS, HoNOS-Secure and LSP-16 were found to be most effective for this task. However, as a patient progresses towards discharge and community reintegration, employing broader needs assessment tools may be more effective than focusing on narrower outcome measures of clinical and forensic/security domains. In this way, by employing outcome measures that capture these broad range of needs,

such tools may assist treating teams to focus on the different needs of patients at various points in their journey towards recovery.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2006). *The Psychology of Criminal Conduct* (4th ed.). Cincinnati, OH.: Anderson Publishing Co.
- Cohen, A., & Eastman, N. (1997). Needs assessment for mentally disordered offenders and others requiring similar services: theoretical issues and a methodological framework. *British Journal of Psychiatry*, *171*, 412 - 416.
- Coid, J., Kahtan, N., Gault, S., Cook, A., & Jarman, B. (2001). Medium secure forensic psychiatry services: Comparison of seven English health regions. *British Journal of Psychiatry*, *178*, 55 – 61.
- Davoren, M., Byrne, O., O’Connell, P., O’Neill, H., O’Reilly, K., & Kennedy, H. G. (2015). Factors affecting length of stay in forensic hospital setting: need for therapeutic security and course of admission. *BMC Psychiatry*, *15*, 301 - 316.
- Ferguson, M. A., Ogloff, J. R. P., & Thomson, L. (2009). Predicting recidivism by mentally disordered offenders using the LSI-R:SV. *Criminal Justice and Behavior*, *36*, 5 - 20.
- Kennedy, H. G., O’Neill, C., Flynn, G., & Gill, P. (2010). *Dangerousness understanding, recovery and urgency manual (The DUNDRUM Quartet) v1.0.22*. Dublin National Forensic Mental Health Service.
- Keulen-de Vos, M., & Schepers, K. (2016). Needs assessment in forensic patients: a review of instrument suites. *International Journal of Forensic Mental Health*, *15*, 283 – 300.

- National Mental Health Information Development Expert Advisory Panel. (2013). *Mental Health National Outcomes and Casemix Collection: NOCC Strategic Directions 2014 - 2024*. Retrieved from Canberra:
- Ogloff, J., Lemphers, A., & Dwyer, C. (2004). Dual diagnosis in an Australian forensic psychiatric hospital: prevalence and implications for services. *Behavioral Sciences and the Law*, 22, 543 – 562.
- Ogloff, J. R. P., Talevski, D., Lemphers, A., Simmons, M., & Wood, M. (2015). Co-occurring psychiatric illness, substance use disorders, and antisocial personality disorder among clients of forensic mental health services. *Psychiatric Rehabilitation Journal*, 38, 16–23.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005a). *Routine measurement of outcomes in Australia's public sector mental health services*. Melbourne: The University of Melbourne.
- Pirkis, J., Burgess, P., Kirk, P., Dodson, S., & Coombs, T. (2005b). *Review of standardised measures used in the National Outcomes and Casemix Collection (NOCC)*. Sydney: NSW Institute of Psychiatry.
- Ruffles, J. (2010). *The management of forensic patients in Victoria: The more things change, the more they remain the same*. (PhD Thesis), Monash University, Melbourne.
- Shaw, J. (2002). Needs assessment for mentally disordered offenders is different. *The journal of Forensic Psychiatry & Psychology*, 13, 14 - 17.

- Shinkfield, G., & Ogloff, J. (2014). A review and analysis of routine outcome measures for forensic mental health services. *International Journal of Forensic Mental Health, 13*, 252 – 271.
- Shinkfield, G., & Ogloff, J. (2015). Use and interpretation of routine outcome measures in forensic mental health. *International journal of mental health nursing, 24*, 11-18.
- Shinkfield, G., & Ogloff, J. (2016). Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital. *The Journal of Forensic Psychiatry & Psychology, 27*, 867 – 885.
- Thomas, S. D. M., Slade, M., McCrone, P., Harty, M.-A., Parrott, J., Thornicroft, G., & Leese, M. (2008). The reliability and validity of the Forensic Camberwell Assessment of Need (CANFOR): a needs assessment for forensic mental health service users. *International Journal of Methods in Psychiatric Research, 17*, 111 – 120.
- Trauer, T. (Ed.) (2010). *Outcome measurement in mental health: theory and practice*. Cambridge: University Press.
- Turner, T., & Salter, M. (2008). Forensic psychiatry and general psychiatry: re-examining the relationship. *Psychiatric Bulletin, 32*, 2 – 6.
- Victorian Law Reform Commission. (2014). *Review of the Crimes (Mental Impairment and Unfitness to be Tried) Act 1997: Consultation Paper*. Melbourne: Victorian Law Reform Commission.

		CLINICAL/SOCIAL NEEDS	
		High	Low
RISK / FORENSIC NEEDS	High	High security needs/risk of offending, with acute psychiatric needs	High security needs/risk of offending, without prominent psychiatric needs
	Low	Acute psychiatric needs, but low-risk of offending or security needs	No prominent psychiatric needs, with low risk of offending / security needs

Figure One: Schematic showing the four categories of high/low clinical and forensic needs, with description of each domain.

Table 1 *Assessment tools grouped by needs areas*

Needs Area	Outcome Measures
Clinical Scales	HoNOS HoNOS-Secure (Clinical scale) BASIS-32
Security Scales	HoNOS-Sec (Security Scale) HoNOS-Secure (Total score)
Needs Scales	CANFOR - Clinician rated (Met Needs / Unmet Needs) - Patient rated (Met Needs / Unmet Needs) BASIS-32 LSP-16
Risk Scales	LSI-R:SV HCR-20

Table 2 *ANOVA – Clinical scales by ward acuity*

Scale Type	Outcome Measure	Sum of Squares	df	Mean Square	F	Sig.
Clinical Scales	HoNOS	2268.28	2	1134.14	40.41	.000
	HoNOS-Secure (Clinical)	2192.19	2	1096.10	45.38	.000
	BASIS-32	1247.87	2	623.93	2.68	.072
Needs Scales	CANFOR (Patient ratings – Met Needs)	14.45	2	7.22	.86	.425
	CANFOR (Patient ratings – Unmet Needs)	94.98	2	47.49	9.90	.000
	CANFOR (Clinician ratings – Met Needs)	185.55	2	92.77	9.34	.000
	CANFOR (Clinician ratings – Unmet Needs)	579.32	2	289.66	40.25	.000
	LSP-16	3266.15	2	1633.07	38.14	.000
Security Scales	HoNOS-Secure (Security)	1114.18	2	557.09	61.07	.000
	HoNOS-Secure (Total)	6399.52	2	3199.76	73.84	.000
Risk Scale	LSI-R:SV	53.17	2	26.58	8.52	.000

Table 3 Mean scores of patients on acute, subacute and rehabilitation wards for each outcome measurement tools

Scale	Acute (a)					Subacute (b)					Rehabilitation (c)				
	n	M	SD	95% CI		n	M	SD	95% CI		n	M	SD	95% CI	
				LL	UL				LL	UL				LL	UL
Clinical Scales															
HoNOS (Adult)	89	10.2 ^{b,c}	6.7	8.8	11.6	78	4.8 ^{a,c}	4.1	3.8	5.7	35	1.7 ^{a,b}	2.8	0.7	2.7
HoNOS-Secure (Clinical)	89	10.9 ^{b,c}	5.9	9.7	12.2	78	6.2 ^{a,c}	4.4	5.2	7.2	35	2.1 ^{a,b}	2.5	1.2	3
BASIS-32	48	15.6	18.4	10.3	21	61	10.9	13.4	7.4	14.3	25	7.3	12.8	2	12.6
Needs Scales															
CANFOR (Patient: Met Needs)	89	3.6	3.1	3	4.3	78	3.2	2.9	2.5	3.8	35	2.9	2.1	2.2	3.7
CANFOR (Patient: Unmet Needs)	89	2.3 ^{b,c}	2.6	1.8	2.9	78	1.1 ^a	1.6	0.8	1.5	35	0.7 ^a	2.3	0	1.5
CANFOR (Clinician: Met Needs)	89	7.2 ^{b,c}	3	6.5	7.8	78	5.6 ^a	3.3	4.9	6.4	35	4.7 ^a	3.1	3.6	5.8
CANFOR (Clinician: Unmet Needs)	89	4.6 ^{b,c}	3.3	3.9	5.3	78	1.9 ^a	2.4	1.4	2.4	35	1.9 ^a	2.4	1.4	2.4
LSP-16	86	15.9 ^{b,c}	7.4	14.3	17.5	77	9.5 ^{a,c}	6.3	8.1	10.9	35	5.5 ^{a,b}	4.4	4	7
Security/Forensic Scale															
HoNOS-Secure (Security)	89	9.8 ^{b,c}	3.7	9.1	10.6	78	5.9 ^{a,c}	2.6	5.4	6.6	35	3.9 ^{a,b}	1.5	3.4	4.4
Risk Scale															
LSI-R:SV	89	3.8 ^{b,c}	2.0	3.4	4.3	78	3.1 ^a	1.4	2.7	3.4	35	2.5 ^a	2.0	3.0	3.6

Note. Superscript letters denote that the group mean differed significantly ($p < 0.05$) from: ^a = acute, ^b = subacute or ^c = rehabilitation

Table 4 *Proportion of patients with high/low clinical and forensic needs, by ward acuity*

		Acute (<i>n</i> = 89)		Subacute (<i>n</i> = 78)		Rehabilitation (<i>n</i> = 35)		Total Sample (<i>n</i> = 202)	
		Clinical/Social Needs		Clinical/Social Needs		Clinical/Social Needs		Clinical/Social Needs	
		High	Low	High	Low	High	Low	High	Low
Security / Forensic Needs	High	59 (67.0%)	12 (13.6%)	15 (19.0%)	17 (21.5%)	0 (0%)	1 (2.9%)	108 (53.5%)	42 (20.8%)
	Low	11 (12.5%)	6 (6.8%)	16 (20.3%)	31 (39.2%)	1 (2.9%)	33 (94.2%)	13 (6.4%)	39 (19.3%)

Note. Clinical/Social need were identified via HoNOS scores, with 0 – 5 being low and 6+ being high need. Security/Forensic needs were identified via HoNOS-Secure scores on the ‘Security Scale’, with 0 – 6 being low and 7+ being high needs.

Chapter Eight: Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital

8.1 Overview of Chapter Eight

This chapter introduces the third empirical study of this thesis.

8.2 Preamble to Published Paper: “Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital”

The Health of the Nation Outcome Scale (HoNOS) is a widely used tool for monitoring consumer outcomes within mental health services. However, concern about the suitability of this tool in forensic mental health settings led to the development of a forensic version of this measure known as the *HoNOS-Secure*. To date, no direct comparison of these versions has appeared in the empirical literature. In the present study, a cohort of forensic mental health patients were rated using the HoNOS and HoNOS-secure. Pearson correlations were generated to compare the tools at both a total score and item level. Logistic regression was employed to evaluate how well these tools would categorise patients on a range of measurable outcomes. HoNOS scores were also compared against civil mental health patients to evaluate differences between these populations.

The findings of this study indicated that the HoNOS/HoNOS-Secure correlated strongly at the total score level, but demonstrated variable correlations at the item level. Logistic regression suggested that the HoNOS-Secure ‘clinical and social functioning scale’ adds little to the HoNOS in a forensic setting; however, the HoNOS-Secure ‘security scale’ added significant benefit to both versions. Results remained stable when re-

evaluated over time. Forensic and civil mental health patients were found to demonstrate the same degree of psychopathology at the point of admission; however, they differed at review and discharge collection occasions. Implications for clinical practice and policy are explored.

The following article was published in the *Journal of Forensic Psychiatry and Psychology* (ISSN 1478-9949 [Print], 1478-9957 [Online]). This is a peer-reviewed journal which has been published bi-monthly since 1990. In 2015 the *Journal of Forensic Psychiatry and Psychology* had an impact factor of 0.810.

8.3 Authorship Indication Form: Chapter Eight**Swinburne Research**

Authorship Indication Form

For PhD (including associated papers) candidates

NOTE

This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each associated 'paper'. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each associated paper to be included in your thesis.

DECLARATION

We hereby declare our contribution to the publication of the 'paper' entitled:

Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital

First Author

Name: Gregg Shinkfield

Signature: _____

Percentage of contribution: 85%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Reviewed literature, obtained ethics approval, collected and coded data, conducted analysis, prepared and revised manuscript.*

Second Author

Name: Professor James Ogloff

Signature: _____

Percentage of contribution: 15%

Date: __/__/____

Brief description of contribution to the 'paper' and your central responsibilities/role on project:

- *Assisted with conceptualisation of study and manuscript, revised manuscript.*

Principal Coordinating Supervisor: Professor James Ogloff

Signature: _____

Date: _____

8.4 Declaration by Co-authors

The undersigned hereby certify that:

- a) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
- b) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- c) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- d) there are no other authors of the publication according to these criteria;
- e) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- f) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s):

Centre for Forensic Behavioural Science,
Swinburne University of Technology and Forensicare
505 Hoddle Street, Clifton Hill
Victoria, Australia

Professor J. Ogloff:		Date:
----------------------	--	-------

8.5 Published Paper Four: “Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital”

Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital

Gregg Shinkfield  and James Ogloff

Centre for Forensic Behavioural Science, Victorian Institute of Forensic Mental Health (Forensicare),
Swinburne University, Fairfield, Australia

ABSTRACT

The Health of the Nation Outcome Scale (HoNOS) is a widely used tool for monitoring consumer outcomes within mental health services. However, concern about its suitability in forensic mental health settings led to the development of a forensic version of this tool (HoNOS-Secure). To date, no direct comparison of these versions has appeared in the empirical literature. In the present study, a cohort of forensic mental health consumers was rated using the HoNOS and HoNOS-Secure. Pearson correlations were generated to compare the tools at a total score and item level. Logistic regression was employed to evaluate how well these tools categorise patients on a range of measurable outcomes. HoNOS scores were also compared against civil mental health consumers to evaluate differences between these populations. The HoNOS/HoNOS-Secure correlated strongly at the total score level, but demonstrated variable correlations at the item level. Logistic regression suggested that the HoNOS-Secure 'clinical and social functioning scale' adds little to the HoNOS in a forensic setting; however, the HoNOS-Secure 'security scale' added significant benefit to both versions. Results remained stable when re-evaluated over time. Forensic and civil mental health patients were found to demonstrate the same degree of psychopathology at the point of admission; however, they differed at review and discharge collection occasions. Implications for clinical practice and policy are explored.

ARTICLE HISTORY Received 6 January 2016; Accepted 25 September 2016

KEYWORDS HoNOS; HoNOS-Secure; outcome measure; forensic mental health

Introduction

The Health of the Nation Outcome Scale (HoNOS; Wing et al., 1998) is a widely used tool designed to monitor patient outcomes within mental health services. Since its development in 1998, the HoNOS has come to be mandated as a routine outcome measure (ROM) in several international jurisdictions, including the United Kingdom (Dickens, Sugarman, Picchioni, & Long, 2010), Australia

CONTACT Gregg Shinkfield  gregg.shinkfield@forensicare.vic.gov.au

© 2016 Informa UK Limited, trading as Taylor & Francis Group

and New Zealand (Shinkfield & Ogloff, 2014). A sizeable literature shows that the HoNOS performs well in civil mental health settings with regard to its sensitivity, specificity and predictive validity (e.g. Pirkis et al., 2005; Shinkfield & Ogloff, 2014). However, despite the utility of these tools when used with a civil population, several difficulties have been reported when attempting to apply the HoNOS in specialist mental health settings, including forensic mental health (Dickens et al., 2010).

Within a forensic or secure environment, two main factors have been identified that may reduce the utility of the HoNOS with this client group. In the first instance, several authors have noted that the broad needs of a forensic mental health population are not entirely analogous to those of civil mental health consumers (e.g. Dickens, Sugarman, & Walker, 2007; Ogloff, Lemphers, & Dwyer, 2004; Ogloff, Talevski, Lemphers, Simmons, & Wood, 2015; Shinkfield & Ogloff, 2015). In particular, differences exist regarding the level of security, risk and risk management procedures required for these client groups (Kennedy, O'Neill, Flynn, & Gill, 2010; Shaw, 2002). Indeed, forensic mental health patients typically remain in psychiatric care longer than civil patients, due to the perceived or actual risk profile of this group (e.g. Davoren et al., 2015; Shinkfield & Ogloff, 2014). Therefore, these represent important outcome domains for forensic consumers; however, they are not represented in the HoNOS. Secondly, in a secure environment, several HoNOS items appear less meaningful and more difficult to interpret than in civil settings. For example, Item 3 of the HoNOS focuses on a patient's use of substances over a two-week period. For most patients in secure settings, access to substances may be limited by environmental constraints; yet, the underlying problem may be demonstrated via cravings for substances, medication-seeking behaviour or other markers not captured by the HoNOS (Shinkfield & Ogloff, 2015).

Despite these limitations, research has demonstrated that the HoNOS can be rated reliably by staff within a forensic setting and it is considered sensitive enough to detect clinical change in secure populations (Shinkfield & Ogloff, 2015). As such, although the specific needs of forensic and civil mental health populations may differ, the constructs underpinning the HoNOS appear to be meaningful as a broad measure of clinical and social functioning for forensic consumers. However, no studies have examined the differences between HoNOS scores obtained by civil and forensic populations, nor has it been determined whether it is possible to compare the needs of these groups on the basis of their HoNOS scores.

The HoNOS-Secure

To increase the applicability of the HoNOS to a wider range of clinical groups, this tool has been adapted for use with several specialist populations, including children, older adults, people with a learning disability/acquired brain injury

and those within forensic mental health. The earliest adaptation of the HoNOS for use with forensic populations was the HoNOS-MDO (Mentally Disordered Offenders; Sugarman & Everest, 1999). This was subsequently expanded and refined over two iterations into the *Health of the Nation Outcome Scales for Users of Secure and Forensic Services* (HoNOS-Secure) to provide a means of tracking the clinical, social and security needs of users of secure psychiatric services, prisons and forensic community services (Sugarman, Walker, & Dickens, 2009).

As shown in Table 1, the HoNOS-Secure contained the original 12 'clinical and social functioning' items of the HoNOS, which were modified to account for the environmental conditions typically found in a secure setting (Dickens et al., 2007). In addition, a seven-item 'security scale' was included to monitor changes in a client's need for risk and security management procedures (Long et al., 2010). As with the original version of the HoNOS, the HoNOS-Secure 'clinical and social functioning scale' was designed to be rated retrospectively; based on a period of two weeks prior to the day on which tool was completed. Whereas, the 'security scale' was designed to be rated prospectively for the period 'in the near future' (Dickens et al., 2007). For a full description of the development of the HoNOS-Secure, readers are referred to Dickens et al. (2007). The HoNOS-Secure is currently freely available from the Royal College of Psychiatrists website (<http://www.rcpsych.ac.uk/quality/honos/secure.aspx>).

Previous research suggests that the HoNOS-Secure is a reliable tool (Dickens et al., 2007) that can effectively track the needs of forensic mental health consumers over time (Long et al., 2010). Moreover, it has been used to provide a measure of service delivery when combined with other performance indicators. Being able to correctly classify consumers on outcomes relevant to their need for ongoing mental health care and containment of risk (i.e. predictive validity) is an important function of an outcome measurement tool, particularly in the context of casemix evaluation (Pirkis, Burgess, Kirk, Dodson, & Coombs, 2005; Sugarman et al., 2009).

The original version of the HoNOS-Secure (i.e. HoNOS-MDO; Sugarman & Everest, 1999) was found to correlate strongly with the HoNOS; however, to the authors' knowledge, this finding has not been repeated with the current version of the HoNOS-Secure. Moreover, little research has appeared in the literature that directly compares the HoNOS-Secure and the original HoNOS, particularly in terms of the ability of these tools to accurately classify forensic patients with respect to real-life outcomes. As such, whilst the HoNOS-Secure appears to be an effective tool for use in forensic and secure environments, the question of whether it out-performs the original HoNOS in such settings has not yet been empirically evaluated.

Table 1. Structure and item descriptors of the HoNOS and HoNOS-Secure.

HoNOS (Working aged adult version)			HoNOS-Secure	
Item	Descriptor	Item	Descriptor	
		<i>Security scale</i>		
1	Overactive, aggressive, disruptive or agitated behaviour	A	Risk of harm to adults or children	
2	Non-accidental self-injury	B	Risk of self-harm (deliberate or accidental)	
3	Problem drinking or drug taking	C	Need for building security to prevent escape	
4	Cognitive problems	D	Need for a safely staffed living environment	
5	Physical illness or disability problems	E	Need for escort on leave (beyond the secure perimeter)	
6	Problems associated with hallucinations and delusions	F	Risk to individual from others	
7	Problems with depressed mood	G	Need for risk management procedures	
8	Other mental and behavioural problems			
9	Problems with relationships			
10	Problems with activities of daily living			
11	Problems with living conditions			
12	Problems with occupation and activities			
		<i>Clinical and social functioning scale</i>		
		1	Overactive, aggressive, disruptive or agitated behaviour	
		2	Non-accidental self-injury	
		3	Problem drinking or drug taking	
		4	Cognitive problems	
		5	Physical illness or disability problems	
		6	Problems associated with hallucinations and delusions	
		7	Problems with depressed mood	
		8	Other mental and behavioural problems	
		9	Problems with relationships	
		10	Problems with activities of daily living	
		11	Problems with living conditions	
		12	Problems with occupation and activities	

Australian context

Within Australia, all public mental health services are mandated to use the HoNOS (or an age specific variant of the HoNOS) as part of a suite of ROM tools, known as the National Outcomes Casemix Collection (NOCC; Burgess, Pirkis, & Coombs, 2015; Pirkis, Burgess, Kirk et al., 2005). The nationwide protocol for collection of NOCC data specifies that ROMs, including the HoNOS, be completed for each patient on admission, at discharge and every 91 days whilst they remain within a service (Burgess et al., 2015). The collection of ROMs by mental health services is also supported by a nationwide system for reporting and analysis of outcomes data (Burgess et al., 2015). This system provides public access to aggregated data submitted by each state and territory, and enables the data-set to be freely interrogated with regard to a variety of high-level descriptors (e.g. age, gender, legal status). These data are available via the Web Decision Support Tools (wDST), which can be accessed on the Australian Mental Health Outcomes and Classification Network's website (<http://wdst.amhocn.org/>).

During the inception of the NOCC, the need to investigate measures that might be used for consumers of specialist services was explicitly acknowledged. Furthermore, it was noted that the applicability of outcome measures designed for use in civil adult mental health settings should be evaluated in a forensic context, to ensure that they effectively capture the needs of this group (National Mental Health Working Group, 2003). This position was reiterated in a report issued by the Victorian Government entitled 'Because Mental Health Matters', which further placed focus on addressing the needs of consumers of specialist services (Department of Human Services, 2008). Within this report, the burgeoning demand for forensic psychiatric services within Australia was acknowledged, and it was also noted that a significant proportion of people within the criminal justice system experienced psychiatric difficulties (Ogloff, Davis, Rivers, & Ross, 2006). Amongst the goals for mental health service reform outlined in this report, was the need to obtain common assessment tools suitable for measuring the range of needs possessed by a forensic psychiatric population (Department of Human Services, 2008). However, in the most recent review of NOCC, it was reported by the National Mental Health Information Development Expert Advisory Panel (2013) that no such measure had been identified nor had an evaluation of the existing tools been undertaken within a forensic context. It was asserted that a clear gap remains in the measures employed for forensic services with respect to outcomes relating to risk, security and legal issues. The present study therefore seeks to address this gap in our knowledge.

Aims and hypotheses

The present study had three main aims. In the first instance, we aimed to investigate whether differences exist between HoNOS scores obtained by civil and

forensic mental health populations. That is, whether the mean scores obtained by both populations were different at each collection occasion (admission, 91-day review and discharge). It was hypothesised that civil and forensic populations would not show differences in mean HoNOS scores at the point of admission, however, it was anticipated that differences would emerge over the course of admission due to the longer period of care received by forensic consumers (hypothesis one).

Within several international jurisdictions, including Australia, the ability to compare ROMs across time and treatment setting is frequently cited as being an important feature of such tools (Pirkis, Burgess, Kirk et al., 2005). As patients may move between civil and forensic settings, if separate versions of the HoNOS were used in each environment (i.e. HoNOS and HoNOS-Secure), tracking progress across settings would only be possible if the two versions correlate strongly. Therefore, the second aim of this study was to evaluate the degree to which the HoNOS and the 'clinical and social functioning scale' of the HoNOS-Secure correlate with each other. This was investigated both at the item and total score level. It was hypothesised that there would be a high degree of concordance between the ratings on the HoNOS and the HoNOS-Secure 'clinical and social functioning scale', and the two scales would not be statistically different (hypothesis two).

Finally, regarding the gap in outcome measurement tools for forensic mental health patients, particularly across the domains of risk and security needs, the present study aimed to evaluate whether the HoNOS or HoNOS-Secure demonstrates better predictive validity on these factors. In the present context, predictive validity was considered in terms of a tool's ability to correctly categorise patients on measurable outcomes, namely: acuity, risk and freedom of movement. It was hypothesised that both the 'clinical and social functioning' and 'security scales' of the HoNOS-secure would more accurately categorise patients on these variables, and account for a greater amount of the overall variance, than that HoNOS alone (hypothesis three).

Methods

Setting and source population

The study was conducted at the Thomas Embling Hospital (TEH), the sole forensic mental health inpatient facility within the state of Victoria, Australia. The hospital provides secure care for up to 116 patients across seven wards. The wards are structured to encompass the spectrum of patient recovery from acute care to community reintegration. All patients within the hospital are detained under involuntary treatment orders, broadly separated into two main categories: forensic patients, who have been found either unfit to stand trial or not guilty of an offence on the grounds of mental impairment; and security patients, who are prisoners requiring assessment or treatment for mental health disorder. A

small proportion of patients are also detained under civil involuntary hospitalisation orders.

Data collection and analysis

To investigate the aims of this study, clinical staff within TEH completed the HoNOS and HoNOS-Secure for all patients who consented to participate. Ratings were completed for patients on admission, discharge and every 91 days that they remained within the hospital, as per the NOCC protocol. All ratings were undertaken by mental health clinicians (e.g. psychiatric nurses, psychologists, occupational therapists and social workers) who had received training in the use of these tools to increase reliability of ratings (Rock & Preston, 2001). All HoNOS and HoNOS-Secure ratings were made in accordance to the rating manuals for these tools (e.g. Sugarman & Walker, 2007; Wing et al., 1998). On each rating occasion, the HoNOS and HoNOS-Secure were rated by separate clinicians (i.e. two clinicians were used at each collection occasion, with one clinician rating the HoNOS and the other rating the HoNOS-Secure), with ratings being based on the patient's presentation over the same two-week period. Data were also recorded regarding a patient's freedom of movement (restricted/unrestricted access to the campus), ward placement (residing on an acute/subacute unit) and number of risk incidents during the two-week rating period (aggression, self-harm and substance use).

Data collection occurred in two phases, with the initial phase occurring between 1 July 2010 and 1 January 2011. To evaluate the stability of findings over time, a second period of data collection occurred between 1 December 2014 and 1 May 2015.

To investigate the first aim, the mean HoNOS score obtained by the forensic sample was compared with the mean HoNOS score of all mental health consumers within the state of Victoria. Data for the state-wide sample were accessed via the wDST, using the following reference criteria: Jurisdiction: *Victoria*, Age Group: *Adult*, Service Setting: *Inpatient*, Financial Year: *July 2010–June 2011*. *Level of Analysis* was specified as *Collection Occasion*, to permit comparison of data collected on admission, 91-day review and discharge; as well as a global average across all collection occasions. It was observed that the forensic cohort within TEH was skewed heavily towards male consumers (85.7%). It was therefore uncertain if the female component of the sample would be representative of female civil mental health consumers generally. As such, the data obtained from the wDST was further restricted to *male consumers*, and only the male portion of the forensic sample was used. Likewise, as the HoNOS was designed for use with 'working age adults', the sample was restricted to consumers aged 18–65. The remaining variables of diagnosis and legal status were set to *All*. Data obtained via the wDST were described in terms of sample size, mean scores and standard deviation. Comparison of mean scores generated by the civil and forensic

samples was undertaken using two-tailed t-tests. To investigate the effect size of any difference observed between the two means, Cohen's d statistics were generated post hoc to provide a standardised measure of similarity between the two means.

To investigate the second aim, Pearson correlations were generated for item pairs between the HoNOS and HoNOS-Secure (clinical and social functioning scale), using data generated from the forensic mental health sample. Cohen's d statistics were generated post hoc to further evaluate any difference observed.

To investigate the question of whether the HoNOS-Secure performs equal to or better than the HoNOS within forensic mental health settings in terms of its predictive validity, a series of logistic regression analyses was performed. Three dependent variables were used as markers of mental health acuity and risk: ward placement (i.e. whether the participant resided on an acute or subacute unit during the period of review), freedom of movement status (i.e. whether the participant had restricted or unrestricted access to the hospital campus) and risk incidents (i.e. occurrence of aggression, self-harm and/or substance use).

In all cases, each of the three HoNOS components (i.e. the original HoNOS and the 'clinical and social functioning' and 'security' scales of the HoNOS-Secure) was employed as independent variables and entered together as one block into the regression analysis. Standardised beta weights for each scale were examined to determine their relative contribution to the classification of patients on the dependent variables. To investigate the stability of results over time, this analysis was repeated with data obtained from a second sample of patients, collected three years after the initial sample. Finally, a post hoc investigation was undertaken, in which the HoNOS-Secure 'security scale' was combined with the HoNOS and a further regression was conducted using the HoNOS-Secure (clinical and security scales) and the HoNOS with security scales added.

All statistical analyses were performed using Statistical Package for Social Sciences (SPSS) software (version XX, SPSS, Inc, Chicago, IL, USA).

Results

Sample characteristics

At the conclusion of the initial phase of data collection (June 2010–January 2011), 253 HoNOS-Secure assessments had been completed for the forensic mental health sample. Of these, 39 occurred on admission, 195 on review and 16 at discharge. As detailed in Table 2, most patients within the sample were male ($n = 217$, 85.7%), with assessments occurring fairly evenly across the acute ($n = 135$) and subacute ($n = 118$) units, 53.4 and 46.6%, respectively. Of the patients for whom a HoNOS-Secure was completed, additional data regarding HoNOS scores, risk incidents, freedom of movement and acuity were only available in 202 cases. Of these, 170 were male (84.1%), 89 acute and 113 subacute

Table 2. Sample and hospital population.

	Initial data collection		Repeated data collection	Hospital population <i>n</i> (%)
	HoNOS-Secure only <i>n</i> (%)	HoNOS-Secure and additional variables ^a <i>n</i> (%)	HoNOS-Secure and additional variables ^a <i>n</i> (%)	
Total	253 (100)	202 (100)	50 (100)	118 (100)
Male	217 (85.7)	170 (84.2)	37 (83)	100 (84.7)
Female	36 (14.2)	32 (15.8)	13 (16)	18 (15.3)
Acute unit	135 (53.4)	89 (44.1)	60 (50.8)	60 (51.0)
Subacute unit	118 (46.6)	113 (55.1)	58 (49.2)	58 (49.0)

^aAdditional variables: HoNOS, freedom of movement, ward acuity and risk incidents.

Table 3. Mean scores state wide and forensic populations.

	Admission μ (SD, <i>n</i>)	Review μ (SD, <i>n</i>)	Discharge μ (SD, <i>n</i>)	Total μ (SD, <i>n</i>)
Civil	15.9 (6.8, 4754)	14.1 (7.5, 142)	7.7 (5.3, 3920)	12.2 (7.4, 8816)
Forensic	13.6 (7.20, 31)	5.2 (4.9, 167)	1 (.93, 16)	6.4 (5.9, 213)
<i>P</i>	.0609	.0001	.0001	.0001
<i>Cohen's d</i>	.311	1.393	1.760	.866

(44.1 and 55.1%, respectively). At the time of data collection, the total patient population of TEH was 118, of which 100 were male (84.7%), with the hospital being divided into 60 acute beds and 58 subacute beds (51 and 49%, respectively). As such, it was considered that the sample obtained was representative of the hospital population.

To investigate the stability of results over time, a second period of data collection occurred between 1 December 2014 and 1 May 2015. In total, 50 additional sets of data were generated, with each comprising HoNOS, HoNOS-Secure, acuity, freedom of movement and risk incident information. Of these participants, 37 (83%) were male, 60 (50.2%) acute and 58 (49.8%) subacute.

Comparison of mean HoNOS scores between forensic and civil mental health patients

Interrogation of the wDST revealed for the period July 2010–June 2011 a total of 8816 HoNOS assessments was conducted in adult inpatient mental health services throughout Victoria for consumers matching the criteria specified. Of these, 4754 occurred on admission, 142 on review, and 3920 at discharge. The mean scores and standard deviations for each period are reported in Table 3 for both the civil and forensic populations (males aged 18–65 only).

Comparison of these mean scores using two-tailed t-tests indicated that the scores obtained by the forensic and civil populations on *admission* were not statistically different ($p = .06$). However, all other means were found to be significantly different from each other (i.e. review, discharge and total population

Table 4. Correlations between HoNOS and HoNOS-Secure items/total score with effect sizes.

Item number	N	HoNOS μ (SD)	HoNOS-Secure μ (SD)	r	p	d
Total score	202	6.65 (6.24)	7.58 (5.89)	.818	.001	.153
Item 1	202	.54 (.96)	.55 (.87)	.730	.001	.011
Item 2	202	.11 (.54)	.19 (.61)	.549	.001	.129
Item 3	202	.18 (.68)	.19 (.63)	.582	.001	.015
Item 4	202	.62 (.90)	.67 (.90)	.633	.001	.055
Item 5	202	.60 (.91)	.89 (1.00)	.669	.001	.293
Item 6	202	1.12 (1.30)	1.19 (1.28)	.755	.001	.054
Item 7	202	.50 (.80)	.63 (.79)	.581	.001	.159
Item 8	202	.72 (1.12)	.80 (1.10)	.479	.001	.068
Item 9	202	1.03 (1.14)	1.18 (1.02)	.492	.001	.138
Item 10	202	.72 (1.05)	.84 (1.01)	.547	.001	.115
Item 11	202	.09 (.44)	.15 (.46)	.341	.001	.134
Item 12	202	.40 (.84)	.33 (.67)	.281	.001	.092

mean; $p < .01$). Moreover, post hoc analysis of effect sizes via *Cohen's d* indicated that the effect size at admission was small ($d = .31$), but large at all other occasions (i.e. review, discharge and total sample mean were 1.39, 1.76 and 0.87, respectively). This finding supported hypothesis one.

Correlation of HoNOS and HoNOS-Secure total score and items

As shown in Table 4, full scale scores generated by each tool were observed to correlate strongly ($r = .81$). Moreover, mean total scores generated by the HoNOS and HoNOS-Secure 'clinical scale' were found not to be statistically different (HoNOS: $\mu = 6.65$, $SD = 6.24$; HoNOS = Secure clinical scale: $\mu = 7.58$, $SD = 5.89$; $p = .13$). Post-hoc analysis via *Cohen's d* indicated that the magnitude of the effect size between these mean total scores was small ($d = .15$). These findings supported hypothesis two. In addition, each of the 12 item pairs also demonstrated significant correlations in the expected direction ($p = .001$). However, the strength of the correlations varied across items, ranging from 0.28 to 0.76. Effect sizes for each item pair were also small.

Predictive ability of HoNOS and HoNOS-Secure

As demonstrated in Table 5, the HoNOS and HoNOS-Secure 'security scale' showed a statistically significant relationship to the dependent variables when investigated via logistic regression. However, the HoNOS-Secure 'clinical scale' was not found to contribute significantly to the correct classification of patients for each of the dependent variables. In each case, on the variables of ward acuity, freedom of movement and any risk incidents, a model containing the HoNOS and HoNOS-Secure 'security scale' best predicted classification of patients at 78.7, 86.6 and 79.2%, respectively.

The above findings only partially supported hypothesis three, in which it was anticipated that the HoNOS-Secure would more accurately categorise patients

Table 5. Logistic regression (initial analysis).

	β	Wald (χ^2)	p	OR	Correctly classified (%)	R^2
<i>Ward acuity</i>						
HSsec	.365	26.023**	.000	1.440	78.7	.541
HSclin	.064	1.374	.241	1.066		
HoNOS	.129	6.084*	.014	1.138		
(Constant)	-4.245	49.243	.000	.014		
<i>Freedom of movement</i>						
HSsec	.153	4.674**	.031	1.165	86.6	.337
HSclin	.065	1.314	.252	1.068		
HoNOS	.099	3.822**	.050	1.105		
(Constant)	-4.787	45.269	.000	.008		
<i>Any risk incidents</i>						
HSsec	.165	8.540**	.003	1.180	79.2	.241
HSclin	-.035	.550	.458	.965		
HoNOS	.118	7.040**	.008	1.125		
(Constant)	-3.211	43.804	.000	.040		

Notes: β = beta weight, S.E. = Standard error, χ^2 = Chi square, df = degrees of freedom, p = significance of result, OR = odds ratio, R^2 = Nagelkerke R squared, * $p < .10$, ** $p < .05$, HSsec = HoNOS-Secure (security scale), HSclin = HoNOS-Secure (clinical scale).

Table 6. Logistic regression (repeated analysis).

	B	Wald (χ^2)	p	OR	Correctly classified (%)	R^2
<i>Ward acuity</i>						
HSsec	.504	6.654**	.010	1.656	90.0	.768
HSclin	-.253	1.906	.167	.776		
HoNOS	.451	4.138**	.042	1.570		
(Constant)	-4.745	10.863	.001	.009		
<i>Freedom of movement</i>						
HSsec	.854	4.711**	.030	2.349	92.0	.818
HSclin	-.141	.371	.542	.868		
HoNOS	.214	.892	.345	1.239		
(Constant)	-10.347	6.430	.011	.000		
<i>Any risk incidents</i>						
HSsec	.153	1.799	.180	1.166	80.0	.254
HSclin	.077	.483	.487	1.080		
HoNOS	-.059	.207	.649	.943		
(Constant)	-3.261	10.099	.001	.038		

Notes: β = beta weight, S.E. = Standard error, χ^2 = Chi square, df = degrees of freedom, p = significance of result, OR = odds ratio, R^2 = Nagelkerke R squared, * $p < .10$, ** $p < .05$, HSsec = HoNOS-Secure (security scale), HSclin = HoNOS-Secure (clinical scale).

than that HoNOS alone. Specifically, the results suggested the HoNOS-Secure 'clinical and social functioning scale' did not perform better than or equal to the original HoNOS scale. As such, the analysis was replicated using a second set of data to test whether or not this result was consistent over time and with a separate cohort of patients. As demonstrated in Table 6, the second data-set produced analogous results to the original cohort, suggesting this finding was indeed stable and repeatable. The HoNOS and HoNOS-Secure 'security scale' continued to account for the greatest amount of variance in the data. With the exception of the variable *any risk incident*, these relationships were significant

Table 7. Logistic regression (combined scales).

	<i>B</i>	Wald (χ^2)	<i>p</i>	OR	Correctly classified (%)	<i>R</i> ²
<i>Ward acuity</i>						
HoNOS + HSsec	.167	4.948**	.026	1.181	79.2	.493
HoNOS-Secure (C + S)	.059	.620	.431	1.060		
(Constant)	-3.368	28.127	.000	.034		
<i>Freedom of movement</i>						
HoNOS + HSsec	.098	3.894**	.048	1.103	86.6	.337
HoNOS-Secure (C + S)	.062	1.513	.0219	1.064		
(Constant)	-4.812	49.935	.000	.008		
<i>Any risk incident</i>						
HoNOS + HSsec	.127	8.541**	.003	1.136	79.2	.234
HoNOS-Secure (C + S)	-.015	.113	.737	.986		
(Constant)	-3.032	46.168	.000	.048		

Notes: β = beta weight, S.E. = Standard error, χ^2 = Chi square, *df* = degrees of freedom, *p* = significance of result, OR = odds ratio, *R*² = Nagelkerke *R* squared, **p* < .10, ***p* < .05, HoNOS-Secure (C + S) = HoNOS-Secure clinical and security scales, HSsec = HoNOS-Secure (security scale).

and in the expected direction. By contrast, the HoNOS-Secure 'clinical scale' demonstrated weak, non-significant associations. Using the HoNOS and HoNOS-Secure 'security scale', patients were correctly classified on the variables of ward acuity, freedom of movement and risk incidents at 90, 92 and 80%, respectively.

To confirm these findings, a post hoc investigation was conducted, in which the HoNOS-Secure 'security scale' was combined with the HoNOS and a further hierarchical regression was conducted using the HoNOS-Secure (clinical and security scales) and 'HoNOS + security scales'. This was undertaken to directly compare the performance of the HoNOS/HoNOS-Secure, if the security scale was added to either version of this tool. As demonstrated in Table 7, the same pattern of results was observed. For all three dependent variables, a combination of the HoNOS and HoNOS-Secure 'security scale' proved the most effective model, producing significant relationships in the expected direction; whereas, the HoNOS-Secure total scale (i.e. clinical and security scales) demonstrated a non-significant relationship.

Discussion

The present study investigated the ability of a commonly used mental health outcome measure to monitor mental health and security needs of a forensic inpatient population. Specifically, the original version of the HoNOS was compared with the *forensic* adaptation of this tool, known as the HoNOS-Secure.

The first aim of the study was to establish whether the mental health needs of a forensic population were demonstrably different to those of civil psychiatric patients when evaluated by the HoNOS. As noted, it has been observed by several authors that consumers of forensic mental health services typically remain in secure care far longer than their mainstream counterparts, and may even remain in the absence of mental health difficulties (e.g. Shinkfield & Ogloff, 2014;

Turner & Salter, 2008). As anticipated, the results confirmed that when the mean HoNOS scores for civil and forensic populations were compared at the point of admission, the mean scores of the two populations were not statistically different. However, over the course of admission, clear differences emerged between the groups, with post hoc analysis indicating a large effect size between the two populations. To the authors' knowledge, this finding has not previously been demonstrated with respect to HoNOS scores.

In the field of outcome measurement, the ability to track and evaluate factors pertinent to a client group's recovery is paramount. Forensic mental health consumers typically remain in secure care for a period of time dictated by their level of risk and security needs. Therefore, monitoring these needs is an important aspect of outcome measurement for this population (Sugarman et al., 2009).

When both versions of the HoNOS were used in the present study to monitor the same forensic population, the total scores of the HoNOS and HoNOS-Secure 'clinical and social functioning scale' did indeed correlate strongly ($R = .82$, $p = .001$). This suggested that both versions vary in a systematic way according to the functional and clinical difficulties experienced by a consumer. Moreover, the mean scores generated by the HoNOS and HoNOS-Secure 'clinical and social functioning' scale were not statistically different, and the effect size observed between the two scales was low, suggesting that if the total score was considered alone, either version of the HoNOS could be used to obtain analogous results. However, at an item level, individual item pairs within the different versions of the tool demonstrated broad variation in the degree to which they correlated. Items that correlated most strongly were Item 1 (*Overactive, aggressive, disruptive or agitated behaviour*; $r = .73$) and Item 6 (*Problems associated with hallucinations and delusions*; $R = .755$). Whereas the weakest relationships were observed with Item 11 (*Problems with living conditions*; $R = .34$) and Item 12 (*Problems with occupation and activities*; $R = .28$). All other item pairs produced moderate correlations between 0.48 and 0.58. Despite this, effect sizes between item pairs suggested the overall impact on mean scores was small.

On examining of the wording changes made to each of the item pairs, it did not appear that there was any systematic relationship between the extent to which changes had been made and the strength of the relationship between the item pair. While it was noted that poorly correlated items (e.g. 11 and 12) contained extensive adaptations, so too did several of the more strongly correlated items (e.g. 1 and 3). It was also noted that previous research had identified items 11 and 12 as being particularly problematic with respect to their reliability and validity for other inpatient samples (Pirkis et al., 2005), which may account for these weaker relationships. It might also be suggested difficulties may emerge in these items as a function of forensic patients receiving a longer period of care than their civil counterparts. That is, as admission lengths for civil patients rarely extend beyond a few weeks, items such as item 9 (relationships), item 11 (living conditions) and 12 (occupation and activity) are generally rated in relation

to their home-based relationships and community environment. However, as forensic patients may remain within a secure setting for years, these same items within a forensic service could arguably be considered less of a measure of individual outcome, but rather provide a reflection of service provision and the opportunities available to the consumer within their restricted environment. This may also go some way to explaining the lack of difference in HoNOS scores between civil and forensic patients at the point of admission, as well as the increasing divide in scores over time.

Taken together, there are two key considerations that can be derived from these data. Firstly, at a global level, both versions of the tool appear to identify a similar level of clinical and social impairment. Secondly, it appears that individual items of the HoNOS-Secure 'clinical scale' contribute variably to the total score, to a greater or lesser extent than they did in the original version of this tool. This may suggest that the adaptations made to individual items altered the way these items are interpreted, resulting in the two versions not being analogous. This is unlikely to matter if the HoNOS-Secure was used as the only version of this tool within a service, with results only being compared with other HoNOS-Secure data. However, if a patient moved between a forensic and civil psychiatric setting, it may not be possible to compare HoNOS/HoNOS-Secure item scores across settings/time.

The final component of the present study was to examine the extent to which the HoNOS and HoNOS-Secure clinical/security scales were able to accurately classify the needs of forensic mental health patients. Overall, the logistic regression analyses produced consistent results across all three outcome variables (e.g. ward placement, freedom of movement, risk incidents). In each case, the strongest model was a combination of the original HoNOS and the 'security scales' of the HoNOS-Secure. Interestingly, in all cases, the HoNOS-Secure 'clinical and social functioning scale' was found to contribute little to the overall classification of patients and was observed to 'drop out' of the model. This was contrary to the hypothesised result, in which it was anticipated that the HoNOS-Secure 'clinical and social functioning scale' would outperform the HoNOS in its ability to correctly classify patients on a number of clinical and risk-related needs domains. However, this result remained stable over time and was replicated with a second cohort of forensic mental health patients; some four years after the initial data-set was collected.

There are a number of ways to consider this finding. Perhaps the most parsimonious explanation would be that this represents a true difference in the performance of the two tools. In this instance, it could be said that the original HoNOS performs better than the HoNOS-Secure in forensic settings, particularly when it is used in combination with the 'security scale'. However, it should also be noted that when the 'security scale' was combined with both versions of the HoNOS, even though the HoNOS-Secure 'clinical and social functioning scale' relationship was found to be non-significant, the difference in odds ratios

between the two forms of the tool was small (e.g. ward security $\Delta OR = .121$; freedom of movement $\Delta OR = .039$; risk incident $\Delta OR = .150$; Table 7). In the earlier part of this study, it was observed that the two forms of the tool correlated strongly at a total item level. As such, it may be that the HoNOS performs only marginally better than the HoNOS-Secure, but when both were forced into the regression model, there was a little variance remaining in data to be explained by the HoNOS-Secure 'clinical and social functioning scale' that had not already been accounted for by the HoNOS. That is, in this scenario, either tool could conceivably be used.

Finally, it might be considered that the results obtained may have been influenced by some systematic difference in the way the clinicians who performed the ratings used the tools or interpreted the items. This possibility will be considered further in the limitations section below.

Limitations

The limitations within the present study should be acknowledged. Most significantly, the cohort of forensic mental health patients upon which this study was based was obtained from only one service setting. Therefore, it was not possible to state with certainty that the findings would generalise to other forensic psychiatric facilities. It was also noted that within the forensic setting used, the HoNOS was routinely employed as a measure of patient outcome. As such, staff were already familiar with this tool; whereas, the HoNOS-Secure was new to many staff. Despite being trained to use the HoNOS-Secure for the purpose of this study, this comparative lack of familiarity may have influenced ratings. It might be further hypothesised that staff within the forensic setting have learned to adapt or interpret the wording of HoNOS items in a manner that enables them to rate these items in a secure environment. However, any 'reinterpretation' of items is likely to have been a non-explicit process, without formal operationalisation of anchor points; as occurred when developing the HoNOS-Secure. Exploring the question of whether/how clinicians adapt or interpret tools to 'fit' with their own service may be a fruitful source of enquiry in the future.

Regarding the comparison of HoNOS scores obtained by consumers of civil and forensic mental health services, as raw data were only obtained from a forensic sample, it was not possible to compare these cohorts directly. Therefore, analysis relied on comparison data obtained via a reporting tool which consolidates and reports on the data of all consumers within the state of Victoria. It is acknowledged that this state-wide sample would have also included the data generated by the cohort of forensic patients used within the study (although they would have represented only a very small fraction of ratings, approximately 2% of the sample). Due to the way these data are reported by the wDST, it was not possible to disentangle these two groups. However, any contamination

of the state-wide sample by data from forensic consumers would have been negligible and unlikely to have affected the results overall.

Finally, it is also acknowledged that the 'real life outcomes' employed in the logistic regression analysis were proxies for the gamut of outcomes that are pertinent to forensic mental health consumers. Moreover, no standardised measure of mental health and/or social functioning was used to provide a metric against which to test the convergent/predictive validity of these tools. Rather the present study relied on three variables that could be observed for patients and were conceptualised as being related to a patient's progression towards discharge (e.g. ward placement, freedom of movement, and risk incidents). With respect to the variable 'ward placement', within the hospital transfer between acute/subacute units is based primarily on a patient's recovery from mental health difficulties. As such, ward placement was considered to reflect the acuity of a patient's mental health difficulties, rather than level of risk they present to themselves or others. However, if a greater sample size was available, a more fine-grained analysis would have been possible via linear regression techniques. In this way, acuity could be examined across multiple stages in a patient's journey through the hospital (e.g. acute, subacute, rehabilitation and community integration units). Conversely, a patient's freedom of movement was considered to reflect a combination of the potential risks a patient poses to themselves/others, as well as their capacity to navigate the social environment within the hospital grounds. Finally, the presence of risk incidents was considered a direct measure of behaviour requiring specific risk management strategies. Overall, it is acknowledged that there are limitations in this study, as the authors were unable to look at all variables relevant to outcome and discharge from a secure forensic setting. However, despite these confounding factors, it is noted that the study still generated significant results. As such, it is recommended that further investigation be conducted utilising a broader range of factors and outcomes pertinent to the mental health and forensic needs of this population.

Conclusion

A number of conclusions can be drawn from the present study. Firstly, this study supports the notion that forensic and civil mental health service users present with comparable levels of clinical and social functioning at the point of admission. However, over the course of an admission, as clinical and social difficulties abate, differences in the risk and security needs of these two groups become more apparent. For the forensic group, consumers can remain in psychiatric care even following amelioration of acute mental health difficulties since legislation that detains them requires that they only be discharged to the community when they no longer represent a 'serious endangerment' to the community (e.g. *Crimes (Mental Impairment and Unfitness to be Tried) Act*, 1997). As such, it is imperative

that ROMs for forensic consumers take into account risk and security needs in addition to clinical factors and functional impairment.

With respect to the use of the HoNOS family of tools as outcome measures in forensic settings, the findings of the present study suggest several things. Firstly, both the HoNOS and HoNOS-Secure appear to measure clinical and functional disability to a comparable extent. However, individual differences amongst the items pairs of the HoNOS/HoNOS-Secure make it difficult to compare these tools directly at an item level. Moreover, when examined via logistic regression, the HoNOS outperformed the HoNOS-Secure 'clinical and social functioning scale' in correctly classifying patients on a range of real-life outcomes. Yet, the difference between the tools was relatively minor and may not be significant enough to reject the use of the HoNOS-Secure in preference for the HoNOS outright. Where this may be of significance could be within jurisdictions in which patients move between forensic and civil mental health settings and there is concern about the ability to directly compare the results from these measures over time. In this instance, it is suggested that the HoNOS be retained in preference of the HoNOS-Secure 'clinical scale', as there does not appear to be significant benefit from implementing the HoNOS-Secure 'clinical and social functioning scale' over the original HoNOS. This would be a particular consideration, if there was a significant cost associated with altering the data collection and reporting infrastructure to enable the HoNOS-Secure to be included.

Despite the lack of favourable outcome for the 'clinical and social functioning scale' of the HoNOS-Secure, the 'security scale' was observed to add significant incremental validity to case classification for patients in a forensic hospital. Regardless of which version of the HoNOS is used in secure settings, based on the findings generated by the present study, the addition of the HoNOS-Secure 'security scale' is recommended. This echoes comments by the authors of the HoNOS-Secure, who have previously suggested it may be possible to combine the 'security scale' with other versions of the HoNOS as required (Sugarman et al., 2009). The present study provides empirical support for this suggestion.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Gregg Shinkfield  <http://orcid.org/0000-0003-1131-5369>

References

- Burgess, P., Pirkis, J., & Coombs, T. (2015). Routine outcome measurement in Australia. *International Review of Psychiatry*, 27, 264–275.

- Davoren, M., Byrne, O., O'Connell, P., O'Neill, H., O'Reilly, K., & Kennedy, H. G. (2015). Factors affecting length of stay in forensic hospital setting: need for therapeutic security and course of admission. *BMC Psychiatry*, 15, 301–316.
- Department of Human Services. (2008). *Because mental health matters: A new focus for mental health and wellbeing in Victoria, Consultation paper*. Melbourne: Victorian Government Department of Human Services.
- Dickens, G., Sugarman, P., Picchioni, M., & Long, C. G. (2010). HoNOS-Secure: Tracking risk and recovery for men in secure care. *The British Journal of Forensic Practice*, 12, 36–46.
- Dickens, G., Sugarman, P., & Walker, L. (2007). HoNOS-Secure: A reliable outcome measure for users of secure and forensic mental health services. *Journal of Forensic Psychiatry & Psychology*, 18, 507–514.
- Kennedy, H. G., O'Neill, C., Flynn, G., & Gill, P. (2010). *Dangerousness understanding, recovery and urgency manual (the DUNDRUM quartet) v1.0.22*. Dublin: National Forensic Mental Health Service
- Long, C. G., Dickens, G., Sugarman, P., Craig, L., Mochty, U., & Hollin, C. R. (2010). Tracking Risk profiles and outcome in a medium secure service for women: Use of the HoNOS-Secure. *International Journal of Forensic Mental Health*, 9, 215–225.
- National Mental Health Information Development Expert Advisory Panel. (2013). *Mental health national outcomes and casemix collection: NOCC strategic directions 2014–2024*. Canberra: Commonwealth of Australia.
- National Mental Health Working Group. (2003). *Mental Health National Outcomes and Casemix Collection: Technical Specification of State and Territory Reporting Requirements for the Outcomes and Casemix Components of 'Agreed Data', Version 1.50*. Canberra: Commonwealth Department of Health and Ageing.
- Ogloff, J., Davis, M., Rivers, G., & Ross, S. (2006). *The Identification of Mental Disorders in the Criminal Justice System*. Canberra: Australian Institute of Criminology.
- Ogloff, J., Lemphers, A., & Dwyer, C. (2004). Dual diagnosis in an Australian forensic psychiatric hospital: Prevalence and implications for services. *Behavioral Sciences and the Law*, 22, 543–562.
- Ogloff, J. R. P., Talevski, D., Lemphers, A., Simmons, M., & Wood, M. (2015). Co-occurring mental illness, substance use disorders, and antisocial personality disorder among clients of forensic mental health services. *Psychiatric Rehabilitation Journal*, 38, 16–23.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005). *Routine measurement of outcomes in Australia's public sector mental health services*. Melbourne: The University of Melbourne.
- Pirkis, J., Burgess, P., Kirk, P., Dodson, S., & Coombs, T. (2005). *Review of standardised measures used in the National Outcomes and Casemix Collection (NOCC)*. Sydney: NSW Institute of Psychiatry.
- Rock, D., & Preston, N. (2001). HoNOS: Is there any point in training clinicians? *Journal of Psychiatric and Mental Health Nursing*, 8, 405–409.
- Shaw, J. (2002). Needs assessment for mentally disordered offenders is different. *The Journal of Forensic Psychiatry*, 13, 14–17.
- Shinkfield, G., & Ogloff, J. (2014). A review and analysis of routine outcome measures for forensic mental health services. *International Journal of Forensic Mental Health*, 13, 252–271.
- Shinkfield, G., & Ogloff, J. (2015). Use and interpretation of routine outcome measures in forensic mental health. *International Journal of Mental Health Nursing*, 24, 11–18.
- Sugarman, P., & Everest, H. (1999). An audit of five regional secure psychiatric services: South Thames NHS Multi-authority Clinical Audit Group.

- Sugarman, P., & Walker, L. (2007). Health of the Nation Scale for Users of Secure/Forensic Services (version 2b). Retrieved from <http://www.rcpsych.ac.uk/researchandtrainingunit/honos/secure.aspx>
- Sugarman, P., Walker, L., & Dickens, G. (2009). Managing outcome performance in mental health using HoNOS: Experience at St Andrew's Healthcare. *Psychiatric Bulletin*, 33, 285–288.
- Turner, T., & Salter, M. (2008). Forensic psychiatry and general psychiatry: Re-examining the relationship. *Psychiatric Bulletin*, 32, 2–6.
- Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *The British Journal of Psychiatry*, 172, 11–18.

PART D: DISCUSSION

Chapter Nine: Integrated discussion

9.1 Overview of the Research

Within the field of forensic mental health, a clear gap exists regarding the availability and use of routine outcome measurement (ROM) tools designed to monitor the clinical, psychosocial and forensic needs of mentally disordered offenders. Although such tools have been employed within civil mental health populations for several decades, and indeed these tools have typically demonstrated a sound ability to monitor the clinical and social functioning needs of mainstream client groups (Pirkis et al., 2005b), the efficacy of these tools with forensic populations has not yet been evaluated. Moreover, the tools currently mandated for use in Australia with forensic mental health populations do not encompass the range of additional needs pertinent to this population, particularly those relating to a client's risk, security and legal issues (National Mental Health Information Development Expert Advisory Panel, 2013). Despite this, forensic mental health services across Australia continue to be required to complete and report on this set of measures.

The potential limitations of these tools were recognised by the Australian Mental Health Outcome Classification Network during the development of the National Outcomes Classification Collection (NOCC). As such, the need to identify and evaluate ROM tools suitable for forensic psychiatric populations was noted (Department of Human Services, 2008). However, to date this evaluation has not occurred. The present thesis was therefore conceived as a means of answering the call from the Australian government to identify, evaluate and compare a range of forensic focused ROM tools against the existing NOCC suite of measures. Within this context, this thesis sought to evaluate the efficacy of the

ROM tools that are currently mandated for use in forensic mental health services. In turn, this thesis then sought to identify and evaluate a range of alternative outcome measures that had been designed for use with forensic mental health populations, and determine which tool or tools were most suitable for this task.

In addition to evaluating forensic and non-forensic ROMs, two further aims were addressed within this body of work. The first of these being to determine whether forensic and civil mental health patients demonstrate differences in their clinical needs when evaluated with the NOCC tools. Secondly, to determine the extent to which the needs of a forensic mental health population are heterogeneous, and moreover, how they might change over the course of an admission to a secure facility.

The goals of this thesis were achieved by completing a body of research to address three key objectives. The first objective was to review the existing literature pertaining to mental health outcome measures to identify tools that could potentially be used to monitor the clinical, social and risk associated needs of patients in a forensic mental health environment. The impetus for this component of the study was the recognition that few forensic ROM tools had been available in the late 1990's when the National Outcomes Casemix Collection suite of measures was first compiled (Chambers et al., 2009). However, due to the increased clinical focus on this population since that time, it was anticipated that new tools were likely to have been developed for this purpose over the ensuing two decades. It was further considered that if newly developed tools were identified by the present study, such tools might feasibly be included within the NOCC suite of tools. As such, tools identified from the existing literature were subsequently evaluated against a set of seven criteria specified by the Australian government. These

criteria had been stipulated as the parameters that such tools were required to meet in order to be included in the NOCC suite of measures (Burgess et al., 2010). Specifically, these criteria stated that any ROM tool included in the NOCC suite should:

- Explicitly measure domains related to functioning;
- Be brief and easy to use;
- Yield quantitative data;
- Have been scientifically scrutinized and used in two or more peer reviewed studies;
- Be applicable to the local jurisdiction;
- Be applicable for both inpatient and outpatient environments; and
- Demonstrate sound psychometric properties (internal consistency, validity, reliability and sensitivity to change).

Three hypotheses were generated in relation to objective one, namely:

Hypothesis 1: It was hypothesised that new forensic mental health focused outcome measures tools would have appeared in the literature during the two decades since the NOCC suite of ROMs was developed.

Hypothesis 2: That the majority of forensic specific ROM tools would focus on the evaluation and monitoring of risk, rather than mental health and general functioning.

Hypothesis 3: Based on previous reviews of outcome measurement tools for inclusion in the Australian NOCC suite of measures, it was hypothesised that a small proportion of the tools identified from this review will meet NOCC inclusion criteria.

The second objective of this thesis was to evaluate the reliability and validity of the currently mandated ROM tools being employed in forensic mental health settings. This objective sought to investigate concerns that had been expressed by a number of authors regarding potential difficulties that may arise when using these particular tools with a forensic mental health population (e.g., National Mental Health Working Group, 2003; Shinkfield & Brennan, 2010; Owens, 2010). Specifically, this thesis aimed to address the question of whether these tools were indeed limited in their utility when used with forensic mental health populations, or whether the reported difficulties may in fact be a function of the way they are being used in such environments (i.e., whether they were being completed in a reliable and valid manner).

Three hypotheses were generated in relation to objective two, namely:

Hypothesis 4: That analysis of NOCC outcome measures completed within TEH would reveal reporting rates to be below the 85% compliance target set by the Department of Health and Ageing (2009).

Hypothesis 5: Based on the findings of previous research, (e.g., Pirkis et al., 2005) it was hypothesised that a high degree of inter-rater reliability would be observed between clinician and auditor ratings in the research sample.

Hypothesis 6: That the use of a standardised protocol may assist in the training or provision of feedback for clinicians using ROM tools as part of their routine clinical practice.

The third and final objective of this thesis was quite broad and sought to answer a number of interrelated questions pertaining to the clinical, social and risk related needs of forensic mental health clients. On the back of these investigations, the main component of objective three was to evaluate which tool or tools would be most useful in capturing the needs of this population in a valid and reliable manner.

Objective three was initially addressed by reviewing the extant literature regarding forensic mental health populations to explore and describe the needs of such clients and, moreover, to consider whether these needs differ from non-forensic client groups. Secondly, an investigation was conducted to investigate how the needs of forensic mental health patients change over the course of their admission. It was considered that if change was observed across a number of domains that are not currently captured by the existing NOCC suite of tools, this would provide additional support for employing additional measures that capture these broader areas of clinical and forensic need. The third subcomponent of objective three aimed to determine if the needs of a forensic mental health population were homogenous, or whether sub-groups of forensic clients might be observed when the population is examined as a whole.

Having concluded the above investigations, a number of forensic and non-forensic ROM tools were investigated to determine which best captured the needs of this population. To achieve this goal, a selected subset of forensic outcome measures that had been identified during objective one, were subsequently compared against the existing NOCC tools. In doing so, the thesis sought to determine whether there would be significant benefit in adopting these forensic specific measures in the place of, or in addition to, the existing NOCC suite of measures.

Three hypotheses were generated in relation to objective three, namely:

Hypothesis 7: That tools developed specifically for use with a forensic population will better differentiate forensic mental health patients at different stages of progress towards recovery and discharge than tools developed for use in civil mental health settings

Hypothesis 8: That clinical/social needs would be most prominent at the point of admission for forensic mental health patients, with forensic/security needs becoming the primary focus towards discharge.

Hypothesis 9: That when the needs of forensic mental health patients were examined collectively, distinct groups of patients would be identified amongst this cohort.

9.2 Overview of Main Findings

A number of key findings emerged from the thesis that will be briefly summarised in this section. This will be followed by a review of the findings obtained by each of the separate studies.

In the first instance, it was determined that at the time of reviewing the outcome measurement literature in 2011, six tools existed that could potentially serve as ROMs in forensic mental health settings. These tools were: the Camberwell Assessment of Needs (Forensic Version); DUNDRUM Quartet; Health of the Nation Outcome Scale for Users of Secure / Forensic Services; Illness Management and Recovery Scales; Mental Health Recovery Measure; and the Short-Term Assessment of Risk and Treatability. However,

following analysis of these six tools, it was noted that none was able to assess all domains of interest (clinical/social functioning, risk, recovery and placement pathway; see Chapter 2), nor did they all fully meet each of the criteria specified by the Australian government for inclusion in the NOCC suite. However, of these, the most promising candidates for further examination in the study setting appeared to be the CANFOR and HoNOS-Secure. Ideally, all six tools might have been included for evaluation; however given the scope of a doctoral thesis, it was not possible to do so.

The second key finding of this thesis was the observation that the ROM tools currently mandated for use in forensic mental health services across Australia were indeed able to be completed by forensic clinicians in a valid and reliable manner. However, limitations were also observed with the use of these tools, particularly with regards to those measures containing items that were strongly influenced by a client's environment (e.g., access to substances, living conditions, and access to meaningful occupations). It was also confirmed that none of these tools contained items that were conducive to monitoring outcomes pertinent to forensic/security needs.

The third key finding of this thesis demonstrated that when forensic and civil mental health populations were assessed via the HoNOS over the course of an admission, significant differences were observed in the level of clinical and social impairment between these two populations. While no statistically significant differences were observed between these two populations at the point of admission, the mean HoNOS scores diverged significantly over the course of admission. It was further observed that forensic clients typically remain in secure mental health care far longer than civil consumers (Ruffles, 2010; Turner & Salter, 2008), at times even in the absence of mental health

symptomatology. As such, given that the NOCC tools do not provide a means of monitoring the additional ‘forensic’ needs of this client group, this set of tools were considered limited in their capacity to track the full range of needs that represent the grounds upon which a client may remain in secure care (i.e., due to forensic and risk related needs).

With respect to forensic mental health patients as a whole, the present study demonstrated that the needs of this population were quite diverse; with patients within the study sample varying significantly in terms of their levels of clinical and risk need. As such, it was concluded that any tools used to track changes in this population would need to be sensitive to a full range of both clinical/social and forensic needs. Moreover, the needs of this population appeared to change over the course of admission, with a combination of clinical and risk needs being most prominent at the point of admission, but changing to focus on social functioning, life skills and broader forensic needs as they progressed towards discharge.

Finally, the thesis evaluated a subset of forensic measures identified from the review of ROM literature (objective one), against the existing NOCC tools currently employed in Australia. Amongst the findings from this component of the thesis, it was demonstrated that it may be possible to use a number of the extant ROM tools to track the needs of this client group. In particular the HoNOS, HoNOS-Secure and LSP-16 were found to be most effective at performing the task of monitoring changes pertinent to a client’s progress towards discharge in the forensic mental health sample.

Flowing from the above findings, a separate study subsequently focused on evaluating the forensic and non-forensic version of the HoNOS (i.e., *Working Age Adults Version* and the *HoNOS for Users of Secure Services*). The findings of this study demonstrated that at a total score level, the clinical scales of these tools were able to be used with some degree of interchangeability. However, at an individual item level, there was a large degree of variation in the extent to which items correlated with each other. This suggested that direct comparison of these tools at an item level may not be meaningful. Moreover, the key finding of this study demonstrated that when the HoNOS and HoNOS-Secure ('clinical' and 'security' scales) were compared via logistic regression, the HoNOS-Secure 'clinical scale' was not found to contribute significantly to the correct classification of patients in terms of their acuity, freedom of movement, or frequency of engaging in risk behaviours (i.e., interpersonal aggression, self-harm or substance use). Ultimately, it was concluded that there was no significant benefit in using the HoNOS-Secure in its entirety in place of the HoNOS. However, there was good evidence to suggest that combining the 'security scale' of the HoNOS-Secure with the original version of the HoNOS might provide the greatest ability to correctly classify patients on both their clinical and forensic/risk need.

9.3 Main Findings of Each Study

The following section explores the main findings of each study in greater detail than was presented above.

9.3.1 Study one

The first objective of this thesis was achieved by reviewing the extant outcome measurement literature to identify any tools that could potentially be applied as measures of *clinical/social functioning, risk, recovery* and *placement pathways* in a forensic mental health context. From the review of the literature, nineteen instruments were initially identified that were considered to be potentially relevant to forensic mental health populations. In most cases, these instruments had been developed in the years following the creation of the NOCC suite, which was consistent with hypothesis one of this thesis. Following detailed analysis of these nineteen instruments, it was concluded that although none of these tools assessed all domains of interest, nor did they each fully meet all seven of the inclusion criteria specified by the Australian government (described above and also in Chapter 2), six tools were considered to have potential utility as outcome measures for users of forensic mental health services. This confirmed hypothesis three. The instruments identified were:

- Camberwell Assessment of Needs: Forensic Version;
- DUNDRUM Quartet;
- Health of the Nation Outcome Scale for Users of Secure / Forensic Services;
- Illness Management and Recovery Scales;
- Mental Health Recovery Measure; and
- Short-Term Assessment of Risk and Treatability.

A subcomponent of objective one was also to review the literature regarding needs that are frequently demonstrated by forensic mental health patients. A full description of this literature was presented in chapters one and two of this thesis. However, in brief, it was found that users of forensic mental health services appear to present with not only the

mental health difficulties and functional impairments seen in civil mental health settings, but can also demonstrate any or all of the following: a history of criminal behaviour; violent or sexual offending; severe personality and behavioural disturbances; self-harm; and/or co-morbid substance use (Coid et al., 2001; Ogloff, Lemphers, & Dwyer, 2004; Ogloff et al., 2015). In addition, clinicians working with such clients are also frequently required to consider additional areas of need, including: level of security required to maintain the safety of their clients and other people; as well as the dangerousness and risk management needs required for this client group (Kennedy et al., 2010; Shaw, 2002). As such, outcomes for forensic mental health patients typically encompass a wide variety of problem areas, including those beyond narrowly defined mental health outcomes (Cohen & Eastman, 2000). It was also considered that offending behaviour within this population could arise from factors that may not be causally related to mental health difficulties. Rather, such behaviour may stem from criminological factors found within non-mental health forensic populations (Cohen & Eastman, 2000; Dickens, Sugarman & Walker, 2007). Given this, treatment within forensic mental health services therefore seeks not only to provide symptomatic relief from mental illness, but also amelioration of the additional risks that these clients present to themselves and others (Andreasson et al., 2014; Davoren et al., 2015; Mullen, 2006).

9.3.2 Study two

Objective two of the thesis sought to critically evaluate the two clinician rated tools currently contained within the NOCC suite, which are currently mandated for use in Australian forensic mental health settings (Chapter 6). Specifically, the Health of the Nation Outcome Scales (HoNOS) and the Life Skills Profile (16 item version; LSP-16)

were selected for this evaluation and their reliability in a forensic mental health context was scrutinised.

A subcomponent of objective two was to develop an audit protocol that could be used not only within the present study, but might also be translated into clinical practice as a means of monitoring and providing feedback to mental health clinicians' regarding the accuracy of their ROM assessments.

On the basis of the data obtained from objective two, it was concluded that despite the HoNOS and LSP-16 having been developed for use in civil mental health settings, they could indeed be reliability interpreted and rated in a forensic mental health context. This finding confirmed hypothesis five of the thesis. The ability of forensic mental health staff to use these tools in a reliable manner also likely contributed to clinicians completing these tools and meeting the reporting requirements of the NOCC protocol. In this manner, the findings of the present thesis did not support hypothesis four, in that reporting rates of the NOCC measures were in fact found to be higher than the 85% minimum standard specified by AMHOCN. However, due to the differences that typically exist between forensic and civil mental health populations noted above, a number of limitations were also identified regarding the use of these tools in a forensic mental health environment. As described, limitations arose with respect to the interpretation of items that are directly influenced by the environment, such as monitoring a client's substance related needs on the basis of how frequently they have drugs and alcohol over the rating period. It was also noted that these measures do not provide information regarding changes in treatment needs such as risk of harm to others, offending behaviour, or level of security required to maintain the patient's safety. As such, at the conclusion of the study, it was recommended that further

investigation be undertaken to ascertain whether other tools that had been developed specifically for use with forensic mental health populations could be effectively substituted in the place of the HoNOS and LSP-16 (see further directions section of this chapter).

Finally, in terms of objective two, it was also demonstrated that the audit protocol developed for this study was effective in evaluating the accuracy with which clinicians interpret and rate ROM items. To this end, it was observed that the HoNOS and LSP-16 could be reliably rated on the basis of file review by a senior clinician, which could in turn be used to provide feedback to the original assessing clinician regarding the accuracy of their ratings. This provided support for hypothesis six; however, further investigation will be required to confirm the utility of this protocol in clinical practice (see further directions section of this chapter).

9.3.3 Study three

The third objective of this thesis was to examine and evaluate a subset of the forensic ROM tools that had been identified during objective one. These tools were then evaluated against the currently mandated mental health tools, to determine which would demonstrate greater ability to identify and monitor the broad range of needs possessed by a forensic mental health population. This portion of the thesis was divided into two empirical studies (study three and study four). In both studies additional data was collected regarding a patient's 'real life outcomes' across the following domains: ward acuity, freedom of movement, and number of risk incidents accrued during the rating period. These markers were used as independent variables against which the forensic/non-forensic tools were examined.

Empirical study three of this thesis (Chapter 7) evaluated three forensic (HoNOS-Secure, CANFOR, LSI-R:SV) and three non-forensic (HoNOS, LSP-16, BASIS-32) tools for their clinical utility with a forensic mental health population. In addition, this study also investigated the degree to which the needs of this forensic population were heterogeneous, as well as the extent to which the needs of forensic clients differ over the course of treatment from admission to discharge.

With respect to the range of needs possessed by patients in a forensic mental health setting, whilst the majority of patients demonstrated high levels of clinical and risk related need (see Chapter 7), a significant proportion of clients had *high* levels of need in only *one* domain (i.e., either clinical or forensic/security needs). Furthermore, approximately a fifth of the population was considered to have low levels of need across both domains. This finding confirmed hypothesis nine, insofar as distinct groups of patients and patient needs were observed within a single cohort of forensic mental health patients. Moreover, these data highlighted the complex nature of patient composition within a forensic setting and the importance of using tools that are sensitive enough to detect a range of needs across both clinical and forensic/security domains to effectively target and monitor interventions for all patients.

It was further observed that the needs of clients at different points of admission differed markedly in terms of their focus of treatment. To this end, it was noted that patients at the acute/newly admitted end of the hospital typically presented with a greater degree of need in the clinical/mental health domain, whereas patients in the subacute and rehabilitation wards (who were closer to discharge) presented with greater needs in the

forensic and psychosocial functioning domains. This finding provided support for hypothesis eight.

The final component of study three evaluated three forensic and three non-forensic tools for their utility in monitoring the needs of a forensic mental health population. From the data generated, it was observed that, with the exception of the BASIS-32 and the 'Met Needs' subscale of the patient rated version of the CANFOR, the majority of mean scores obtained on these measures differed significantly for forensic patients at different levels of ward acuity. However, only the HoNOS, HoNOS-Secure ('clinical' and 'security' scales), and the LSP-16 were able to differentiate amongst patients across all three levels of acuity. The mean scores obtained by clients in the acute, subacute and rehabilitation wards did not overlap for the HoNOS, HoNOS-Secure and LSP-16. Whereas, the remaining tools (i.e., CANFOR and LSI-R:SV) demonstrated differences in the mean scores obtained by patients on the acute unit and the subacute/rehabilitation wards; but, neither tool was sensitive enough to detect differences between patients residing on the subacute and rehabilitation wards. As such, it was concluded that of those tools examined in this study, the HoNOS, HoNOS-Secure and LSP-16 were the most effective for differentiating between forensic mental health clients at different levels of acuity. This provided partial support for hypothesis seven, as one of the three tools that performed best as a measure of change for forensic clients had been developed specifically for use with forensic populations (i.e., the HoNOS-Secure). However, contrary to hypothesis seven, the remaining tools were non-forensic ROMs.

9.3.4 Study four

To conclude this thesis, empirical study four (Chapter 8) focused specifically on two tools that had been identified as showing particular promise for use as ROMs in forensic mental health settings during the earlier components of this thesis, namely the *Health of the Nation Outcome Scales* (HoNOS) and the forensic version of this tool known as the *HoNOS-Secure* (see figure 2). As noted, despite the HoNOS-secure having been developed in 2007 (Dickens et al., 2007), no direct comparisons of these different versions of this tool has appeared in the empirical literature.

Study four also investigated the extent to which the two versions of the HoNOS could be used interchangeably. This was considered particularly pertinent, as clients may move from forensic to civil settings over time. As such, if the two versions of the tool were found to produce analogous results, their overall pattern of clinical need could be monitored across time and setting using the same measure.

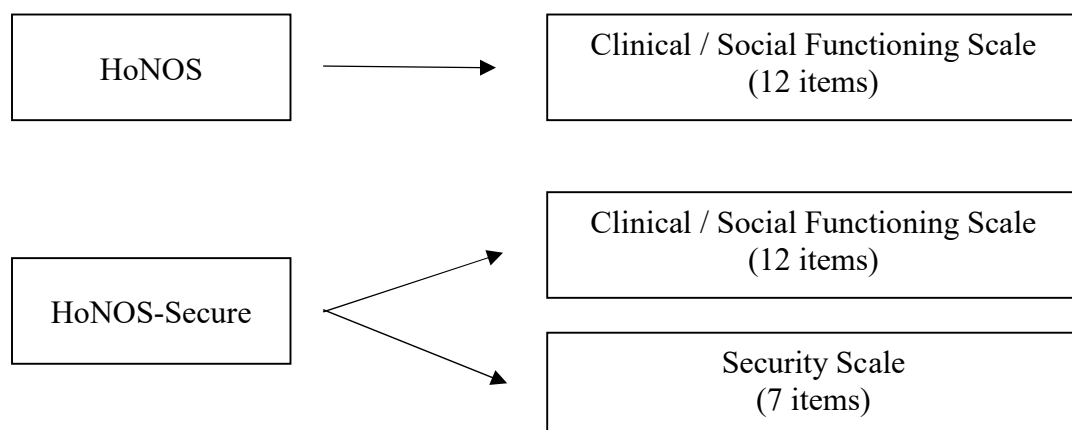


Figure 2: Composition of HoNOS and HoNOS-Secure tools

Logistic regression was subsequently employed to evaluate how well these tools were able to categorise patients on a range of measurable outcomes, including psychiatric acuity, risk related behaviour and freedom of movement.

The final component of study four compared the HoNOS scores generated by civil and forensic mental health patients, to evaluate whether differences might be observed to arise between these populations on the basis of mean scores at different points in admission.

The results of study four, which are presented in the fourth publication of this thesis (Chapter 8), indicated that the HoNOS/HoNOS-Secure correlated strongly at the total score level. However, these tools were also found to show variable correlations at the item level. That is, although each tool produced a similar overall result, the individual items within each tool varied considerably in how they were rated. Therefore, it was considered that individual items may be less comparable on the two versions of the tool, and that direct comparison at an item level may not be meaningful.

With regards to the comparison of the mean HoNOS scores obtained by forensic and civil mental health patients, these populations were not found to demonstrate a statistical difference in their degree of clinical acuity at the point of admission; however, they differed to a significant extent at both the review and discharge collection occasions. Specifically, when assessed during the review and discharge periods, the forensic cohort generated significantly lower mean scores on the HoNOS than the civil mental health population. Moreover, at the point of discharge, the mean HoNOS score of the forensic

sample was lower than is typically found in outpatient services and comparable with that found in the general population (e.g., Audin et al., 2001).

Regarding the logistic regression analysis, it was observed that the HoNOS and HoNOS-Secure ‘security scale’ demonstrated a statistically significant relationship to the independent variables (i.e., ward acuity, freedom of movement, and presence of risk behaviour). However, the HoNOS-secure ‘clinical and social functioning scale’ was not found to contribute significantly to the correct classification of patients for each of the independent variables. In each case, a model containing the HoNOS and HoNOS-Secure ‘security scale’ best predicted classification of patients at around 80% - 86%.

The above analysis was subsequently replicated using a second set of data to test whether this result was consistent over time and with a separate cohort of patients. Analogous results were obtained from the follow-up analysis, suggesting this finding was indeed stable and repeatable.

To further confirm these findings, a post-hoc investigation was conducted, in which the HoNOS-secure ‘security scale’ was combined with the HoNOS and an additional logistic regression was conducted using the HoNOS-Secure (‘clinical’ and ‘security’ scales) and HoNOS plus ‘security scales’. For all three independent variables, a combination of the HoNOS and HoNOS-Secure ‘security scale’ proved the most effective model, producing significant relationships in the expected direction. Whereas, the HoNOS-Secure ‘total scale’ (i.e., clinical and security scales combined) demonstrated a non-significant relationship.

Taken together, the results of study four suggested that in the present context the HoNOS-Secure ‘clinical scale’ did not perform better than the original HoNOS scale. Moreover, there did not appear to be any significant benefit from using the HoNOS-Secure in place of the HoNOS. Therefore, on the basis of the evidence generated by this analysis, it was concluded that combining the ‘security scale’ of the HoNOS-Secure with the original HoNOS would likely provide the greatest ability to correctly classify patients on both their clinical and forensic/risk needs (see figure 3).

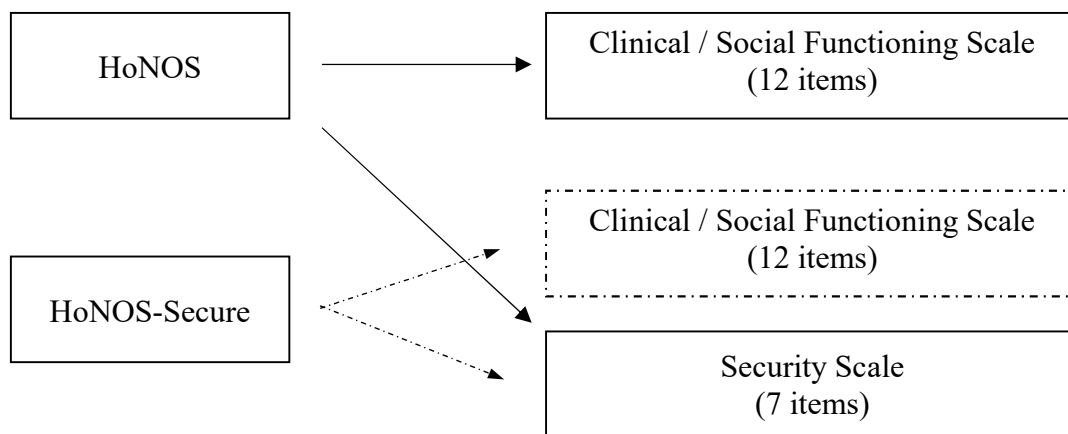


Figure 3: Most effective model for predicting classification of patients (HoNOS plus ‘security scale’).

9.4 Integrated Interpretation of Findings

Taking a high level view of the research as a whole, the present thesis describes a journey of exploration from identifying, collating and analysing existing tools that were designed for use in a forensic mental health setting; evaluating the existing NOCC tools for their reliability and validity in a forensic context; identifying any areas of difficulty with these existing tools; and ultimately evaluating a number of tools that were developed specifically for use with a forensic population against the currently mandated tools.

Whilst the above investigations provided the backbone for this thesis, these investigations were supported by a series of subcomponents designed to examine and understand the needs of forensic clients as a population, as well as to determine the differences in the needs demonstrated by forensic and non-forensic mental health populations. This second subcomponent of the thesis subsequently involved investigating the notion that clients in a forensic mental health facility may not be homogenous; and indeed, may differ in terms of clinical/social functioning and risk related needs.

In terms of integrating the findings of this thesis, it is perhaps pertinent to begin by considering the ancillary findings of this work first; upon which the discussion of outcome measures in forensic mental health must surely rest. That is, the comparison of HoNOS scores obtained by civil and forensic mental health consumers at different points of admission, as well as the identification of needs observed amongst forensic patients with respect to their clinical and risk related domains.

In many ways, the keystone finding of this thesis was the observation that a difference exists in mean HoNOS scores obtained by forensic and civil mental health patients. That is, when the mean HoNOS scores of these two populations were compared, both client groups were found to be statistically indistinguishable at the point of admission; however, differences in their level of clinical need emerged during the review and discharge periods. From a clinical perspective, this finding implies that at the point of admission the clinical needs of both populations are likely to overlap to a large degree, irrespective of the setting in which the client resides. Moreover, consistent with data generated during other components of this thesis, it was also noted that the NOCC tools did indeed appear to be capturing some aspects of the forensic population's presentation (i.e.,

their overall clinical needs, with the degree of need trending downwards as patients move towards discharge, see Figure 4). But, given the disparity between the two populations, these data also provided an initial indication that these tools may not be sufficiently capturing the raft of needs possessed by a forensic mental health population, particularly those needs which represent the reason that such individuals remain in secure care. That is, the additional forensic needs and risks that such clients present to themselves and others, which have led them to be detained longer than they may otherwise have been in a civil mental health setting. As such, it was considered that by neglecting to monitor these needs, the ROM tools had failed to account for changes in domains of high importance to this population, which have a direct impact on their capacity to be discharged out of secure care or to move towards greater autonomy and personal liberty within the hospital setting.

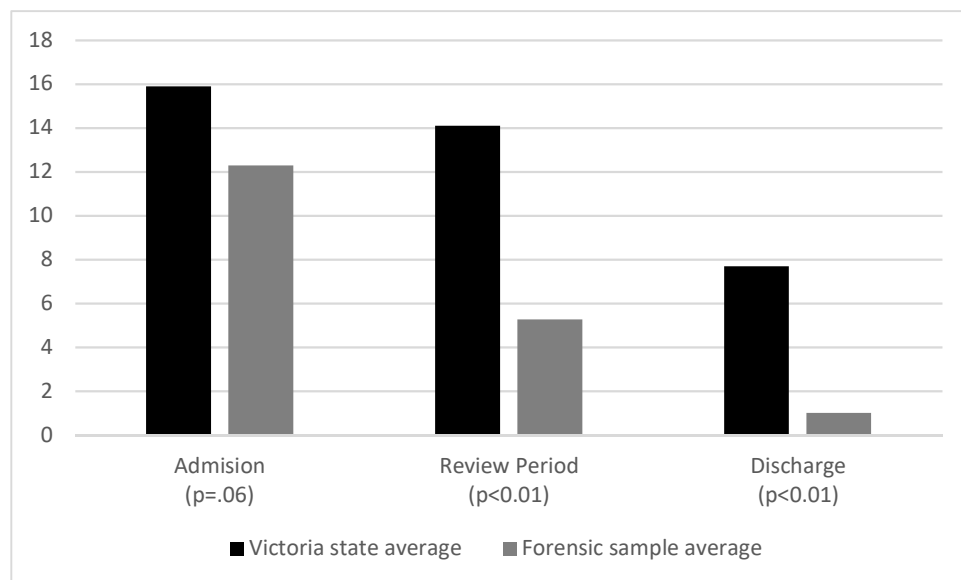


Figure 4: Mean HoNOS scores obtained by the Forensic and State-wide sample, showing t-statistics of statistical significance at each rating occasion (Graphical representation of data presented in Chapter 8 of this thesis).

The above analysis was subsequently strengthened by the findings of study three, which provided empirical data to support the notion that the needs of forensic patients were heterogeneous and that they differed with respect to the prominence of clinical and forensic/risk related domains. This has been described by previous authors (e.g., Cohen & Eastman, 1997; Keulen-de Vos & Schepers, 2016), but to date there has been a lack of empirical evidence to support this claim. The findings of study three indicated that, within the Victorian forensic mental health system, for some patients, issues pertaining to stabilisation of mental health appear to be the primary treatment concern, with their mental health difficulties contributing directly to offending behaviour. However, for others, mental health issues may not be the key factor underpinning their offending; with their criminogenic needs being more akin to non-mentally disordered offenders (Andrews & Bonta, 2006). In the latter case, mental health issues may be considered to be a comorbid issue, which impacts on the individual's daily and long-term functioning, but may not contribute directly to their risk of reoffending.

Given the diversity of needs amongst this population, the above findings suggested that selecting an outcome measurement tool or tools to monitor changes in clinical, social and forensic/risk needs may be a more complex task than for civil mental health populations. To this end, tools selected for this task would need to be able to capture a diverse range of needs, whilst being sensitive enough to detect change in these domains, and also be considered useful by both clinicians and patients (Happell, 2008; Kwan & Rickwood, 2015). It also became apparent throughout the series of studies generated by this thesis, that tools developed primarily as risk assessment instruments (i.e., to assess and monitor specific risk domains, such as general recidivism, violence and/or sexual offending) were not typically suited as ROMs. This was discussed at length in articles one

and three, particularly with respect to the limited utility of the HCR-20 and LSI-R:SV for monitoring changes in a broad range of risk/forensic needs. Consistent with hypothesis two, it was indeed found that a significant portion of the outcome measurement tools that had been developed for use in forensic settings had focused largely on the issue of monitoring risk, but neglected to provide a measure of broader clinical and psychosocial needs. However, risk assessment tools that have been designed to take a broad view of risk (e.g., the START), or tools that do not assess risk *per se*, but rather draw upon the information gathered by risk assessment tools to inform their ratings of need (e.g., the HoNOS-Secure ‘security scale’), did appear to be effective for this task. This mirrors the opinion expressed by the lead author of the HCR-20, version 2, who has cautioned against mandating the use of any *specific* risk assessment tool for use with all service users. Rather, a more flexible model is proposed, whereby clinicians should be permitted to select whichever measure of risk is most appropriately employed in each case (personal communication, C. Webster, June 2011) and utilise broader measures of need for the purposes of collating information pertaining to client progress and outcome.

The question that therefore arose from the above findings was, given the broad range of needs possessed by a forensic mental health population and the differences in mean scores obtained when compared to civil mental health consumers, how reliable are the tools that are currently mandated for use in Australian forensic settings.

Paper two sought to directly answer this question by conducting an investigation of the interrater reliability for the HoNOS and LSP-16. As noted above, the findings of this study demonstrated that these tools were indeed able to be completed by clinicians in a reliable and valid manner, with high levels of precision being obtained regarding the

ratings recorded on each item. This suggested that although these tools had not been designed with a forensic mental health population in mind, they were constructed in such a way as to be amenable to interpretation in a consistent and reliable manner in this setting. However, as was also discussed in paper two, due to several inherent differences between forensic and civil mental health settings, a number of limitations were identified with the use of the HoNOS and LSP-16 in a forensic mental health environment. Specifically, limitations arose with respect to items that were influenced directly by the environment, such as *Problems with living conditions*, *Problems with occupation and activities*, and *Problems with drinking or drug taking*. For example, Item 3 of the HoNOS focuses on a patient's use of substances over a two-week period. For most patients in secure settings, access to substances may be limited by environmental constraints; yet, the underlying problem may be demonstrated via cravings for substances, medication seeking behaviour, and/or other markers of dependence not captured by the HoNOS. This observation is consistent with findings of a large scale field trial of the HoNOS in Victoria (Trauer et al., 1999), in which items pertaining to living conditions and occupation were also identified as problematic (see also Pirkis et al., 2005a). Moreover, it was also noted that these measures did not provide information regarding treatment needs that are specific to a forensic environment, such as risk of harm to others, offending behaviour, and level of security required. Therefore, despite the finding that clinicians can utilise item criteria in a precise and reliable manner, questions were raised about the utility of the general adult version of the HoNOS in a forensic mental health setting.

At this stage of the thesis, it was considered that there was evidence to support the notion that although the existing tools could be used reliably in forensic mental health settings, they were not capturing the full range of needs pertinent to this client group. As

such, the thesis then turned to explore the extant literature with the aim of identifying any tools that had been already developed for this task. On the basis of this review, six tools were ultimately identified as showing potential for this task; both in terms of the need domains they monitored and the range of criteria they met for inclusion in the NOCC suite of measures, as required by the Australian government. Of these six tools, which were described in detail in chapter two of this thesis, two tools were selected for further evaluation. These were the *Health of the National Scale for Users of Secure Services (HoNOS-Secure)* and the *Camberwell Assessment of Need: Forensic Version (CANFOR)*. Whilst other measures showed potential for use in this setting (e.g., the *Short Term Assessment of Risk and Treatability [START]*, and the *DUNDRUM Quartet [DUNDRUM]*), it was not possible to evaluate all tools of interest within the scope of this thesis. This remains a limitation of the thesis and will be discussed further in the *Limitations* and *Future Direction* sections of this chapter (9.5 and 9.7, respectively). As such, the HoNOS-Secure and CANFOR were subsequently evaluated against the existing NOCC suite of measures for their ability to correctly identify and classify the needs of forensic mental health patients at different stages of their trajectory through the hospital setting.

The main evaluations of the HoNOS-Secure and CANFOR were presented in papers three and four of this thesis (Chapters 7 and 8) and drew on data obtained from the third and final study of this body of work.

Paper three sought to determine which tool was best able to correctly classify patients in terms of their acuity, freedom of movement (i.e., whether they were able to access the main campus freely, without the need for supervision by a staff member), as

well as the frequency of risk behaviour such as interpersonal aggression, self-harm and substance use. These tools were compared against the existing NOCC measures (i.e., HoNOS, LSP-16 and BASIS-32).

The findings of this study were particularly instructive in terms of understanding the relative strengths and weaknesses of both the NOCC and forensic ROM tools. In the first instance, it was observed that the HoNOS, HoNOS-Secure ‘clinical’ and ‘security’ scales, and the LSP-16 were able to differentiate effectively among patients at different levels of ward acuity. Moreover, the CANFOR and LSI-R:SV demonstrated the ability to differentiate between the needs possessed by patients on the acute unit and the subacute/rehabilitation wards; however, neither tool was sensitive enough to detect differences between patients residing on the subacute and rehabilitation wards. It was also noted that analysis of the BASIS-32 suggested that this tool failed to detect differences between clients at any of the levels of acuity. Moreover, when the HoNOS and HoNOS-Secure were investigated further, it was found that the HoNOS-Secure, as a whole, did not perform better than the HoNOS. Rather, a combination of the ‘security scale’ of the HoNOS-Secure and the original 12 items of the HoNOS produced the best model for differentiating amongst groups of forensic mental health patients.

The above finding regarding the HoNOS performing more effectively than the clinical scale of the HoNOS-Secure was somewhat unexpected; particularly as the HoNOS-Secure had been developed specifically for use with a forensic mental health population. Moreover, the ‘clinical and social functioning scale’ of the HoNOS-Secure was developed from the 12 items of the original HoNOS, with the wording of items being adapted with the aim of making them more readily interpreted in a secure environment.

However, findings generated by study four of this thesis suggest that these wording changes may in fact have impacted upon the way they are interpreted and rated, and in doing so, possibly reduced the utility of this scale.

A comparison of HoNOS/HoNOS-Secure wording changes for each of the 12 clinical items has been presented in Appendix L. At a surface level, the content of the 12 ‘clinical and social functioning’ items of the HoNOS-Secure appear to replicate the intent of the 12 HoNOS items. Indeed at least one previous study has reported that this tool demonstrates good face validity for these items (Dickens et al., 2007). Yet, close inspection of the 12 item pairs reveals that some of the wording changes may indeed have subtly changed the meaning of several items.

To illustrate this, the wording changes of item 2 (Non-accidental self-injury) have been presented in table 6 for specific consideration. In the first instance, item 2 of the HoNOS-Secure contains several anchor points that have been changed to include non-behavioural markers relating to self-harm (e.g., “persistent or worrying thoughts about self-harm”; anchor point two). As such, this requires the clinician to make an assessment of an aspect of the client’s experience that is not directly observable. Not only is it necessary that the clinician has specifically asked about thoughts of self-harm (which arguably should be part of routine clinical practice), but also that the client has been *willing to disclose* these thoughts to the clinician. In addition, the clinician then has to subjectively determine if these thoughts are ‘persistent or worrying’. Moreover, the fourth anchor point of item two includes reference to a person having ‘*intended to*’ cause serious self-harm. As such, this requires the clinician to make a subjective assessment of a person’s *intention*, rather than simply reporting observable and objective markers; or utilising other risk assessment

processes to determine the highest level of risk present during the rating period. It was noted in paper four (Chapter 8) that the correlation between the HoNOS/HoNOS-Secure for item two fell within the high range ($r = .549$); suggesting a reasonable degree of agreement in how these item pairs had been rated. However, the mean scores for the two versions of this item were found to differ to a statistically significant extent (HoNOS: $\mu = .11$, $SD = .54$; HoNOS-Secure: $\mu = .19$, $SD = .61$; $p = .001$). In this instance, the mean score on the HoNOS item was closer to a value of ‘one’ (i.e., “*minor problem requiring no action*”) whereas the mean score on the HoNOS-Secure item was closer to a value of ‘two’ (i.e., “*mild problem but definitely present*”; Sugarman & Walker, 2007; Wing et al., 1998). Post-hoc analysis via Cohen’s d indicated that the magnitude of the effect size between these mean total scores was small ($d = .129$), yet taken as a whole this may suggest there may be a tendency for scores obtained on item two of the HoNOS-Secure to be higher than the analogous item on the HoNOS. A similar pattern was observed across the majority of the 12 item pairs (see Chapter 8, table 4) and may suggest that the changes made to the wording of the HoNOS-Secure ‘clinical and social functioning’ scale may have impacted upon its overall reliability and comparability with the original HoNOS tool.

Table 6: Wording changes between the HoNOS and HoNOS-Secure of item two (Non-accidental self-injury).

HoNOS (Working Aged Adult Version)	HoNOS-Secure
2 Non-accidental self-injury	2 Non-accidental self-injury
Do not include accidental self-injury (due to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and the injury at Scale 5. Do not include illness or injury as a direct consequence of drug or alcohol use rated at Scale 3, (eg, cirrhosis of the liver or injury resulting from drunk driving are rated at Scale 5).	<u>Do not</u> include <u>accidental</u> self-injury (due to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and injury at Scale 5. <u>Do not</u> include illness/injury as a direct consequence of drug/alcohol use rated at Scale 3 (e.g., cirrhosis of liver or injury resulting from drunk driving are rated at Scale 5).
0 No problem of this kind during the period rated.	0. No problem of this kind during the period rated.
1 Fleeting thoughts about ending it all, but little risk during the period rated; no self-harm.	1. Fleeting thoughts about self-harm or suicide, but little risk; no self-harm.
2 Mild risk during period; includes non-hazardous self-harm eg, wrist– scratching.	2. Mild risk during period; includes non-hazardous self-harm (e.g., wrist scratching, not requiring physical treatment); persistent or worrying thoughts about self-harm.
3 Moderate to serious risk of deliberate self-harm during the period rated; includes preparatory acts eg, collecting tablets.	3. Moderate to serious risk of deliberate self-harm; includes preparatory acts (e.g., collecting tablets, secreting razor blade, making nooses, suicide notes).
4 Serious suicidal attempt or serious deliberate self-injury during the period rated.	4. Serious suicidal attempt and/or serious deliberate self-harm during period (i.e., person seriously harmed self, or intended to, or risk death by their actions).

Note. Highlighted text has been added by the author to signify changes between the tools

9.5 Limitations of the Research

The limitations of the empirical studies in this thesis have been discussed in detail in Chapters 6, 7 and 8, and so will only be considered briefly in this section. The most significant limitation, particularly for studies three and four, relates to the size and composition of the forensic mental health sample obtained. With respect to this, it is noted that the total capacity of patients within the study setting (Thomas Embling Hospital, TEH) at any given time was 116 beds. In addition, as has been demonstrated, the admission length of forensic mental health patients is typically far longer than in civil mental health settings (Ruffles, 2010; Turner & Salter, 2008). As such, with a relatively small proportion of the patient population being discharged/admitted, it was only possible to obtain a sample size of 202 complete data sets over the six-month period of data collection. Whilst the data obtained from this sample provides useful information about the use of ROMs tools in this setting, given the heterogeneity of this client group, it was not possible to identify how effectively these tools may work with specific subgroups of forensic patients.

To expand upon the above, it is recognised that the patient population of TEH is comprised of individuals who are detained under different involuntary treatment orders, broadly separated into two main categories: forensic patients, who have been found either unfit to stand trial, or of being not guilty of an offence on the grounds of mental impairment; and security patients, who are prisoners requiring assessment and/or treatment for mental health disorder. A small proportion of patients are also detained under civil involuntary hospitalisation orders. In the first instance, forensic patients typically remain within Thomas Embling Hospital for several years and, as long as they are deemed to pose a *serious endangerment to the public*, may remain even in the absence of active mental health symptoms (Ruffles, 2010; Turner & Salter, 2008). Whereas security patients,

typically remain in hospital for a period of weeks to months. Security patients generally leave the hospital either at the end of their sentence or by being discharged back to prison once their symptoms resolve, whichever occur first.

The presence of these two client groups within the study sample may have produced a confounding effect regarding the needs of patients at different data collection periods (i.e., admission, review and discharge). Specifically, it may be anticipated that the longer period of time spent by forensic patients in the hospital may provide a greater probability that their mental health symptoms would resolve and their underlying forensic and psychosocial needs may emerge. This may have influenced the findings of the study to the extent that the trend observed in the data may have been an artefact of these different subpopulations. However, whilst the greater numbers of security patients residing in the acute units of the hospital, and greater numbers of forensic patients in the subacute/rehabilitation units, may have influenced these results to some extent, it is noted that both sets of patients are monitored via the same ROM tools within this setting. Therefore, it is still pertinent that the broad range of needs of this diverse group be monitored by these tools.

A further consideration regarding the study sample is the finding that there was a significant gender imbalance amongst the patients in the hospital setting. As discussed in paper four, it was observed that the forensic cohort within Thomas Embling Hospital was skewed heavily towards males (85.7%). It is therefore uncertain if the findings of this study were able to generalise to female forensic mental health patients generally. However, it has been also noted that across correctional and forensic services internationally, men are uniformly observed in far greater numbers than women (Burman, Batchelor & Brown,

2001; de Vogel & Nicholls, 2016). As such, it may be considered that the present sample is representative of a typical patient-gender composition that exists within forensic mental health samples.

The final consideration with respect to the study sample relates to the data being obtained from a single forensic mental health facility. As such, without replicating this study in other forensic services, it cannot be stated for certain that these findings would generalise to other forensic settings. It is also acknowledged that the data upon which this thesis has been based was collected during two periods occurring between two and four years ago. As such, it could be possible that the findings of this study may not reflect any progress or change in clinical practice that has occurred since that time.

More broadly, given the scope of a doctoral thesis, the present studies were limited in the number of tools that were able to be evaluated. Whilst a number of other tools were identified by this thesis that showed potential for use as ROMs in forensic mental health (e.g., the *Short Term Assessment of Risk and Treatability [START]*; *DUNDRUM Quartet [DUNDRUM]*; *Illness Management and Recovery Scales [IMR]*; and the *Mental Health Recovery Measure [MHRM]*), it was not possible to evaluate all tools of interest at this time. Although the tools selected for evaluation were those which appeared most likely candidates based on review of the literature to be applicable to the setting in which the study was conducted, the remaining four tools cannot be dismissed outright and will require further research in the future (Gallagher & Teesson, 2000). Indeed, significant bodies of work have been published in recent years regarding the recovery focused tools identified within this study (e.g., IMR and MHRM), which already serve to address this gap (e.g., Burgess et al., 2010; Mental Health Information Strategy Standing Committee,

2015). Likewise, as was discussed at the conclusion of chapter one, since publication of the initial scoping study for this thesis, new empirical papers have appeared in the literature pertaining to the DUNDRUM quartet. The above discussion is expanded upon in the "Future Directions" section of this chapter (9.7).

9.6 Implications of this Research

Throughout this thesis, a number of potential implications have been raised in relation to this research. The following section will draw these together and discuss each in detail.

Overall, the findings of this thesis appear to confirm and provide support for the scoping work by Andrews, Peters, and Teesson (1994), which was conducted during the initial development of the NOCC suite; as well as the field testing (Stedman, et al., 1997) and subsequent review (Pirkis et al., 2005a) of these tools within Australia. That is, of the measures investigated, the HoNOS and LSP-16 do indeed appear to have ongoing utility as ROMs tools within forensic mental health settings. However, while these tools remain a useful component of forensic outcome measurement, they are in themselves necessary – but not sufficient – to capture the full range of needs pertinent to this client group.

Whilst undertaking this thesis, opportunity arose to present the findings generated by this work to the Forensic Mental Health Information Development Expert Advisory Panel (a transnational working party of AMHOCN, tasked with providing recommendations for the future of outcome measurement in Australian forensic mental health settings). Through contact with this working party, and members of AMHOCN more broadly, it became clear

that whilst the needs of forensic consumers were considered important and certainly required recognition within the NOCC framework, any recommendations pertaining to the ROM tools used for this population would necessarily have to sit within the requirements of the broader AMHOCN framework. That is, a recommendation to have separate ROMs for forensic and non-forensic populations would not likely be supported at this stage.

As discussed in chapters one and two, there has been considerable time, effort and resources invested in the development of outcome measurement systems for Australian mental health services (Burgess et al., 2015). Indeed, these efforts have led this country to be considered a world leader in this field (Eagar, Trauer & Mellsop, 2005; Pirkis et al., 2005; Slade, 2002b). Moreover, the original NOCC suite was conceived as a means of tracking patient progress over time, regardless of the setting in which they were receiving treatment (Pirkis et al., 2005b). Therefore, if separate measures were employed within forensic and civil mental health settings, this would impact negatively upon this goal. As such, in order to maintain the integrity of the NOCC framework, as well as the database that has been used to collect NOCC data over the past two decades, it would be necessary to retain the existing tools to enable comparisons of mental health status across forensic and civil psychiatric patient groups. In this sense, to capture the additional needs of this clients group would mean adding additional forensic focused tools that proved most effective for this task. However, alternatively, if it could be demonstrated that a forensic version of an existing tool could be used interchangeably in place of the existing NOCC measures (e.g., the HoNOS-Secure), this may have proved to be an adequate alternative option.

Given these considerations, an important aspect of the present thesis was to evaluate a range of tools identified for this purpose and propose the most parsimonious means of capturing the needs of forensic mental health patients in a manner that would retain the integrity of the NOCC framework. It was further considered that any recommendation to retain the existing NOCC measures and add additional forensic focused tools would have significant implications for the workload of clinicians responsible for conducting these assessments. Within forensic mental health, as with other areas of mental health service provision, resources are already stretched (RANZCP, 2010; Saxena, Thornicroft, Knapp & Whiteford, 2007). As such, whilst there may be a range of tools that would provide useful and pertinent information about the needs of forensic clients, this must be balanced by considering what can reasonably be expected of the clinical workforce.

As such, the findings of this thesis appear to present a useful way forward in terms of the future of ROM in Australian forensic mental health settings. Specifically, it was demonstrated that the most parsimonious means of adapting the NOCC suite for forensic patients would be to add the 7-item security scale of the HoNOS-Secure, whilst retaining the currently used clinician rated measures. While it was initially anticipated that the HoNOS-Secure (as a complete measure, with both the ‘clinical’ and ‘security’ scales) might serve this purpose most effectively, this was not supported by the findings of this thesis. Rather, the HoNOS-Secure ‘clinical’ scale did not appear to perform as well as the original HoNOS; but both versions of the tool appeared to benefit from the inclusion of the seven-item ‘security’ scale. This finding was replicated over time and with different cohorts of patients. Hence the recommendation of this thesis is to retain the HoNOS for the time being and add the additional seven items of the security scale as the primary clinician rated ROM tool. As discussed in section 9.4 of this chapter, it is likely that the wording

changes of the HoNOS-Secure ‘clinical scale’ impacted negatively on the overall utility of this facet of the tool.

Within Australia, outcome measurement in public mental health services are supported by electronic systems for the collection, collation and analysis of data (Burgess et al, 2015). If it were therefore possible to add an additional module to the database⁵, which would enable the collection of HoNOS-Secure security data, this would likely present the least disruptive means of achieving a more useful data set for this client group. Moreover, this would also mean that the existing dataset would maintain consistency and contiguity of data across mental health services (forensic or otherwise).

While the above findings regarding the clinician rated ROMs appears fairly straightforward, both in terms of the overall recommendations and potential for being implemented in the least disruptive manner, the findings with respect to consumer rated tools were more problematic. In the first instance, it was found that the currently used consumer measure (i.e., BASIS-32) was not effective at detecting change or differentiating

⁵ Collection of ROM data within Victoria occurs via the Wellbeing Module of the Department of Human Service’s Client Management Interface / Operational Data Store (RAPID - CMI-ODS) electronic database. The CMI/ODS is the information system for Victoria’s public mental health system and consists of two main components: the Client Management Interface (CMI) and the Operational Data Store (ODS). Additional information regarding the CMI/ODS can be obtained via the Victoria State Government website (Victoria State Government, 2016).

amongst forensic mental health clients across the different levels of ward acuity. Although the data generated by this body of work did not permit investigation of the reasons that underpin this to be explored directly, it is considered likely that the BASIS-32 may suffer from the same problems that bedevil many self-report questionnaires. Namely, a lack of insight experienced by acutely unwell clients and a range of personal barriers that impact upon completion rates when clients are invited to undertake self-directed tasks (e.g., Kazantzis, & Shinkfield, 2007; Shinkfield, 2006)⁶.

For those clients at the acute end of the hospital, a lack of insight into their condition may impact upon their capacity and/or motivation to complete such tools. Moreover, even for clients who are both aware of their mental health difficulties and indeed wish to express these to their clinical team via a reporting method such as the BASIS-32, previous research has demonstrated that when clients are asked to engage in tasks without the benefit of direct assistance of a clinician or other form of support person, rates of participation in such activities are significantly decreased (e.g., Paulhus & Vazire, 2009; Dattilio, Kazantzis, Shinkfield & Carr, 2011). Both of these factors potentially reduce the capacity of tools to obtain reliable data that would assist in differentiating between clients at different levels of acuity. As such, whilst this tool may provide some useful information by

⁶ Previously identified barriers to completing such tasks include: lack of comprehension by the client, task perceived as being too difficult, negative beliefs about the task held by the client, poor explanation of task by clinician, perception that the task will not be beneficial by the client, acuity of mental health problems, lack of collaboration between client and clinician, poor therapeutic alliance, practical obstacles such as time or resources to complete the task (Shinkfield, 2006).

way of a qualitative understanding of an individual's needs (i.e., by examining which items an individual patient has endorsed), from a quantitative perspective, it may be considered that its performance was less than adequate.

In contrast to the BASIS-32, the CANFOR was found to be able to differentiate between the needs of acute patients and those on subacute/rehabilitation wards. This is the first study published in which the CANFOR's sensitivity to detect change has been investigated (c.f. Segal et al., 2010). Indeed, this facet of the tool was questioned during the initial review of forensic measures that was presented in article one (Chapter 2) of this thesis. While the CANFOR was limited in its ability to differentiate amongst those patients at the subacute/rehabilitation end of the hospital, it still may prove more useful than the BASIS-32 on the whole.

Adding to the above dilemma regarding consumer rated ROMs, were the findings generated by several authors that have reported poor uptake and completion rates of consumer rated tools within the NOCC suite (Brophy & Moeller-Saxone, 2012; Pirkis & Callaly, 2010). Indeed, it has been shown by these studies that that majority of patients opt out of such assessment processes. However, of the two consumer rated outcome measures examined in this thesis (i.e., CANFOR and BASIS-32), when patients were asked to complete these tools, the CANFOR was completed at almost twice the rate of the BASIS-32. Whilst this finding was not overtly discussed in article three (Chapter 7), it was demonstrated in table three of the aforementioned paper that the CANFOR was completed by all 202 (100%) members of the sample, whereas the BASIS-32 was only completed by 134 (66%). Table 7 displays a summary of these data extracted from article three for the reader's reference.

Table 7: Number of patients on acute, subacute and rehabilitation wards who completed the BASIS-32 and CANFOR. (Data extracted from table three of article three).

Consumer rated outcome measure	Acute n (%)	Subacute n (%)	Rehabilitation n (%)	Total n (%)
BASIS-32	48 (54%)	61 (78%)	25 (71%)	134 (66%)
CANFOR	89 (100%)	78 (100%)	35 (100%)	202 (100%)

There may be a number of factors that could account for the above finding. In the first instance, it is noted that during the consent process for this study, participants were informed that they would be given the opportunity to complete an outcome measure; which would be compared against the existing tools currently used within the hospital setting. In doing so, there is a potential that they were primed to the idea that they would be participating in a novel assessment. However, it is noted that all participants were also offered the BASIS-32 at the same time as the CANFOR and no distinction was made as to which was the subject of the evaluation. It might also be considered that many of the patients within the sample, particularly those present at the review and discharge periods, would have been asked to have completed the BASIS-32 at several points in their admission. Indeed, for those participants who had resided within the hospital for several years, they would have been presented with this tool on a three-monthly basis throughout their admission, as per NOCC protocol. It might be hypothesised that this cohort of patients may have developed assessment fatigue in relation to this tool, or may otherwise have developed/held beliefs that did not support their participation in this task (Shinkfield, 2006). However, these questions were not assessed directly in the present study and therefore cannot be extrapolated here. It is also noted that approximately half of the patients at the admission collection period also opted not to complete the BASIS-32, whereas *all* chose to complete the CANFOR. As such, this would suggest that the above hypotheses about the BASIS-32 not being completed solely by those patients who were

already familiar with the tool, or indeed those who were acutely unwell, may not be supported.

An alternative explanation regarding the higher uptake of the CANFOR may have to do with the manner in which these two tools were presented to patients. In the case of the BASIS-32, patients in the study setting were typically presented with the questionnaire and simply asked to complete it. It is also noted that at the top of the BASIS-32 form used within the study setting, there is a note within the instructions that advises the patient that they may choose not to complete the tool if they do not wish to. Despite clinicians being encouraged to sit with their patients and assist them to complete the form, *ad hoc* observation of clinical practice within the study setting and anecdotal reports obtained from staff indicated that this rarely occurred. In contrast, by the very nature of the CANFOR, in order to complete the tool a clinician is required to sit and discuss each area with a patient. On the basis of their discussion, the clinician then rates each item according to the clients' responses. In this manner, the CANFOR is considered a tool that serves to capture consumer opinion, but the onus for completing the form is placed upon the clinician. Perhaps this facet of the CANFOR, as a means of generating discussion about a range of needs areas that are not typically canvassed in everyday interactions with patients, is the key to the high level of uptake by consumers. Patients are actively engaged in a conversation about areas of importance to them. This enables patients to experience their needs as being heard and understood by their treating team; in a manner that it not achieved via completing a self-report outcome measure form. This aspect of the CANFOR/BASIS-32 was not explored directly in the present thesis and will be discussed in the next section of this chapter as a potential area of future research.

Returning to the question of recommendations for future developments of the NOCC suite, on the grounds that the CANFOR was completed more readily than the BASIS-32, and indeed demonstrated great capacity to differentiate clients at different levels of acuity, it is recommended that the CANFOR be considered for further investigation with a view to being included as a consumer measure in the NOCC suite of forensic outcome measures. This should, however, be balanced against the factors identified above regarding the integrity of the NOCC framework and the impact that a tool which takes longer to complete the BASIS-32 may have upon clinicians' abilities to comply with the reporting requirements of the NOCC protocol.

9.7 Future Directions

As with most research, over the course of completing this thesis a number of additional areas of investigation became apparent. These include: replication of the study across multiple (interstate and possibly international) forensic mental health services; expand the number of tools field tested to include all six of instruments identified during the review of the literature; repeat and update the review of the literature regarding forensic ROM tools available, prospectively collect data from civil and forensic mental health services to compare and contrast the needs of these client groups; further investigation of the reason for higher completion rates of the CANFOR versus the BASIS-32; and field testing of the audit tool developed by this study for its utility in clinical practice.

In the first instance, as noted in the limitations section, data for this study were collected from one forensic mental health setting. As such, it is possible that the findings of this study may not generalise to other forensic facilities. It would therefore be pertinent to

extend this study across multiple forensic mental health services to provide a robust analysis of the needs of forensic patients and evaluate the utility of the ROM tools across multiple settings. In doing so, it would be possible to test the findings of this thesis and either confirm or refute their accuracy. It is also considered that a more robust investigation of the forensic focused tools would likely be necessary in order for them to be included in the NOCC suite of measures, in a similar manner to the extensive field testing that was undertaken during the development of the existing NOCC suite (e.g., Stedman, et al., 1997; Trauer, et al., 1999).

It is further considered that, with additional resources, it would be pertinent to expand the number of tools field tested to include those other tools identified from the review and analysis of forensic ROMs literature (i.e., the *Short Term Assessment of Risk and Treatability*; *DUNDRUM Quartet*; *Illness Management and Recovery Scales*; and the *Mental Health Recovery Measure*). It is noted that during the six years over which this study has been conducted, increased urgency and focus has been given to integrating a recovery framework throughout mental health services in Australia. Given this, it is particularly pertinent that the recovery focused tools be investigated and field tested in a forensic context (see also Burgess, et al., 2010). Moreover, at the point of embarking upon this thesis, the DUNDRUM quartet was in its infancy and had only recently been published. Since then, a number of studies have emerged regarding this tool, which have indicated that it may have good utility in forensic mental health settings (Abidin, Davoren, Naughton, Gibbons, Nulty & Kennedy, 2013; Davoren, Abidin, Naughton, Gibbons, Nulty, Wright & Kennedy, 2013; Davoren, O'Dwyer, Abidin, Naughton, Flynn, O'Neill, McInerney & Kennedy, 2011; Gibbons & Doyle, 2012; Flynn, O'Neill & Kennedy, 2011; Kennedy, O'Neill, Flynn & Gill, 2010; O'Dwyer, Davoren et al., 2011). Indeed, it is noted

that such an evaluation is currently being conducted in New South Wales (Australia) by the Justice Health and Forensic Mental Health Network (NSW Agency for Clinical Innovation, 2015). Likewise, the START has also generated additional empirical support since the review was conducted within this thesis of forensic ROMs (the reader is referred to <http://www.bcmhsus.ca/start> for a list of publications in relation to this tool) and a recent publication by Dickens and O'Shea (2017) reports on the use of the HoNOS-Secure as a reliable and clinically significant measure of outcome for forensic mental health inpatients. As such, each of these tools appear to warrant further investigation.

Consistent with the above, it is also noted that in the intervening years since the publication of article one of this thesis (Chapter 2), additional tools have already emerged in the literature seeking to address the gap in forensic ROM measures. Examples of such tools are: the *Security Needs Assessment Profile* (SNAP; Davies, Collins & Ashwell, 2012), and the *Instrument for Forensic Treatment Evaluation* (IFTE; Schuringa, Spreen & Bogaerts, 2014; Schuringa, Heininga, Spreen & Bogaerts, 2016). Moreover, of those tools that were identified during study one, but were rejected from the hierarchical analysis on the basis of limited empirical support in the literature, it is recommended that these tools be monitored for future developments. For example, since the publication of the review, the Atascadero Skills Profile has appeared in a new publication (Hakvoort, Bogaerts & Marinus, 2012) and may now meet criteria for further investigation.

With respect to the findings of article four, specifically regarding the differences in clinical acuity observed between the forensic and civil samples when assessed via the HoNOS, it is considered that there may be utility in replicating this study in the future. In particular, utilising a prospective research design, whereby HoNOS assessments are

completed concurrently with both a forensic and civil sample. Moreover, it would also behoove future researchers to collect such data from a variety of sources, including multiple forensic and civil mental health facilities.

It was noted in the limitations section that a difference in completion rates was observed between the CANFOR and BASIS-32 consumer rated measures. While several hypotheses were proffered as possible explanations for this observation, it was not possible to explore and formally test these hypotheses using the data collected for the present thesis. However, given the difficulties reported within the literature regarding patient completion of such self-report tools (Paulhus & Vazire, 2009), this may prove a fruitful source of enquiry for future investigation.

Finally, as discussed in chapter six, the audit protocol developed to gather data for investigating the reliability and precision of HoNOS/LSP-16 assessments appeared to show some promise as a tool for providing feedback to clinicians regarding their completion of these tools. It is noted that the utility of such tools is only as strong as the manner in which they are used. As such, if clinicians do not complete ROMs accurately and validly, the data obtained will be spurious at best. Within this context, it is proposed that the audit tool be further evaluated in clinical practice to investigate if there is any potential benefit in employing this as a teaching aide for clinicians.

9.8 Conclusion

The future directions and limitations notwithstanding, the present thesis achieved the goals specified at the outset of this body of work. It was found that while the existing

NOCC tools demonstrate some capacity to monitor the clinical needs of forensic mental health patients, and indeed could be used with an adequate degree of reliability and precision in such settings, there were significant needs not captured by these tools that are considered highly pertinent to this population. As such, comparing forensic and general psychiatric patients using the existing NOCC suite may present an incomplete picture of acuity and treatment needs of these groups. While forensic patients, as a cohort, demonstrate greater levels of psychiatric stability by the time they are discharged, the existing measures are considered to not adequately capture the broader range of forensic needs that are pertinent to this population

To be useful, ROMs must be valid and reliable, sensitive to change, comparable across relevant client groups and service types, and be meaningful to both clients and clinicians (Happell, 2008; Kwan & Rickwood, 2015). As such, based on the findings of this body of work, the recommendation of this thesis is that, for the time being, the HoNOS be retained for use with forensic mental health populations. Yet, in order for this measure to be effective with this population it should be supplemented with the ‘security scale’ of the HoNOS-Secure. In addition, it is recommended that consideration be given to substituting the existing consumer measure (BASIS-32) for the Camberwell Assessment of Need: Forensic Version (CANFOR).

It is, however, believed that the above recommendations should only be viewed as a way forward in the *short term*. Ultimately it is recommended that effort now needs to be expended on evaluating the remaining measures identified within this body of work (i.e., the *Short Term Assessment of Risk and Treatability*; *DUNDRUM Quartet*; *Illness Management and Recovery Scales*; and the *Mental Health Recovery Measure*), with a view

to developing a NOCC suite tailored for use specifically with forensic patients. Given the demonstrable differences between forensic and civil mental health populations, failing to capture those needs pertinent to a forensic population undermines the utility of outcome measurement with this client group as a whole. It is therefore considered imperative for a suite of tools to be developed that provides adequate overlap with the existing NOCC suite, particularly with respect to clinical needs, but also adequately evaluates the forensic/risk needs pertinent to this group.

Over the course of completing this thesis, it had become apparent that the fields of forensic mental health and routine health outcome measurement continue to evolve. In this changing landscape, new tools continue to be created and subsequently evaluated as a means of better capturing the needs of mental health patients across all populations. As such, the final recommended of this thesis is that the ROM tools used within Australia should be reviewed at least every 10 years. Any new technology that emerges should be evaluated for its potential use within the National Outcomes Case Collection suite of measures. Moreover, if found to be of greater utility to all stakeholders (i.e., patients, clinicians and government), effort should be made to incorporate such tools into the NOCC framework.

PART E: THESIS REFERENCES

- Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H.G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry*, 13, 197 – 215.
- Andreasson, H., Nyman, M. Krona, H., Anckarsäter, H., Nilsson, T., & Hofvander, B. (2014). Predictors of length of stay in forensic psychiatry: The influence of perceived risk of violence. *The International Journal of Law and Psychiatry*, 37, 635 – 642.
- Andresen, R., Caputi, P., & Oades, L. G. (2010). Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Research*, 177, 309 - 317.
- Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (1998). *The Level of Service Inventory–Revised: Screening Version*. Toronto: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati, OH: Anderson Publishing Co.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2009). The level of service (LS) Assessment of adults and older adults. In R. Otto & K. S. Douglas (Eds.), *Handbook of Violence Risk Assessment*. New York, NY: Taylor & Francis.
- Andrews, D. A., Bonta, J., & Wormith, S. J. (2004). *The Level of Service/Case Management Inventory*. Toronto: Multi-Health Systems.

- Andrews, G., Peters, L., & Teesson, M. (1994). *Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee*. Sydney: Clinical Research Unit for Anxiety Disorders.
- Audin, K., Margison, F. R., Clark, J. M., & Barkham, M. (2001). Value of HoNOS in assessing patient change in NHS psychotherapy and psychological treatment services. *The British Journal of Psychiatry*, 178, 561 - 566.
- Australian Bureau of Statistics. (2016). *3101.0 - Australian Demographic Statistics (Mar 2016)*. Retrieved from <http://www.abs.gov.au/ausstats/abs@.nsf/mf/3101.0>
- Australian Health Ministers. (1993). *National Mental Health Plan*. Canberra: Commonwealth Department of Health and Family Services.
- Australian Health Ministers. (1998). *Second National Mental Health Plan*. Canberra: Commonwealth Department of Health and Family Services.
- Australian Health Ministers. (2003). *National Mental Health Plan 2003–2008*. Canberra: Commonwealth Department of Health and Family Services.
- Australian Health Ministers. (2008). *National Mental Health Policy 2008*. Canberra: Commonwealth Department of Health and Family Services.
- Australian Institute of Health and Welfare. (2005). *Report on the Evaluation of the National Minimum Data Set for Admitted Patient Mental Health Care*. Canberra.
- Australian Mental Health Outcomes & Classification Network. (2005). *Adult National Outcomes & Casemix Collection Standard Reports. 1st Edition, version 1.1*. Retrieved from <http://www.amhocn.org/publications/adult-nocc-standard-reports>
- Australian Mental Health Outcomes & Classification Network. (2008). *National Mental Health Benchmarking Project Forensic: Review of Key Performance Indicators*. Retrieved from http://amhocn.org/static/files/assets/7fe17da1/2008_FORENSIC_KPI_Review_DRAFT_FINAL.pdf

- Australian Mental Health Outcomes & Classification Network. (2009). *AMHOCN NOCC Reports Portal. 2.1.2 Core Clinical Summary Score Profiles*. Retrieved from http://reports.mhnocc.org/cgi-bin/menu.cgi?plugin=2_1_2
- Australian Mental Health Outcomes & Classification Network. (2013). *NOCC Volume & Percentage Clinical Ratings: Australia*. Retrieved from <http://amhocn.org/analysis-reporting/nocc-volume-percentage-clinical-ratings-australia/>
- Australian Mental Health Outcomes & Classification Network. (2016). *Australian Mental Health Outcomes and Classification Network*. Retrieved from <http://amhocn.org/>.
- Australian Mental Health Outcomes & Classification Network. (2016a). *Mental Health National Outcomes and Casemix Collection: Technical specification of State and Territory reporting requirements, Version 1.90*. Canberra: Department of Health.
- Australian Mental Health Outcomes & Classification Network. (2016b). *NOCC Data Volume & Completeness: National & State/Territory. Adults by Setting by Collection Occasion 2011-12 through 2014-15*. Retrieved from <http://www.amhocn.org/publications/nocc-data-volume-and-completeness-national-stateterritory-adults>.
- Australian Mental Health Outcomes & Classification Network. (2011). *AMHOCN - Useful Resources*. Retrieved from <http://amhocn.org/home/useful-resources>.
- Beveridge, J., Papps, R., Bower, A., & Smith, M. (2012). *Mental Health Outcomes Information Collection Protocol – HONOS Family (Version 2.1)*. Retrieved from <http://www.tepou.co.nz/download/asset/502>.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk – 20: Professional guidelines for assessing risk of sexual violence*. Vancouver, B.C.: The Mental Health, Law, and Policy Institute.

- Bonta, J., & Andrews, D. A. (2012). Viewing offender assessment and rehabilitation through the lens of the risk-need-responsivity model. In F. McNeill, P. Raynor, & C. Trotter (Eds.), *Offender Supervision: New Directions in Theory, Research and Practice* (pp. 19 - 40). New York: Routledge.
- Borum, R., Bartel, P., & Forth, A. (2003). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)* (1.1 ed.). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute.
- Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *The International Journal of Forensic Mental Health*, 9, 271 – 281.
- Brophy, L., & Moeller-Saxone, K. (2012). *Using outcome measures in Mind Australia: Discussion paper*. Retrieved from <https://www.mindaustralia.org.au/assets/docs/research/Outcome%20measures%20in%20Mind%20Australia%20November%202012.pdf>.
- Brown, P., & Pirkis, J. (2009). Mental health quality and outcome measurement and improvement in Australia. *Current Opinion in Psychiatry*, 22, 610 – 618.
- Buckingham, W., Burgess, P., Solomon, S., Pirkis, J., & Eagar, K. (1998). *Developing a Casemix Classification for Mental Health Services (Vol. 2: Resource Materials)*. Canberra: Commonwealth Department of Health and Family Services.
- Bullock, W. A. (2005). The Mental Health Recovery Measure. In T. Campbell-Orde, Chamberlin, J. Carpenter, J., & Leff, H. S. (Ed.), *Measuring the Promise of Recovery: A Compendium of Recovery Measures. Volume II*. Cambridge, MA: The Evaluation Center.
- Bullock, W. A., O'Rourke, M., Farrer, E., Breedlove, A., M.K., S., & Claggett, A. (2005). *Evaluation of the Illness Management and Recovery (IMR) Program*. Paper

- presented at the 113th annual meeting of the American Psychological Association Meeting, Washington, D.C.
- Bullock, W. A., Sage, J., Hupp, D., O'Rourke, M., & Smith, M. K. (2009). From illness management and recovery (IMR) to wellness management and recovery (WMR): Implementation and evaluation of Ohio's WMR program. *New Research in Mental Health, 18*, 312 – 321.
- Bullock, W. A., Wuttke, G. H., Klein, M., Bechtoldt, H., & Martin, J. (2002). Promoting mental health recovery in an urban county jail. *New Research in Mental Health, 15*, 305 – 314.
- Burgess, P., & Coombs, T. (2011). *National Outcomes & Casemix Collection: Volume and Percentage Valid Clinical Ratings*. Brisbane, Queensland: Australian Mental Health Outcomes and Classification Network.
- Burgess, P., Coombs, T., Dickson, R., & Pirkis, J. (2012). Achievements in mental health outcome measurement in Australia: Reflections on progress made by the Australian Mental Health Outcomes and Classification Network (AMHOCN). *International Journal of Mental Health Systems, 6*, 1 – 12.
- Burgess, P., Pirkis, J., Buckingham, W., Eagar, K., & Solomon, S. (1999). Developing a casemix classification for specialist mental health services. *Casemix, 1*, 4 - 20.
- Burgess, P., Pirkis, J., & Coombs, T. (2015). Routine outcome measurement in Australia. *International Review of Psychiatry, 27*, 264 – 275.
- Burgess, P., Pirkis, J., & Coombs, T. (2006). Do adults in contact with Australia's public sector mental health services get better? *Australia and New Zealand Health Policy, 3*, 1 – 7.
- Burgess, P., Pirkis, J., Coombs, T., & Rosen, A. (2010). *Review of recovery measures (version 1.01)*. Retrieved from <http://amhocn.org/static/files/assets/80e8befc/>

Review_of_Recovery_Measures.pdf

- Burgess, P., Trauer, T., Coombs, T., McKay, R., & Pirkis, J. (2009). *What does clinical significance' mean in the context of the Health of the Nation Outcome Scales? Australasian Psychiatry, 17*, 141 – 148.
- Burlingame, G. M., Lambert, M. J., Reisinger, C. W., Neff, W. M., & Mosier, J. (1995). Pragmatics of tracking mental health outcomes in a managed care setting. *Journal of Mental Health Administration, 22*, 226 – 236.
- Burman, M. J., Batchelor, S. A., & Brown, J. A. (2001). Researching girls and violence: Facing the dilemmas of fieldwork. *British Journal of Criminology, 41*, 443 – 459.
- Burns, A., Beevor, A., Lelliott, P., Wing, J., Blakey, A., Orrell, M., Mulinga, J., & Hadden, S. (1999). Health of the Nation Outcome Scales for Elderly People (HoNOS 65+). Glossary for HoNOS65+ score sheet. *The British Journal of Psychiatry, 174*, 435 – 438.
- Butler, T., & Allnutt, S. (2003). *Mental illness among New South Wales prisoners*. New South Wales: Corrections Health Service.
- Campbell-Orde, T., Chamberlin, J., Carpenter, J., & Leff, H. S. (2005). *Measuring the promise: A compendium of recovery measures*. Cambridge, MA: Human Services Research Institute.
- Carroll, A. (2008). Risk assessment and management in practice: the Forensicare Risk Assessment and Management Exercise. *Australasian Psychiatry, 16*, 412 - 417.
- Chambers, J. C., Yiend, J., Barrett, B., Burns, T., Doll, H., & Fazel, S. (2009). Outcome measures used in forensic mental health research: a structured review. *Criminal Behaviour and Mental Health, 19*, 9 – 27.

- Cohen, A., & Eastman, N. (1997). Needs assessment for mentally disordered offenders and others requiring similar services: theoretical issues and a methodological framework. *British Journal of Psychiatry*, 171, 412 – 416.
- Cohen, A., & Eastman, N. (2000). Need assessment for mentally disordered offenders: measurement of 'ability to benefit' and outcome. *British Journal of Psychiatry*, 117, 493 – 498.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37 – 46.
- Coid, J., Kahtan, N., Gault, S., Cook, A., & Jarman, B. (2001). Medium secure forensic psychiatry services: Comparison of seven English health regions. *British Journal of Psychiatry*, 178, 55 – 61.
- Commonwealth of Australia. (2006). *Select Committee on Mental Health: A national approach to mental health - from crisis to community*. Retrieved from www.aph.gov.au/~media/wopapub/senate/Committee/mentalhealth_ctte/report/report_pdf.ashx
- Coombs, T., & Meehan, T. (2003). Mental health outcomes in Australia: Issues for mental health nurses. *International Journal of Mental Health Nursing*, 12, 163 – 164.
- Corrado, R. R., Cohen, I. M., Hart, S. D., & Roesch, R. (2000). Diagnosing mental disorders in offenders: Conceptual and methodological issues. *Criminal Behaviour and Mental Health*, 10, 29 – 39.
- Crocker, A. G., Braithwaite, E., Laferriere, D., Gagnon, D., C., V., & Jenkins, T. (2011). START Changing Practice: Implementing a risk assessment and management tool in a civil psychiatric setting. *International Journal of Forensic Mental Health*, 10, 13 – 28.

- Daffern, M., Ogloff, J. R. P., Ferguson, M., & Thomson, L. (2005). Assessing risk for aggression in a forensic psychiatric hospital using the Level of Service Inventory-Revised: Screening version. *International Journal of Forensic Mental Health, 4*, 201 – 206.
- Dambro, M. R., & Weiss, B. D. (1988). Assessing the quality of data entry in a computerized medical records system. *Journal of Medical Systems, 12*, 181 – 187.
- Dattilio, F., Kazantzis, N., Shinkfield, G., & Carr, A. (2011). A survey of homework use, experience of barriers to homework, and attitudes about the barriers to homework among couples and family therapists. *Journal of marital and family therapy, 37*, 121 – 136.
- Davies, S., Collins, M., & Ashwell, C. (2012). Validation of the Security Needs Assessment Profile (SNAP) by a national survey of secure units in England. *The Psychiatrist, 36*, 366 – 371.
- Davoren, M., Abidin, Z., Naughton, L., Gibbons, O., Nulty, A., Wright, B., & Kennedy, H. (2013). Prospective study of factors influencing conditional discharge from a forensic hospital: the DUNDRUM-3 programme completion and DUNDRUM-4 recovery structured professional judgement instruments and risk. *BMC Psychiatry, 13*, 185 - 201.
- Davoren, M., Byrne, O., O'Connell, P., O'Neill, H., O'Reilly, K., & Kennedy, H. G. (2015). Factors affecting length of stay in forensic hospital setting: need for therapeutic security and course of admission. *BMC Psychiatry, 15*, 301 – 316.
- Davoren, M., O'Dwyer, S., Abidin, Z., Flynn, G., & Kennedy, H. G. (2012). Cross validation of a structured professional judgement tool for security needs in mental health: the DUNDRUM-1. *BMC Psychiatry, 11*, 1 – 11.

- de Vogel, V., & Nicholls, T. (2016). Gender matters: An introduction to the special issues on women and girls. *International Journal of Forensic Mental Health, 15*, 1 – 25.
- Deady, M. (2009). *A review of screening, assessment and outcome measures for drug and alcohol settings*. Retrieved from http://www.nada.org.au/downloads/Info_Management_Project/Review_of_Measures_09.pdf.
- Department of Health and Ageing. (2003a). *Mental Health National Outcomes and Casemix Collection: Technical specification of state and territory reporting requirements for the outcomes and casemix components of 'Agreed Data', Version 1.50*. Canberra: Commonwealth Department of Health and Ageing.
- Department of Health and Ageing. (2003b). *Mental Health National Outcomes and Casemix Collection: Overview of clinician rated and consumer self-report measures, Version 1.50*. Canberra: Department of Health and Ageing.
- Department of Human Services. (2008). *Because mental health matters: A new focus for mental health and wellbeing in Victoria, Consultation paper*. Melbourne: Victorian Government Department of Human Services.
- Department of Human Services. (2009). *Opportunities in practice: outcome measurement in mental health*. Melbourne: Mental Health and Drugs Division, Department of Human Services.
- Derogatis, L. R. (1993). *Brief Symptom Inventory (BSI): Administration, scoring and procedures manual* (3rd ed.). Minneapolis: NCS Pearson, Inc.
- Desmarais, S. L., Nicholls, T. L., & Brink, J. (2008). *Psychometric properties of the START in practice: Increasing the ecological validity of risk assessment research*. Paper presented at the International Association of Forensic Mental Health Services, Vienna, Austria.

- Dickens, G., & O'Shea, E. (2017). Reliable and Clinically Significant Change in Outcomes for Forensic Mental Health Inpatients: Use of the HoNOS-Secure. *International Journal of Forensic Mental Health, 16*, 161 – 171.
- Dickens, G., Sugarman, P., Picchioni, M., & Long, C. G. (2010). HoNOS-Secure: tracking risk and recovery for men in secure care. *British Journal of Forensic Practice, 12*, 36 – 46.
- Dickens, G., Sugarman, P., & Walker, L. (2007). HoNOS-secure: a reliable outcome measure for users of secure and forensic mental health services. *Journal of Forensic Psychiatry & Psychology, 18*, 507 – 514.
- Douglas, K. S., Blanchard, A. J. E., Guy, L. S., Reeves, K. A., & Weir, J. (2010). *HCR-20 violence risk assessment scheme: overview and annotated bibliography*. Retrieved from <http://kdouglas.files.wordpress.com/2007/10/hcr-20-annotated-biblio-sept-2010.pdf>.
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk of violence – User guide*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S., & Reeves, K. (2009). The HCR-20 violence risk assessment scheme: Overview and review of the research. In R. Otto & K. S. Douglas (Eds.), *Handbook of violence risk assessment*. Oxford, UK: Routledge/Taylor & Francis.
- Doyle, M., Lewis, G., & Brisbane, M. (2008). Implementing the Short-Term Assessment of Risk and Treatability in a forensic mental health service. *Psychiatric Bulletin, 32*, 406 – 408.
- Dwyer, S. O., Davoren, M., Abidin, Z., Doyle, E., McDonnell, K., & Kennedy, H. G. (2011). The DUNDRUM Quartet: validation of structured professional judgement instruments DUNDRUM-3 assessment of programme completion and

- DUNDRUM-4 assessment of recovery in forensic mental health services. *BMC Research Notes*, 4, 1 – 12.
- Eagar, K., Burgess, P., & Buckingham, B. (2003). *Towards national benchmarks for Australian mental health services. ISC discussion paper No 4*. Canberra: Commonwealth Department of Health and Ageing.
- Eagar, K., Gaines, P., Burgess, P., Green, J., Bower, A., Buckingham, B., & Mellsop, G. (2004). Developing a New Zealand casemix classification for mental health services. *World Psychiatry*, 3, 172 – 177.
- Eagar, K., Trauer, T., & Mellsop, G. (2005). Performance of routine outcome measures in adult mental health care. *Australian and New Zealand Journal of Psychiatry*, 39, 713 – 718.
- Eisen, S. V., Dill, D. L., & Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry*, 45, 242 – 247.
- Ellwood, P. (1988). Shattuck Lecture - Outcomes management. A technology of patient experience. *The New England Journal of Medicine*, 318, 1549 – 1556.
- Emmanuel, M. K., & Campbell, M. H. (2009). Comparing staff and client perceptions of needs in a British single-sex medium secure unit. *Psychological Reports*, 104, 1033 – 1034.
- Ferguson, C. J., & Negy, C. (2006). Development and preliminary validation of a defendant and offender screening tool for psychopathology in inmate populations. *Criminal Justice and Behavior*, 33, 325 – 346.
- Ferguson, M. A., Ogloff, J. R. P., & Thomson, L. (2009). Predicting recidivism by mentally disordered offenders using the LSI-R:SV. *Criminal Justice and Behavior*, 36, 5 – 20.

- 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 Technology Assessment, 14*.
- Flynn, G., O'Neill, C., & Kennedy, H. G. (2011). DUNDRUM-2: Validation of a structured professional judgment Instrument assessing priority for admission from the waiting list for a Forensic Mental Health Hospital. *BMC Research Notes, 4*.
- Flynn, G., O'Neill, C., McInerney, C., & Kennedy, H. G. (2011). The DUNDRUM-1 structured professional judgement for triage to appropriate levels of therapeutic security: retrospective-cohort validation study. *BMC Psychiatry, 11*. Retrieved from <http://www.biomedcentral.com/1471-244X/11/43>.
- Gallagher, J., & Teesson, M. (2000). Measuring disability, need and outcome in Australian community mental health services. *Australian and New Zealand Journal of Psychiatry, 34*, 850 – 855.
- Garland, A., Kruse, M., & Aarons, G. (2003). Clinicians and outcome measurement: What's the use? *Journal of Behavioural & Health Services Research, 30*, 393 – 405.
- Garner, B. R., Knight, K., Flynn, P. M., Morey, J. T., & Simpson, D. D. (2007). Measuring offender attributes and engagement in treatment using the Client Evaluation of Self and Treatment. *Criminal Justice and Behavior, 34*, 1113 – 1130.
- Gilbody, S. M., House, A. O., & Sheldon, T. A. (2003). *Outcomes measurement in psychiatry: A critical review of outcomes measurement in psychiatric research and practice*. York: University of York.
- Gowers, S., Harrington, R., Whitton, A., Beever, A., Lelliott, P., Jezzard, R., & Wing, J. (1999). Health of the Nation Outcome Scales for Children and Adolescents

- (HoNOSCA): Glossary for HoNOSCA score sheet. *The British Journal of Psychiatry*, 174, 428 – 431.
- Greenwood, A., & Burt, G. (2001). Validity of the brief psychiatric rating scale within a forensic inpatient hospital. *Expert Evidence*, 8, 15 – 30.
- Groth-Marnat, G. (2003). *Handbook of Psychological Assessment* (4th ed.). Hoboken, New Jersey: John Wiley and Sons, Inc.
- Hakvoort, L., Bogaerts, S., & Marinus, S. (2012). (Dys)Functional behavior in forensic psychiatric patients: Study of analogy between music therapy and group work. *The Arts in Psychotherapy*, 39, 304 – 313.
- Happel, B. (2008). The value of routine outcome measurement for consumers of mental health services: master or servant? *International Journal of Social Psychiatry*, 54, 317 – 327.
- 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*. Vancouver, BC: The British Columbia Institute against Family Violence.
- Hayward, M., & Moran, P. (2008). Comorbidity of personality disorders and mental illnesses. *Psychiatry*, 7, 102 – 104.
- Health Research Council of New Zealand. (2003). *Partnerships for evidence-based policy and practice*. Auckland: Health Research Council of New Zealand.
- Healthcare Management Advisors. (1999). *Coding audit of hospital medical records: First year*. Adelaide: Healthcare Management Advisors Ltd.
- Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York: John Wiley & Sons.

- Holloway, F. (2002). Outcome measurement in mental health - welcome to the revolution. *British Journal of Psychiatry*, 181, 1 – 2.
- Hopewell, S., McDonald, S., Clarke, M. J., & Egger, M. (2007). Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews*.
- Idaiani, S. (2011). Inter-rater reliability of Health of Nations Outcome Scale (HoNOS) among mental health nurses in Aceh. *Health Science Journal of Indonesia*, 2, 53 – 57.
- Jacobs, R. (2009). *Investigating patient outcome measures in mental health*. Retrieved from <http://www.york.ac.uk/che/pdf/rp48.pdf>.
- Jenkins, R. (1990). Towards a system of outcome indicators for mental health care. *British Journal of Psychiatry*, 157, 500 – 514.
- Kay, S. R., Opler, L. A., & Fiszbein, A. (2000). *The Positive and Negative Syndrome Scale (PANSS) Manual*. Toronto, ON: Multi-Health Systems Inc.
- Kazantzis, N., & Shinkfield, G. (2007). Conceptualising patient barriers to non-adherence with homework assignments. *Cognitive and Behavioural Practice*, 14, 317 – 324.
- Kennedy, H. G., O'Neill, C., Flynn, G., & Gill, P. (2010). *Dangerousness understanding, recovery and urgency manual (The DUNDRUM Quartet) v1.0.22*. Dublin National Forensic Mental Health Service.
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. T., Manderscheid, R. W., Walters, E. E., Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, 60, 184 – 189.

- Keulen-de Vos, M., & Schepers, K. (2016). Needs assessment in forensic patients: A review of instrument suites. *International Journal of Forensic Mental Health, 15*, 283 – 300.
- Kroppan, E., Nasset, M. B., Nonstad, K., Pedersen, T. W., Almvik, R., & Palmstierna, T. (2011). Implementation of the Short Term Assessment of Risk and Treatability (START) in a forensic high secure unit. *The International Journal of Forensic Mental Health, 10*, 7 – 12.
- Kwan, B. K., & Rickwood, D. J. (2015). A systematic review of mental health outcome measures for young people aged 12 to 25 years. *BMC Psychiatry, 15*, 279 – 298.
- Leese, M., Thornicroft, G., Shaw, J., Thomas, S., Mohan, R., Harty, M. A., & Dolan, M. (2006). Ethnic differences among patients in high-security psychiatric hospitals in England. *The British Journal of Psychiatry, 188*, 380 – 385.
- Long, C. G., Dickens, G., Sugarman, P., Craig, L., Mochty, U., & Hollin, C. R. (2010). Tracking risk profiles and outcome in a medium secure service for women: Use of the HoNOS-Secure. *International Journal of Forensic Mental Health, 9*, 215 – 225.
- Long, C. G., Webster, P., Waine, J., Motala, J., & Hollin, C. R. (2008). Usefulness of the CANFOR-S for measuring needs among mentally disordered offenders resident in medium or low secure hospital services in the UK: A pilot evaluation. *Criminal Behaviour and Mental Health, 18*, 39 – 48.
- Lowenkamp, C. T., Lovins, B., & Latessa, E. J. (2009). Validating the Level of Service Inventory–Revised and the Level of Service Inventory: Screening Version with a sample of probationers. *The Prison Journal, 89*, 192 – 204.
- Ludwick, D. (2009). *What is the effect of information and computing technology on healthcare?* University of Alberta, Edmonton.

- Maddison, P., Marlee, A. L., Webb, D. E., Berry, A. D. I., & Whitelock, T. (2016). The subscale structure of the HoNOS in a forensic population: A framework for a currency and tariff outcome system in mental health. *International Journal of Forensic Mental Health*, 15, 1 – 7.
- Marshall, M., Hogg, L. I., Gath, D. H., & Lockwood, A. (1995). The Cardinal Needs Schedule - a modified version of the MRC Needs for Care Assessment Schedule. *Psychological Medicine*, 25, 605 – 617.
- McColl, E., Jacoby, A., & Thomas, L. (Eds.). (2006). *Observing exposures and outcomes concurrently: Reporting surveys or cross-sectional studies* (2nd ed.). Philadelphia: American College of Physicians.
- McKay, R., & McDonald, R. (2008). Expensive detour or a way forward? The experience of routine outcome measurement in an aged care psychiatry service. *Australasian Psychiatry*, 16, 428 – 432.
- Mental Health Information Strategy Standing Committee. (2015). *Measuring recovery in Australian specialised mental health services: a status report*. Canberra: Australian Institute of Health and Welfare.
- Mental Health Standing Committee. (2006). *National statement of principles for forensic mental health*. Retrieved from <https://mhsa.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=38654706444>.
- Minichino, A., Francesconi, M., Carrión, R., Delle Chiaie, R., Bevilacqua, A., & Parisi, M. (2017). From neurological soft signs to functional outcome in young individuals in treatment with secondary services for non-psychotic disorders: a path analysis. *Psychological Medicine*, 1 – 12.
- Monahan, T., & Fisher, J. A. (2010). Benefits of ‘observer effects’: lessons from the field. *Qualitative Research*, 10, 357 – 376.

- Mueser, K. T., Gingerich, S., Salyers, M. P., McGuire, A. B., Reyes, R. U., & Cunningham, H. (2004). *The illness management and recovery (IMR) scales (client and clinician versions)*. Concord, New Hampshire: Dartmouth Psychiatric Research Center.
- Mullen, P. (2000). Forensic mental health. *The British Journal of Psychiatry*, 176, 307 – 311.
- Mullen, P. (2006). Schizophrenia and violence: from correlations to preventive strategies. *Advances in Psychiatric Treatment*, 12, 239 – 248.
- Mullen, P., & Ogloff, J. (2009). Providing mental health services to adult offenders in Victoria, Australia: Overcoming barriers. *European Psychiatry*, 24, 395 – 400.
- NACRO. (2006). *Liaison and diversion for mentally disordered offenders*. Retrieved from <https://3bx16p38bchl32s0e12di03h-wpengine.netdna-ssl.com/wp-content/uploads/2015/05/Liaison-and-diversion-for-mentally-disordered-offenders.pdf>.
- Nagi, C., Ostapiuk, E. B., Craig, L. A., Hacker, D., & Beech, A. R. (2009). Using the revised Problem Identification Checklist to predict inpatient and community violence: a pilot study. *British Journal of Forensic Practice*, 11, 4 – 13.
- National Institute for Mental Health in England. (2008). *Mental health outcomes compendium*. Retrieved from http://cpaa.co.uk/files/DH_093677.pdf
- National Mental Health Information Development Expert Advisory Panel. (2013). *Mental health national outcomes and casemix collection: NOCC strategic directions 2014 - 2024*. Canberra: Commonwealth of Australia.
- National Mental Health Working Group. (2003). *Mental health national outcomes and casemix collection: Technical specification of state and territory reporting*

- requirements for the outcomes and casemix components of 'agreed data', Version 1.50.* Canberra: Commonwealth Department of Health and Ageing.
- National Mental Health Working Group. (1999). *Mental health information development: National information priorities and strategies under the second national mental health plan 1998-2003.* Canberra: Commonwealth Department of Health and Aged Care.
- Nicholls, T. L., Brink, J., Desmarais, S., Webster, C. D., & Martin, M.-L. (2006). The Short-Term Assessment of Risk and Treatability (START), A prospective validation study in a forensic sample. *Assessment, 13*, 313 – 327.
- Nicholls, T. L., Viljoen, J., Cruise, K., Desmarais, S. L., & Webster, C. D. (2010). *Short-Term Assessment of Risk and Treatability: Adolescent Version (START:AV) (abbreviated manual).* Port Coquitlam, Canada: BC Mental Health and Addictions Services.
- Nonstad, K., Nasset, M. B., Kroppan, E., Pedersen, T. W., Nøttestad, J., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norwegian high secure hospital. *The International Journal of Forensic Mental Health, 9*, 294 – 299.
- NSW Agency for Clinical Innovation. (2015). NSW forensic patient needs and services project. Retrieved from <https://www.aci.health.nsw.gov.au/innovation-exchange/projects/nsw-forensic-patient-needs-services#>.
- O'Donahoo, J., & Simmonds, J. G. (2016). Forensic patients and forensic mental health in Victoria: legal context, clinical pathways, and practice challenges. *Australian Social Work, 69*, 169 – 180.
- O'Dwyer, S., Davoren, M., Abidin, Z., Doyle, E., McDonnell, K., & Kennedy, H. (2011). The DUNDRUM Quartet: validation of structured professional judgement

- instruments DUNDRUM-3 assessment of programme completion and DUNDRUM-4 assessment of recovery in forensic mental health services. *BMC Research Notes*, 4, 229.
- Ogloff, J., Davis, M., Rivers, G., & Ross, S. (2006). *The identification of mental disorders in the criminal justice system*. Canberra: Australian Institute of Criminology.
- Ogloff, J., Lemphers, A., & Dwyer, C. (2004). Dual diagnosis in an Australian forensic psychiatric hospital: prevalence and implications for services. *Behavioral Sciences and the Law*, 22, 543 – 562.
- Ogloff, J., Talevski, D., Lemphers, A., Wood, M., & Simmons, M. (2015). Co-occurring mental illness, substance use disorders, and antisocial personality disorder among clients of forensic mental health services. *Psychiatric Rehabilitation Journal*, 38, 16 – 23.
- Ohio Department of Mental Health. (2004). *Procedural manual* (6th ed.): Ohio Department of Mental Health.
- Osher, F., Scott, J. E., Steadman, H. J., & Robbins, P. C. (2004). *Validating a Brief Jail Mental Health Screen: Final report for National Institute of Justice*. Retrieved from Washington, DC: www.ncjrs.gov/pdffiles/nij/grants/213805.pdf.
- Ostapiuk, E., Stringer, I., & Craig, L. (2000). *Some modifications and revisions to the Problem Identification Checklist*. Unpublished manuscript. Available from Forensic Psychology Practice Ltd, 98 Sheffield Road, Boldmere, Sutton and Coldfield, Birmingham, UK.
- Otto, R. K., & Douglas, K. S. (Eds.). (2009). *Handbook of violence risk assessment*. New York, NY: Routledge: Taylor & Francis Group.
- Overall, J., & Gorham, D. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports*, 10, 799 – 812.

- Ovretveit, J. (1995). *Purchasing for health*. Buckingham: Open University Press.
- Owens, D. (2010). *The clinical utility of outcome measures in forensic mental health*. Paper presented at the Australian Mental Health Outcomes & Classification Network: National Outcomes Forum 2010, Adelaide, South Australia.
- Paulhus, D. L., & Vazire, S. (2009). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224 – 239). New York: Guilford Press.
- Pillay, S. M., Oliver, B., Butler, L., & Kennedy, H. G. (2008). Risk stratification and the care pathway. *Irish journal of psychological medicine*, 25, 123 – 127.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005). *Routine measurement of outcomes in Australia's public sector mental health services*. Melbourne: The University of Melbourne.
- Pirkis, J., Burgess, P., Kirk, P., Dodson, S., & Coombs, T. (2005). *Review of standardised measures used in the National Outcomes and Casemix Collection (NOCC)*. Sydney: NSW Institute of Psychiatry.
- Pirkis, J. E., Burgess, P. M., Kirk, P. K., Dodson, S., Coombs, T. J., & Williamson, M. K. (2005). A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health and quality of life outcomes*, 3, 1 – 12.
- Pirkis, J., & Callaly, T. (2010). Mental health outcome measurement in Australia. In T. Trauer (Ed.), *Measurement in mental health: Theory and practice* (pp. 15 – 25). Cambridge: Cambridge University Press.
- Private Mental Health Alliance. (2009). *Annual statistical report from the PMHA's centralised data management service regarding the services provided by participating private hospitals with psychiatric beds and private psychiatric day*

- hospitals: 1 July 2007 to 30 June 2008*. Canberra: Strategic Planning Group for Private Psychiatric Services.
- Quality Network for Forensic Mental Health Services. (2009). *Outcome and activity measures working group: Summary Document*. London: Royal College of Psychiatrists.
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- RANZCP. (2010). Mental health deserves a better deal. Retrieved from https://www.ranzcp.org/Files/Resources/Mental_health_deserves_a_better_deal_final-pdf.aspx.
- Relman, A. (1988). Assessment and accountability: the third revolution in health care. *New England Journal of Medicine*, 319, 1220 – 1222.
- Rock, D., & Preston, N. (2001). HoNOS: Is there any point in training clinicians? *Journal of Psychiatric and Mental Health Nursing*, 8, 405 – 409.
- Romeva, G. E., Rubio, L. G. G., Güerre, S. O., Miravet, M. J. R., Cáceres, A. G., & Thomas, S. D. M. (2010). Validación clínica de la escala Camberwell Assessment of Need-Forensic version, CANFOR para la evaluación de necesidades de personas con trastorno mental en el ámbito penitenciario. [Clinical validation of the CANFOR scale (Camberwell Assessment of Need-Forensic version) for the needs assessment of people with mental health problems in the forensic services]. *Actas Españolas de Psiquiatría*, 38, 129 – 137.
- Rosen, A., Hadzi-Pavlovic, D., & Parker, G. (1989). The Life Skills Profile: A measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin*, 15, 325 – 337.

- Rosen, A., Hadzi-Pavlovic, D., Parker, G., & Trauer, T. (2006). *The Life Skills Profile: Background, items and scoring for the LSP-39, LSP-20 and the LSP-16*. Retrieved from <http://www.blackdoginstitute.org.au/docs/LifeSkillsProfile.pdf>.
- Ruffles, J. (2010). *The management of forensic patients in Victoria: The more things change, the more they remain the same*. (PhD Thesis), Monash University, Melbourne.
- Saltman, D. C., Myers, L., Kendrick, J., & Fischer, G. C. (1998). Standardising health outcome measurements in general practice. *Australian Family Physician*, 2, 89 – 93.
- Salyers, M. P., Godfrey, J. L., Mueser, K. T., & Labriola, S. (2007). Measuring illness management outcomes: A psychometric study of clinician and consumer rating scales for illness self-management and recovery. *Community Mental Health Journal*, 43, 459 – 480.
- Saxena, S., Thornicroft, G., Knapp, M., & Whiteford, H. (2007). Resources for mental health: scarcity, inequity, and inefficiency. *Lancet*, 370, 878 – 889.
- Schuringa, E., Heininga, V. E., Spreen, M., & Bogaerts, S. (2016). Concurrent and predictive validity of the Instrument for Forensic Treatment Evaluation: From risk assessment to routine, multidisciplinary treatment evaluation. *International journal of offender therapy and comparative criminology*.
- Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-Rater and test-retest reliability, internal consistency, and factorial structure of the Instrument for Forensic Treatment Evaluation. *Journal of Forensic Psychology Practice*, 14, 127 – 144.
- Sederer, L. I., & Dickey, B. (1996). *Outcomes assessment in clinical practice*. Baltimore: Williams & Wilkins.

- Segal, A., Daffern, M., Thomas, S., & Ferguson, M. (2010). Needs and risks of patients in a statewide inpatient forensic mental health population. *International Journal of Mental Health Nursing*, 19, 223 – 230.
- Sharma, A., Dunn, W., O'Toole, C., & Kennedy, H.G. (2015). The virtual institution: Cross-sectional length of stay in general adult and forensic psychiatry beds. *International Journal of Mental Health Systems*, 9, 25 – 37.
- Shaw, J. (2002). Needs assessment for mentally disordered offenders is different. *The journal of Forensic Psychiatry & Psychology*, 13, 14 – 17.
- Shinkfield, G. (2006). *Therapist's experience of and attitudes towards barriers to the completion of therapeutic homework tasks in children, adolescents and families*. (Masters Thesis), Massey University, Auckland, New Zealand.
- Shinkfield, G., & Brennan, A. (2010). *What are we missing? Using outcome measures to identify rehabilitation needs in a forensic psychiatric population*. Paper presented at the Australian Mental Health Outcomes & Classification Network: Victoria Forum, Melbourne, Victoria.
- Shinkfield, G., & Ogloff, J. (2014). A review and analysis of routine outcome measures for forensic mental health services. *International Journal of Forensic Mental Health*, 13, 252 – 271.
- Shinkfield, G., & Ogloff, J. (2015). Use and interpretation of routine outcome measures in forensic mental health. *International Journal of Mental Health Nursing*, 24, 11 – 18.
- Shinkfield, G., & Ogloff, J. (2016). Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital. *The Journal of Forensic Psychiatry & Psychology*, 27, 867 – 885.

- Shinkfield, G., & Ogloff, J. (under review). Monitoring risk, security needs, clinical and social functioning within a forensic mental health population. *International Journal of Forensic Mental Health*.
- Simpson, D. D. (2005). *TCU treatment assessment forms*. Fort Worth: Texas Christian University, Institute of Behavioural Research.
- Slade, M. (2002). The use of patient-level outcomes to inform treatment. *Epidemiologia e Psichiatria Sociale*, 11, 20 – 27.
- Slade, M. (2002b). Routine outcome assessment in mental health services. *Psychological Medicine*, 32, 1339 – 1343.
- Slade, M., McCrone, P., Kuipers, E., Leese, M., Cahill, S., Parabiaghi, A., Priebe, S., Thornicroft, G. (2006). Use of standardised outcome measures in adult mental health services: Randomised controlled trial. *British Journal of Psychiatry*, 189, 330 – 336.
- Stedman, T., Yellowlees, P., Mellsop, G., Clarke, R., & Drake, S. (1997). *Measuring consumer outcomes in mental health: Field testing of selected measures of consumer outcome in mental health*. Canberra: Department of Health and Family Services.
- Sugarman, P., & Everest, H. (1999). *An audit of five regional secure psychiatric services*. South Thames NHS Multi-Authority Clinical Audit Group (available from the authors).
- Sugarman, P., & Walker, L. (2007). *Health of the Nation Scale for Users of Secure/Forensic Services (version 2b)*. Retrieved from <http://www.rcpsych.ac.uk/researchandtrainingunit/honos/secure.aspx>

- Sugarman, P., Walker, L., & Dickens, G. (2009). Managing outcome performance in mental health using HoNOS: experience at St Andrew's Healthcare. *Psychiatric Bulletin*, 33, 285 – 288.
- Te Pou. (2010). *Kawerongo hiko (e-bulletin) 5 March 2010*. Retrieved from <http://www.blacksheepcreative.co.nz/fanmail/action/view/?link=aJObo5yTZ3GS15qUb5XEK2-Wb2WUmXHAaGVnk5STk2Bxo50>.
- Te Pou. (2012). *The HoNOS family of measures: A technical review of their psychometric properties*. Auckland: Te Pou o Te Whakaaro Nui.
- Thomas, S., Daffern, M., Martin, T., Ogloff, J., Thomson, L., & Ferguson, M. (2009). Factors associated with seclusion in a statewide forensic psychiatric service in Australia over a 2-year period. *International journal of mental health nursing*, 18, 2 – 9.
- Thomas, S., Harty, M., Parrott, J., McCrone, P., Slade, M., & Thornicroft, G. (2003). *Camberwell Assessment of Need (CANFOR): a needs assessment for forensic mental health service users*. London: Gaskell.
- Thomas, S., McCrone, P., & Fahy, T. (2009). How do psychiatric patients on prison healthcare centres differ from inpatients in secure psychiatric inpatient units? *Psychology, Crime & Law*, 15, 729 – 742.
- Thomas, S., Slade, M., McCrone, P., Harty, M., Parrott, J., Thornicroft, G., & Leese, M. (2008). The reliability and validity of the forensic Camberwell Assessment of Need (CANFOR): a needs assessment for forensic mental health service users. *International Journal of Methods in Psychiatric Research*, 17, 111 – 120.
- Timmerman, I. G. H., Vastenburg, N. C., & Emmelkamp, P. M. G. (2001). The Forensic Inpatient Observation Scale (FIOS): development, reliability and validity. *Criminal Behaviour and Mental Health*, 11, 144 – 162.

- Trauer, T. (2003). *Analysis of outcome measurement data from the four Victorian 'round one' agencies*. Victoria: Mental Health Branch, Department of Human Services.
- Trauer, T. (Ed.) (2010). *Outcome measurement in mental health: theory and practice*. Cambridge: University Press.
- Trauer, T., Callaly, T., Hantz, P., Little, J., Shields, R., & Smith, J. (1999). Health of the Nation Outcome Scales: Results of the Victorian field trial. *The British Journal of Psychiatry*, 174, 380 – 388.
- Turner, T., & Salter, M. (2008). Forensic psychiatry and general psychiatry: re-examining the relationship. *Psychiatric Bulletin*, 32, 2 – 6.
- Veit, C. T., & Ware, J. E. J. (1983). The structure of psychological distress and well-being in general populations. *Journal of Clinical and Consulting Psychology*, 51, 730 – 745.
- Velligan, D., Lopez, L., Castillo, D., Manaugh, B., Milam, A., & Miller, A. (2011). Interrater reliability of using brief standardized outcome measures in a community mental health setting. *Psychiatric Services*, 62, 558 – 560.
- Ventura, J., Lukoff, D., Nuechterlein, K. H., Liberman, R. P., Green, M., & Shaner, A. (1993). Brief Psychiatric Rating Scale (BPRS): Expanded version 4.0 scales, anchor points, and administration manual. *International Journal of Methods in Psychiatric Research*, 3, 227 – 243.
- Vermillion, J. M., & Pfeiffer, S. I. (1993). Treatment outcome and continuous quality improvement: two aspects of program evaluation. *Psychiatric Hospital*, 24, 9 – 14.
- Vess, J. (2001). Development and implementation of a functional skills measure for forensic psychiatric inpatients. *Journal of Forensic Psychiatry & Psychology*, 12, 592 – 609.

- Victoria State Government. (2016). *Research and reporting*. Retrieved from <http://www.health.vic.gov.au/mhdr-info/rules.htm>.
- Victorian Institute of Forensic Mental Health. (2007). *Thomas Embling Hospital*. Retrieved from <http://www.forensicare.vic.gov.au/assets/pubs/TEH%202.pdf>.
- Victorian Law Reform Commission. (2014). *Review of the Crimes (Mental Impairment and Unfitness to be Tried) Act 1997: Consultation paper*. Melbourne: Victorian Law Reform Commission.
- Viera, A., & Garrett, J. (2005). Understanding inter-observer agreement: the kappa statistic. *Family medicine*, 37, 360 – 363.
- Wahi, W. M., Parks, D. V., Skeate, R. C., & Goldin, S. B. (2008). Reducing errors from the electronic transcription of data collected on paper forms: a research data case study. *Journal of the American Medical Informatics Association*, 15, 386 – 389.
- Walters, G., Kirkby, K., Marks, I., Whiteford, H., Andrews, G., & Swinson, R. (1996). Outcome measurement: sharing experiences in Australia. *Australasian Psychiatry*, 4, 316 – 318.
- Webster, J., Bretherton, F., Goulter, N., & Fawcett, L. (2013). Does an educational intervention improve the usefulness of the Health of the Nation Outcome Scales in an acute mental health setting? *International Journal of Mental Health Nursing*, 22, 322 – 328.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing the risk for violence (Version 2)*. Vancouver: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Webster, C. D., & Hucker, S. J. (2007). *Violence risk assessment and management*. Chichester, UK: Wiley.

- Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START) (Version 1.1)*. Port Coquitlam, BC: Forensic Psychiatric Services Commission and St. Joseph's Healthcare.
- Webster, C. D., Martin, M.-L., Brink, J. H., Nicholls, T. L., & Middleton, C. (2004). *Short-Term Assessment Risk and Treatability (START)*. Coquitlam, Canada: Forensic Psychiatric Services Commission.
- Webster, C. D., Nicholls, T., Martin, M.-L., Desmarais, S. L., & Brink, J. (2006). Short-Term Assessment of Risk and Treatability (START): The case for a structured professional judgment scheme. *Behavioral Sciences and the Law*, 24, 747 – 766.
- Wilson, C., Desmarais, S. L., Nicholls, T. L., & Brink, J. (2010). The role of client strengths in assessments of violence risk using the Short-Term Assessment of Risk and Treatability (START). *The International Journal of Forensic Mental Health*, 9, 282 – 293.
- Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *British Journal of Psychiatry*, 172, 11 – 18.
- Wong, S., & Gordon, A. (1996). *Violence Risk Scale – Experimental Version 1 (VRS)*. Saskatchewan, Canada: Regional Psychiatric Centre, Solicitor General of Canada.
- Woods, P., & Reed, V. (1999). The Behavioural Status Index (BSI): Some preliminary reliability studies. *International Journal of Psychiatric Nursing Research*, 5, 554 – 561.
- Wright, P., & Webster, C. D. (2011). Implementing structured professional judgment risk assessment schemes: an example of institutional change. *International Journal of Forensic Mental Health*, 10, 1 – 6.

Young, S. L., & Bullock, W. A. (2003). *The mental health recovery measure*. Toledo, OH: University of Toledo, Department of Psychology.

PART F: APPENDICES

Appendix A: Explanatory Statement

Appendix B: Consent Form

Appendix C: Ethics Approval – Monash University

Appendix D: Ethics Approval – Swinburne University of Technology

Appendix E: Operational Approval – Forensicare

Appendix F: Audit Protocol

Appendix G: Data Collection Tool (Phase One) – File Audit and Validity Check

Appendix H: Data Collection Tool (Phase Two) – HoNOS-Secure, LSI-R:SV, CANFOR

Appendix I: Copyright Notice – International Journal of Forensic Mental Health

Appendix J: Copyright Notice – International Journal of Mental Health Nursing

Appendix K: Copyright Notice – Journal of Forensic Psychiatry & Psychology

Appendix L: Comparison of HoNOS and HoNOS-Secure wording

APPENDIX A: EXPLANATORY STATEMENT

Explanatory Statement

Thomas Embling Hospital Consumers

Project Title: Measuring the Progress and Outcome of Patients at the Thomas Embling Hospital

This information sheet is for you to keep

My name is Gregg Shinkfield and I am a psychologist on the Barossa Unit of Thomas Embling Hospital. I am conducting a research project with Professor James Ogloff and Dr Stuart Thomas from the Centre for Forensic Behavioural Science, Monash University. This research will form the basis of a Doctoral Degree at Monash University. The findings from this study are intended to be published in a thesis, equivalent to a 300 page book. We will be inviting all patients currently residing in Thomas Embling Hospital to take part in this project. Participating in this study is completely voluntary and **YOU DO NOT HAVE TO PARTICIPATE**. Sometimes people may feel pressured to participate in research when they are asked by members of hospital staff. However, you are under no obligation to agree to participate in this study and you do not have to give your consent. If you feel uncomfortable about us asking you to participate in this study, please be assured that agreeing or not agreeing to participate will not affect your treatment, access to services, or discharge from Thomas Embling Hospital. Please discuss any concerns you have with your contact nurse.

The aim/purpose of the research

In Australia, public mental health services routinely complete a number of assessments with each of the patients in their service. These assessments were designed to assist with treatment planning and monitoring of patient progress. However, the assessment tools used for this task were developed in general mental health services, which are quite different to Thomas Embling Hospital. Because of this, it's possible that these tools might not be very good at assessing the needs of our consumers and may even give us wrong information. Therefore, the aim of this study is to test these tools at Thomas Embling Hospital, to find out whether they are useful, or if other tools might do a better job. We are also interested in checking how well staff use these tools and how accurately they record information.

Possible benefits

We think this study will benefit consumers of Thomas Embling Hospital, as it will help us identify which tools are best at assessing patient needs. Secondly, by checking how well staff use these tools, we can see if it would be beneficial to fine tune these tools or introduce others that might be more useful.

What does the research involve?

Firstly, to check that staff are using these tools properly, we need some way of measuring this. By comparing the assessments staff have already completed to information in patient files, we can check how accurately they were done. We can also check if they were done at the right times and whether any important information has been left out.

The second part of this research looks at whether other assessment tools might be more useful than those we already have. Over the past year, staff have been completing three additional assessments alongside those they already fill out. We would like to look at the information collected by these tools to see which ones best identified the needs of our consumers and predicted any difficulties that arose over time. Over the next two years, we would also like to check on the progress of participants to see which tools best predict consumer's real life outcomes over time. To do this, we would like to look at basic information held in databases kept by the Departments of Human Services, Justice, and Victoria police, as well as in clinical files here at Thomas Embling.

Specifically, we are seeking access to the following information:

- Routinely collected assessment information (Clinical files / RAPID database)
- Dates of transfer between units / discharge from TEH (Clinical files)
- Dates of contact with mental health services following discharge from TEH (Clinical Files / RAPID database)
- Dates of contact with the police or re-incarceration following discharge from TEH (Department Justice / Victoria Police)

The assessment tools that are being reviewed for this study are:

- HoNOS / HoNOS-Secure: A 12-item staff-rated measure of mental health and social functioning. HoNOS-Secure has an additional 7 items that assess security needs.
- Life Skills Profile-16: A 16-item clinician-rated measure of general functioning.
- CANFOR: A 25-item staff-rated measure that assesses broad domains in mental health and functioning.
- Level of Service Inventory-Revised (screening version): An 8-item clinician-rated measure of risk and consumer needs related to offending behaviour.

The databases that this information will be accessed from are:

- *RAPID Database* (Client Management Interface: CMI): A patient information and administration system used by public mental health services in Victoria to record details of consumer contacts with services and assessment information.
- *Prisoner Information Record* (PIR): An electronic database used by the Department of Justice to track admissions, discharges and transfers within Victoria's prisons.
- *Law Enforcement Assistance Program* (LEAP): An electronic database used by Victoria Police to record contact with individuals and particulars of crimes.

We understand that this information is personal and you may not want to be included in this study. That is why we are asking if you would be willing for us to access this information to help us evaluate these tools. In providing your consent, we offer you the assurance that this information will be used ONLY for the purpose of research. No personal information will be given to anyone, including your treating team or other members of staff. Finally, whether you decide to participate or not, this will have no impact on your treatment, access to services, or discharge from hospital.

How much time will the research take?

To help with this research, you don't have to do anything at all. The information needed for this study is collected about everyone who has contact with a mental health or forensic service. By providing your consent, we will simply access this information via the databases described above and will not ask you for anything else or for any more of your time.

Inconvenience / discomfort

We have made every effort to ensure there are no foreseeable risks to you by participating in this research. However, if you do feel discomfort or concern regarding this research we will remove you from this study, without question, at any time. We will also assist you to obtain any help you may need in relation to your feelings of discomfort regarding your participation.

Can I withdraw from the research?

Being in this study is entirely voluntary and if change your mind at any time you will be removed from the study without question.

Confidentiality

Only the research team will have access to your information. Any information collected will have your name and other identifying features removed before being included for analysis.

At the end of the study, no participants will be identified in any publications arising from this research. All of the information and results will be based on group information.

Storage of data

Storage of the data will adhere to Monash University regulations and will be kept on hospital premises in a locked filing cabinet for 5 years. A report from the study may be submitted for publication, but individual participants will not be identifiable in any such reports.

Use of data for other purposes

Because it's not always possible to think of all of the questions that could be answered when developing a study, with your consent, we may use data from this study to investigate future research questions. As this information will contain no names or identifying numbers, you will not be identifiable in any way.

Results

If you would like to be informed of the results of this study, please contact Professor James Ogloff on 9495-9160 or james.ogloff@med.monash.edu.au. The findings are accessible for 12 months after the end of the study.

If you would like to contact the researchers about any aspect of this study, please contact the Chief Investigator:	If you have a complaint concerning the manner in which this research CF10/3127 – 2010001685 is being conducted, please contact:
Prof. James Ogloff Centre for Forensic Behavioural Science 505 Hoddle Street, Clifton Hill, 3068, Victoria Tel: +61 3 9947 2600 Fax: +61 3 9947 2650 Email: james.ogloff@monash.edu.au	Executive Officer, Human Research Ethics Monash University Human Research Ethics Committee (MUHREC) Building 3e Room 111 Research Office Monash University VIC 3800 Tel: +61 3 9905 2052 Fax: +61 3 9905 3831 Email: muhrec@adm.monash.edu.au

Thank you.

Prof. James Ogloff

Gregg Shinkfield

APPENDIX B: CONSENT FORM

Consent Form

Project Title: Measuring the Progress and Outcome of Patients at Thomas Embling Hospital

NOTE: This consent form will remain with the Monash University researcher for their records

Participants Name: _____

I **consent** / **do not consent** (please circle) to participate in the Monash University research project specified above. I have had the project explained to me, and I have read the Explanatory Statement, which I keep for my records.

I understand that agreeing to take part means that I give permission for the researchers to:

- Access my Forensicare clinical files Yes ☐ No ☐
- Access my mental health records on the Client Management Interface Database (CMI-ODS) Yes ☐ No ☐
- Access my files on the police Law Enforcement Assistance Program Database (LEAP) Yes ☐ No ☐
- Access my files in the Prisoner Information Record (PIR) Yes ☐ No ☐
- I agree to this information being stored in a non-identifiable (de-identified form) for use in future research projects Yes ☐ No ☐

I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and I can withdraw at any stage of the project without being penalised or disadvantaged in any way.

I understand that any data the researcher extracts for use in reports or published findings will, under no circumstances, contain any names or identifying characteristics.

I understand that any information I permit the researcher to access is confidential, and that no information that could identify me will be used in any reports on the project, or given to any other party.

I understand that data from the above named sources will be kept in secure storage and accessible to the research team. I also understand that the data will be destroyed after a 5 year period unless I consent to it being used in future research.

Participant's name

Signature

Date

APPENDIX C: ETHICS APPROVAL (MONASH UNIVERSITY)

Human Ethics Certificate of Approval

Date: 24 March 2011

Project Number: CF10/3127 - 2010001685

Project Title: Measuring the progress and outcome of patients at Thomas Embling Hospital

Chief Investigator: Prof James Ogloff

Approved: From: 24 March 2011 to 24 March 2016

Terms of approval

1. The Chief investigator is responsible for ensuring that permission letters are obtained, if relevant, and a copy forwarded to MUHREC before any data collection can occur at the specified organisation. **Failure to provide permission letters to MUHREC before data collection commences is in breach of the National Statement on Ethical Conduct in Human Research and the Australian Code for the Responsible Conduct of Research.**
2. Approval is only valid whilst you hold a position at Monash University.
3. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval and to ensure the project is conducted as approved by MUHREC.
4. You should notify MUHREC immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.
5. The Explanatory Statement must be on Monash University letterhead and the Monash University complaints clause must contain your project number.
6. **Amendments to the approved project (including changes in personnel):** Requires the submission of a Request for Amendment form to MUHREC and must not begin without written approval from MUHREC. Substantial variations may require a new application.
7. **Future correspondence:** Please quote the project number and project title above in any further correspondence.
8. **Annual reports:** Continued approval of this project is dependent on the submission of an Annual Report. This is determined by the date of your letter of approval.
9. **Final report:** A Final Report should be provided at the conclusion of the project. MUHREC should be notified if the project is discontinued before the expected date of completion.
10. **Monitoring:** Projects may be subject to an audit or any other form of monitoring by MUHREC at any time.
11. **Retention and storage of data:** The Chief Investigator is responsible for the storage and retention of original data pertaining to a project for a minimum period of five years.



Professor Ben Canny
Chair, MUHREC

Cc: Dr Stuart Thomas; Mr Gregg Shinkfield

APPENDIX D: ETHICS APPROVAL
(SWINBURNE UNIVERSITY OF TECHNOLOGY)

From: Astrid Nordmann <anordmann@swin.edu.au>

Date: 16 August 2016 at 14:48:49 AEST

To: James Ogloff <jogloff@swin.edu.au>

Cc: RES Ethics <resethics@swin.edu.au>

Subject: SHR Project 2016/222 - Ethics clearance (expedited approval based on Monash University HREC approval CF10/3127 - 2010001685)

To: Prof. Jim Ogloff, CFBS

SHR Project 2016/222 – Measuring the progress and outcome of patients at Thomas Embling Hospital

Prof J Ogloff AM, CFBS/FHAD and Forensicare; Mr Gregg Shinkfield (Student)

Approved Duration: 16-08-2016 to 08-03-2019

(Monash University HREC ref: CF10/3127 – 2010001685)

I refer to your application for Swinburne ethics clearance for the above project.

Relevant documentation pertaining to the application, as emailed on 15 July 2016 with attachment, was given expedited ethical review on behalf of Swinburne's Human Research Ethics Committee (SUHREC) by a delegate significantly on the basis of the ethical review conducted by the Monash University Human Research Ethics Committee (MUHREC ref: CF10/3127 – 2010001685). In reviewing the documentation, it was noted that while MUHREC ethics clearance was given until 24 March 2016, all data collection was completed prior to Mr Gregg Shinkfield's enrolment at Swinburne University of Technology.

I am pleased to advise that, as submitted to date and as regards Swinburne, ethics clearance has been given for the above project to proceed in line with standard on-going ethics clearance conditions outlined below and as follows. MUHREC may need to be apprised of the Swinburne ethics clearance.

- All human research activity undertaken under Swinburne auspices must conform to Swinburne and external regulatory standards, including the *National Statement on Ethical Conduct in Human Research* and with respect to secure data use, retention and disposal.
- The named Swinburne Chief Investigator/Supervisor remains responsible for any personnel appointed to or associated with the project being made aware of ethics clearance conditions, including research and consent procedures or instruments approved. Any change in chief investigator/supervisor requires timely notification and SUHREC endorsement.
- The above project has been approved as submitted for ethical review by or on behalf of SUHREC. Amendments to approved procedures or instruments ordinarily require prior ethical appraisal/clearance. SUHREC must be notified immediately or as soon as possible thereafter of (a) any serious or unexpected adverse effects on participants and any redress measures; (b) proposed changes in protocols; and (c) unforeseen events which might affect continued ethical acceptability of the project.
- At a minimum, an annual report on the progress of the project is required as well as at the conclusion (or abandonment) of the project. Information on project monitoring, self-audits and progress reports can be found on the Research Intranet [pages](#). (However, formats required by or submissions to Monash University HREC in this regard may be acceptable all things being equal.)
- A duly authorised external or internal audit of the project may be undertaken at any time.

Please contact the Research Ethics Office if you have any queries about on-going ethics clearance as regards Swinburne, citing the Swinburne project number. Please retain a copy of this email as part of project record-keeping.

Yours sincerely,
Astrid Nordmann

Dr Astrid Nordmann | Research Ethics Coordinator
Swinburne Research | Swinburne University of Technology
Ph +61 3 9214 3845 | anordmann@swin.edu.au
Level 1, Swinburne Place South
24 Wakefield St, Hawthorn VIC 3122, Australia
www.swinburne.edu.au



APPENDIX E: OPERATIONAL APPROVAL – FORENSICARE



Forensicare

Victorian Institute of
Forensic Mental Health

Yarra Bend Road
Fairfield Victoria 3078

Locked Bag 10
Fairfield Victorian 3078
DX 212119

Telephone +61 3 9495 9143
Facsimile +61 3 9495 9190
www.forensicare.vic.gov.au

2 September 2010

Dear Mr Shinkfield

Re: Measuring the progress and outcome of patients at Thomas Embling Hospital

The Forensicare Research Committee has given operational approval for your research to be conducted at Forensicare. This approval is subject to approval by the Monash University Human Research Ethics Committee.

You may not commence the research until you provide a letter of approval from the Department of Justice Human Research Ethics Committee, and you receive a letter from the Forensicare Research Committee acknowledging receipt of the approval letter.

Approval is given for the period between the anticipated commencement and completion dates as set out in the documentation. If the study has not been completed by the nominated completion date, an application for extension will be required.

To enable the Committee to meet its obligations in relation to monitoring Forensicare's research program, you are required to provide a report within 12 months or on completion of your project, whichever is earlier.

Forensicare must report ongoing research activities to the Minister of Mental Health quarterly. As such you may be asked to provide information on the progress of your research.

Please ensure that the Research Committee is notified of any matter that arises that may affect the conduct of the approved program.

Should you have any queries please don't hesitate to contact Ms Mitali Gupta on 99472601 or email mitali.gupta@forensicare.vic.gov.au.

Yours sincerely

Professor Mairead Dolan
Assistant Clinical Director (Research)
Forensicare

APPENDIX F: AUDIT PROTOCOL

Note. The protocol manual presented includes all of the variables collected during data collection. A number of variables were not included in the analyses reported in this thesis but will be analysed at a later date. For the purposes of transparency the whole manual is included.



**Measuring the Progress and Outcome of
Patients at the Thomas Embling Hospital:
An Evaluation of Clinical Service**

Dear Auditor,

Thank you for agreeing to assist with data collection for the first phase of this project. The initial task we wish to complete is an audit of current patient files, with the aim of evaluating how accurately and reliably outcome measures are being completed at present. We are focusing this portion of our study on those outcome measures completed during the acute period of a patient's admission to TEH. To standardise the data collected, we have therefore chosen to audit the *first three* outcome measures completed for each patient. Namely, those completed on *admission* and at the *90 and 180 day reviews*.

The methodology for this audit is as follows:

1) Patient Files:

A randomised list of patients whose files are to be audited has been generated.

To conduct the audit, please collect volumes 1 & 2 of the clinical file for each patient identified. These should contain the Outcome Measure record forms covering the period of interest for this study (i.e., *admission*, *90 day* and *180 day* reviews).

For patients who have had lengthy admissions at TEH these files may need to be obtained from file storage.

2) Compliance Audit:

To facilitate data collection, a reporting tool has been developed to assist us to obtain the information required.

a. On *page one* of the data collection tool, please record the following:

Sample Number : _____ 1

Admission Date : _____ 2

Admission Length : Less than 3 months ☐ 3
Between 3 - 7 months ☐ 4
Over 7 months ☐ 5

Outcome Measures Completed : None Completed ☐ 6
On Admission ☐ 7
90 Day Review ☐ 8
180 Day Review ☐ 9

If present, use 90 day review for Validity Check (pp.5) ☐

- i) The *sample number* of the patient (taken from the randomised list)
- ii) *Length of admission* for that patient. This is broken into three bands: less than three months, between three – seven months, over seven months
- iii) Which outcome measures have been *completed* (i.e., none, admission, 90 day review, and/or 180 day review)

b. On pages two, three and four, space has been provided in which to record data for each of the three sets of Outcome Measures audited.

Pages two, three and four each relate to a specific group of Outcome Measures, based on the time they were recorded. That is, page two relates to the initial assessment on admission, page three to the 90 day review, and page four to the 180 day review.

At the top of pages 2, 3, and 4, response boxes have been provided for you to indicate the date on which the outcome measure was completed by the clinician. However, if the data for that period is missing, please use the same space to record whether the patient had been discharged by that date or if the measures had simply not been completed. *Please complete only one box in this section.*

TEH Outcome Measures Compliance Study

ADMISSION

Date
Data Missing
Not in TEH

1) The date on which the outcome measure was completed;

2) Outcome measures are missing or not completed;

3) Or, the patient had been discharged prior to their review.

If an Outcome Measure form has been completed over multiple dates, please record the earliest date.

- c. On pages *two, three and four*, please record information directly from the completed Outcome Measures forms regarding the *admission, 90 and 180 day reviews*.

To complete this section, inspect the original Outcome Measures form (located in the patient's clinical file) and check either the NO/YES box depending on whether or not these aspects have been completed.

Attachment One provided item descriptions to clarify how each item is identified and rated.

HoNOS - (from file)

	0	1	
11 Completed	NO	YES	→ 13 Total Missing <u>2</u>
12 Items Missing	NO	YES	
14 Reason For Completing Identified	NO	YES	
15 Reason For Completing Accurate	NO	YES	
16 BASIS 32 Check Completed	NO	YES	
17 Signed	NO	YES	
18 Designation	NO	YES	
19 Date	NO	YES	

(Note. Repeat the above procedure for the HoNOS, LSP, and BASIS-32)

- d. The RAPID items refer to whether or not it is indicated on the record form that the data has been entered into RAPID.

RAPID		
	0	1
20 Entered	NO	YES
21 Form Dated	NO	YES
22 Form Signed	NO	YES

Note. This is irrespective of whether or not it has indeed been entered. There is no need to check the RAPID CMI database to ascertain the validity of this.

3) Validity Check (Page 5 of Audit Form)

An important aspect of this audit is to assess the *validity* of ratings provided by clinical staff. As an experienced clinician we are seeking to draw upon your expertise to re-rate *one set* of Outcome Measures for each patient audited. This is to be done based on a

review of clinical entries (i.e., continuation notes) covering the same two week period as rated by the original clinician. To perform this validity check, the follow methodology has been established:

- a. One set of outcome measures per patient is selected to be re-rated. To determine which set of measures is to be selected, use the following criteria:
 - i. If the **90 day** review has been *completed* and is *available*, **this is selected**
 - ii. If the 90 day review is missing, but the 180 day review is available, the 180 day review is selected.
 - iii. If neither the 90 or 180 day review has been completed or is available, then select the admission data.
 - iv. If no outcome measures have been completed, or all are unavailable, then record "None Completed"

This information is also displayed in a flow chart located in Attachment 2

TEH Outcome Measures Compliance Study									
VALIDITY CHECK	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 80%; padding: 2px;">Admission</td> <td style="width: 20%; text-align: center; padding: 2px;">1/2/04</td> </tr> <tr> <td style="padding: 2px;">90 day review</td> <td style="text-align: center; padding: 2px;">180</td> </tr> <tr> <td style="padding: 2px;">180 day review</td> <td style="text-align: center; padding: 2px;"></td> </tr> <tr> <td style="padding: 2px;">None Completed</td> <td style="text-align: center; padding: 2px;"></td> </tr> </table>	Admission	1/2/04	90 day review	180	180 day review		None Completed	
Admission	1/2/04								
90 day review	180								
180 day review									
None Completed									

- b. *Date Completed:* Having selected which set of Outcome Measures is to be used to perform the validity check, record the date on which they were first completed
- c. *Period Rated:* Using the calendar charts provided, determine the date TWO WEEKS prior to the completion of the HoNOS. Record the date range covered by this assessment.

(e.g., If the HoNOS was originally completed on the 15/08/09, the period covered by the HoNOS would be 01/08/09 ←→ 15-08-09). This is therefore the period of file entries that are to be reviewed for the validity check.

- d. *Clinical Admin Ratings:* Review the clinical file entries made within the two week period identified above. Based on the information contained within the clinical notes, provide your own independent rating of the patient for each of the Outcome Measure items.

Note.

- i. If you are unable to rate an item based on file information, mark this as "DK" (don't know)
 - ii. Repeat process for both HoNOS and ALSP
- e. *Clinician Rating:* Once you have provided your own independent ratings, please obtain the ratings originally determined by the assessing clinician at the time the Outcome Measures were completed and transpose them in the space provided

TEH Outcome Measures Compliance Study

VALIDITY CHECK

Admission ^{1,2,3}
 90 day review ☒ 100
 180 day review ☐

HoNOS

Date Completed:

02 / 04 / 08 101

Period Rated (2 weeks):

19 / 03 / 08 102 to 02 / 04 / 08 103

#	Item Content	Clinican Rating	Clinical Admin Rating
1	Overactive/Aggressive/Disruptive Behaviour	<u>0</u> 104	<u>1</u> 116
2	Non-Accidental Self-injury	<u>0</u> 105	<u>0</u> 117
3	Problem Drinking or Drug taking	<u>0</u> 106	<u>0</u> 118
4	Cognitive problems	<u>1</u> 107	<u>1</u> 119
5	Physical illness / Disability Problems	<u>0</u> 108	<u>0</u> 120
6	Hallucinations and delusions	<u>3</u> 109	<u>3</u> 121
7	Depressed mood	<u>2</u> 110	<u>1</u> 122
8	Other Mental and Behavioural Problems	<u>2 (B)</u> 111	<u>DK</u> 123
9	Problems with Relationships	<u>2</u> 112	<u>2</u> 124
10	Problems with Activities of Daily Living	<u>1</u> 113	<u>1</u> 125
11	Problems with Living Conditions	<u>0</u> 114	<u>DK</u> 126
12	Problems with Occupation and Activities	<u>1</u> 115	<u>1</u> 127

Attachment One – Item explanations

HoNOS

<i>Completed</i>	Has the HoNOS been filled in?
<i>Items Missing</i>	Have all HoNOS items been rated
- <i>Total Missing</i>	- If items are missing, record the number of HoNOS items not rated
<i>Reason for Completing Identified</i>	Has the rating clinician indicated the reason for completing the HoNOS?
<i>Reason for Completing Accurate</i>	Given the time at which the form was completed, has the rating clinician identified the correct reason for completing the HoNOS (i.e., admission, 90 day review)?
<i>BASIS 32 Check</i>	Has the rating clinician recorded whether or not the BASIS 32 was offered to patient?
<i>Signed</i>	Has the rating clinician signed the completed form?
<i>Designation</i>	Has the clinician indicated their designation (e.g., PSEN, RPN2, NUM, Psychologist, Occupational Therapist)
<i>Date</i>	Has the clinician recorded the date that the form was completed?

RAPID - On the record form, have the following been indicated:

<i>Entered</i>	Is there some indication that this data has been entered into RAPID (e.g., stamped, signed, a note written)?
<i>Dated</i>	Has the date on which this data was entered into Rapid been recorded?
<i>Signed</i>	Has the name of the person who entered this data been recorded, or has the form been signed?

ALSP (Abbreviated Life Skills Profile)

<i>Completed</i>	Has the ALSP been filled in?
<i>Items Missing</i>	Have all ALSP items been rated
- <i>Total Missing</i>	- If items are missing, record the number of HoNOS items not rated
<i>Focus of Care</i>	Has the main focus of care of the past 2 weeks been recorded (e.g. acute, function gain, intensive, maintenance)
<i>Signed</i>	Has the rating clinician signed the completed form?
<i>Designation</i>	Has the clinician indicated their designation (e.g., RPN2, Psychologist, OT)
<i>Date</i>	Has the clinician recorded the date the form was completed

RAPID - On the record form, have the following been indicated:

<i>Entered</i>	Is there some indication that this data has been entered into RAPID (e.g., stamped, signed, a note written)?
<i>Dated</i>	Has the date on which this data was entered into Rapid been recorded?
<i>Signed</i>	Has the name of the person who entered this data been recorded, or has the form been signed?

BASIS-32

<i>Completed</i>	Has the HoNOS been filled in?
<i>Consent Indicated</i>	Has the patient indicated their consent for data being recorded for the purpose of outcome measurement (e.g., strike out either "consent / do not consent")
<i>Consent Portion of Form Signed</i>	Has the patient signed consent portion of the BASIS-32 form?
<i>Date</i>	Has the date on which consent was given been recorded?

RAPID - On the record form, have the following been indicated:

<i>Entered</i>	Is there some indication that this data has been entered into RAPID (e.g., stamped, signed, a note written)?
<i>Dated</i>	Has the date on which this data was entered into Rapid been recorded?
<i>Signed</i>	Has the name of the person who entered this data been recorded, or has the form been signed?

Note. There is no place specified on the BASIS-32 form in which to record if/when this data was entered into RAPID. For the purpose of this study, we are simply interested in whether or not some indication has been made to this effect.

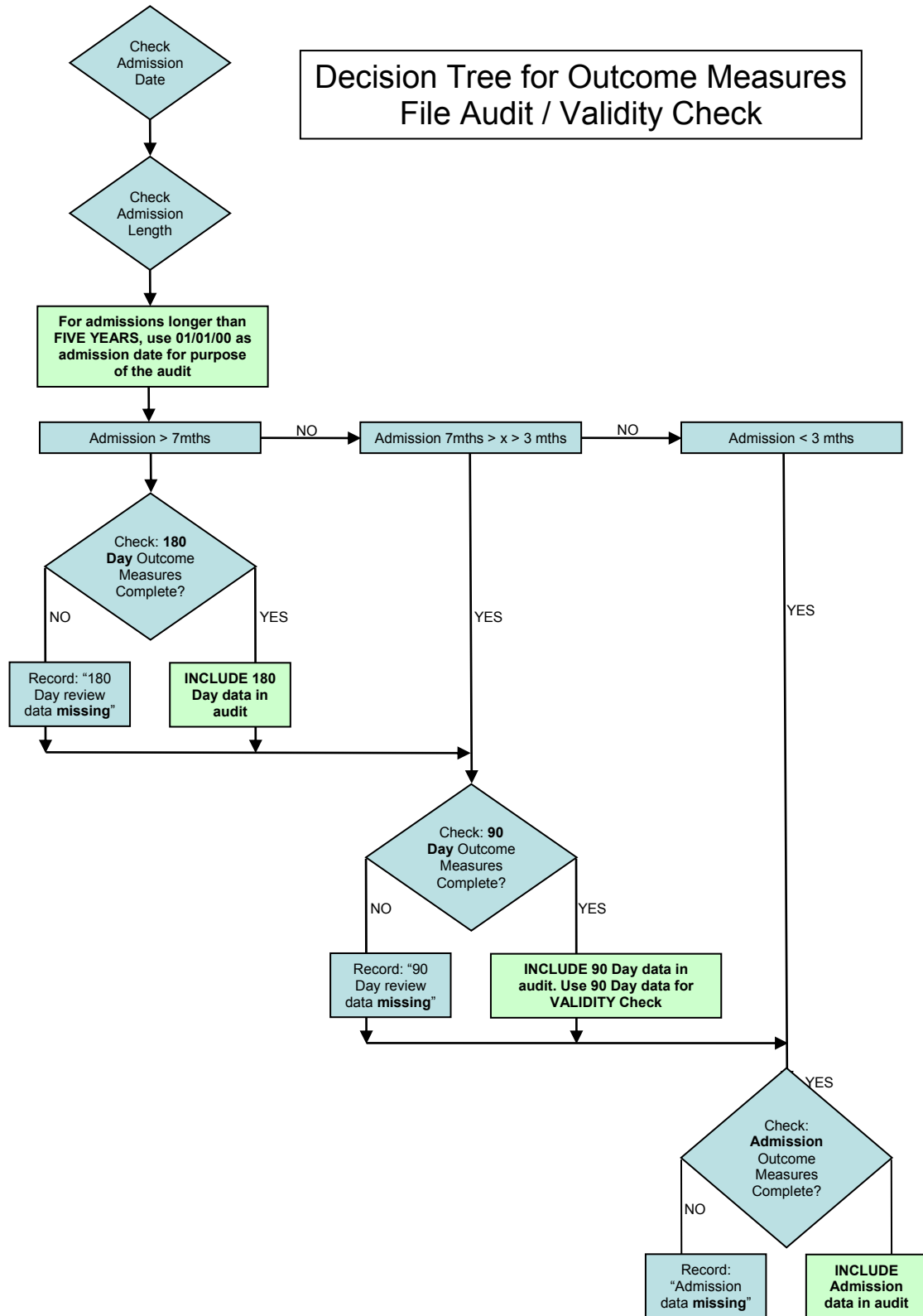
VALIDITY CHECK**HoNOS**

<i>Date Completed</i>	Record the date on which the HoNOS form selected for the Validity check was initially completed by the rating clinician. This is recorded to ensure the correct period is selected for file review.
<i>Period Rated</i>	Using the calendar charts provided determine the date TWO WEEKS prior to the completion of the HoNOS. Record the date range covered by this assessment. (e.g., if the HoNOS was originally completed on the 15/08/09, the period covered by the HoNOS would be 01/08/09 ←→ 15-08-09)
<i>Clinician Rating</i>	These are the original ratings determined by the rating clinician at the time the measure was completed.
<i>Clinical Admin Rating</i>	These are the ratings determined by clinical admin after reviewing clinical entries from the period specified.
<i>Item Content</i>	These are the item descriptors taken directly from the HoNOS form

ALSP

<i>Date Completed</i>	Record the date on which the HoNOS selected for the Validity check was completed by the rating clinician. This is recorded to ensure the correct period is selected for file review.
<i>Period Rated</i>	Using the calendar charts provided determine the date TWO WEEKS prior to the completion of the HoNOS. Record the date range covered by this assessment.
<i>Clinician Rating</i>	These are the original ratings determined by the rating clinician at the time the measure was completed.
<i>Clinical Admin Rating</i>	These are the ratings determined by clinical admin after reviewing clinical entries from the period specified.

Attachment Two – Flow Chart reference for validity check



APPENDIX G: DATA COLLECTION TOOL (PHASE ONE)

FILE AUDIT AND VALIDITY CHECK

TEH Outcome Measures Compliance Study

PILOT STUDY

Sample Number : _____ 1

Admission Date : _____ 2

Admission Legnth : Less than 3 months ☐ 3
 Between 3 - 7 months ☐ 4
 Over 7 months ☐ 5

Outcome Measures Completed : None Completed ☐ 6
 On Admission ☐ 7
 90 Day Review ☐ 8
 180 Day Review ☐ 9

If present, use 90
day review for
Validity Check
(pp.5)
and DASA
(pp. 3)

**FOR PATIENTS ADMITTED PRIOR TO 2000, PLEASE USE 01/01/2000
AS THE DATE OF ADMISSION FOR THIS STUDY**

Notes:

TEH Outcome Measures Compliance Study

ADMISSION

Date

Data Missing 10

Not in TEH

HoNOS - (from file)

- 11 Completed
- 12 Items Missing
- 14 Reason For Completeing *Identified*
- 15 Reason For Completeing *Accurate*
- 16 BASIS 32 Check Completed
- 17 Signed
- 18 Designation
- 19 Date

0	1
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES



13 Total Missing

RAPID

	0	1
20 Entered	NO	YES
21 Form Dated	NO	YES
22 Form Signed	NO	YES

ALSP - (from file)

- 23 Completed
- 24 Items Missing
- 26 Focus Of Care
- 27 CRS Score
- 28 Signed
- 29 Designation
- 30 Date

0	1
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES



25 Total Missing

RAPID

	0	1
31 Entered	NO	YES
32 Form Dated	NO	YES
33 Form Signed	NO	YES

BASIS-32 - (from file)

- 34 Completed
- 35 Signed
- 36 Date

0	1
NO	YES
NO	YES
NO	YES

RAPID

	0	1
37 Entered	NO	YES
38 Form Dated	NO	YES
39 Form Signed	NO	YES

COMPLETE THE FOLLOWING ONLY IF ADMISSION DATA IS USED FOR VALIDITY CHECK

DASA

- 163 Was the DASA completed for this patient during the previous two week period

Skip to next section

- 164 Number of ratings recorded

0	1
NO	YES

DASA Scores

Week 1	Week2
165	172
166	173
167	174
168	175
169	176
170	177
171	178

TEH Outcome Measures Compliance Study

91 DAY REVIEW

Date
 Data Missing 40
 Not in TEH

HoNOS - (from file)

- 41 Completed
- 42 Items Missing
- 44 Reason For Completeing Identified
- 45 Reason For Completeing *Accurate*
- 46 BASIS 32 Check Completed
- 47 Signed
- 48 Designation
- 49 Date

0	1
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES



43 Total Missing

RAPID

	0	1
50 Entered	NO	YES
51 Date	NO	YES
52 Signed	NO	YES

ALSP - (from file)

- 53 Completed
- 54 Items Missing
- 56 Focus Of Care
- 57 CRS Score
- 58 Signed
- 59 Designation
- 60 Date

0	1
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES



55 Total Missing

RAPID

	0	1
61 Entered	NO	YES
62 Date	NO	YES
63 Signed	NO	YES

BASIS-32 - (from file)

- 64 Completed
- 65 Signed
- 66 Date

0	1
NO	YES
NO	YES
NO	YES

RAPID

	0	1
67 Entered	NO	YES
68 Form Dated	NO	YES
69 Form Signed	NO	YES

COMPLETE THE FOLLOWING ONLY IF 90 DAY REVIEW DATA IS USED FOR VALIDITY CHECK

DASA

- 163 Was the DASA completed for this patient during the previous two week period

Skip to next section

- 164 Number of ratings recorded

0	1
NO	YES

DASA Scores

Week 1	Week2
165	172
166	173
167	174
168	175
169	176
170	177
171	178

TEH Outcome Measures Compliance Study

182 DAY REVIEW

Date
 Data Missing 70
 Not in TEH

HoNOS - (from file)

- 71 Completed
- 72 Items Missing
- 74 Reason For Completeing Identified
- 75 Reason For Completeing *Accurate*
- 76 BASIS 32 Check Completed
- 77 Signed
- 78 Designation
- 79 Date

0	1
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES

73 Total Missing

RAPID

	0	1
80 Entered	NO	YES
81 Form Dated	NO	YES
82 Form Signed	NO	YES

ALSP - (from file)

- 83 Completed
- 84 Items Missing
- 86 Focus Of Care
- 87 CRS Score
- 88 Signed
- 89 Designation
- 90 Date

0	1
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES
NO	YES

85 Total Missing

RAPID

	0	1
91 Entered	NO	YES
92 Form Dated	NO	YES
93 Form Signed	NO	YES

BASIS-32 - (from file)

- 94 Completed
- 95 Signed
- 96 Date

0	1
NO	YES
NO	YES
NO	YES

RAPID

	0	1
97 Entered	NO	YES
98 Form Dated	NO	YES
99 Form Signed	NO	YES

COMPLETE THE FOLLOWING ONLY IF 180 DAY REVIEW DATA IS USED FOR VALIDITY CHECK

DASA

- 163 Was the DASA completed for this patient during the previous two week period

Skip to next section

- 164 Number of ratings recorded

0	1
NO	YES

DASA Scores

Week 1	Week2
165	172
166	173
167	174
168	175
169	176
170	177
171	178

TEH Outcome Measures Compliance Study

VALIDITY CHECK

Admission	1,2,3,4	
90 day review		100
180 day review		
None Completed		

HoNOS

Date Completed: _____ 101
 Period Rated (2 weeks): _____ 102 to _____ 103

#	Item Content	Clinican Rating	Clinical Admin Rating
1	Overactive/Aggressive/Disruptive Behaviour	104	116
2	Non-Accidental Self-injury	105	117
3	Problem Drinking or Drug taking	106	118
4	Cognitive problems	107	119
5	Physical illness / Disability Problems	108	120
6	Hallucinations and delusions	109	121
7	Depressed mood	110	122
8	Other Mental and Behavioural Problems	111	123
9	Problems with Relationships	112	124
10	Problems with Activities of Daily Living	113	125
11	Problems with Living Conditions	114	126
12	Problems with Occupation and Activities	115	127

ALSP

Date Completed: _____ 128
 Period Rated (3 months): _____ 129 to _____ 130

#	Item Content	Clinican Rating	Clinical Admin Rating
1	Initiating and Responding to conversation	131	147
2	Withdrawal from social contact	132	148
3	Warmth to others	133	149
4	Personal Grooming	134	150
5	Clean Clothing	135	151
6	Neglect of physical health	136	152
7	Violence to others	137	153
8	Make/keep friendships	138	154
9	Maintenance of adequate diet	139	155
10	Compliance with prescribed medication	140	156
11	Willingness to take medication	141	157
12	Cooperation with health services	142	158
13	Problems with others in household	143	159
14	Offensive behaviour	144	160
15	Irresponsible behaviour	145	161
16	Work capability	146	162

Note. Clinican rating is taken from original outcome measure completed by treating team. Clinical Admin Rating is obtained by reviewing clinical notes from the two week period prior to the date of the clinician rating.

APPENDIX H: DATA COLLECTION TOOL (PHASE TWO)**HoNOS-SECURE, LSI-R:SV, CANFOR**



Outcome Measures Evaluation Study

Local UR No	Unit:
Given Name	Surname:
Rated By:	Date:

HoNOS-Secure

Security Scales:							D. Rate need for a safely-staffed living environment	0	1	2	3	4	9
A. Rate risk of harm to adults or children	0	1	2	3	4	9	E. Rate need for escort on leave	0	1	2	3	4	9
B. Rate risk of self harm	0	1	2	3	4	9	F. Rate risk to individual from others	0	1	2	3	4	9
C. Rate need for building security to prevent escape	0	1	2	3	4	9	G. Rate need for risk management procedures	0	1	2	3	4	9
Clinical Scales							7 Problems with depressed mood	0	1	2	3	4	9
1 Overactive, aggressive, disruptive or agitated	0	1	2	3	4	9	8 Other mental and behavioural problems	0	1	2	3	4	9
2 Non-accidental self-injury	0	1	2	3	4	9	8A Specify disorder [†] (A, B, C, D, E, F, G, H, I or J)						
3 Problem drinking or drug taking	0	1	2	3	4	9	9 Problems with relationships	0	1	2	3	4	9
4 Cognitive problems	0	1	2	3	4	9	10 Problems with activities of daily living	0	1	2	3	4	9
5 Physical illness or disability problems	0	1	2	3	4	9	11 Problems with living conditions	0	1	2	3	4	9
6 Problems with hallucinations/delusions	0	1	2	3	4	9	12 Problems with occupations and activities	0	1	2	3	4	9

[†] Specify Disorder - Only the single most severe problem during the period is rated, using the following categories: A-Phobias, B-Anxiety/Panic, C-Obsessional/Compulsive problems, D-Reactions to severely stressful/traumatic events, E-Dissociative problems, F-Somatisation, G-Appetite (under/over eating), H-Sleep problems, I-Sexual problems, J-Problems not specified elsewhere (e.g., Mania).

LSI-R:SV

1	Two or more prior convictions	YES	NO	OMIT	Total Score Risk/Needs 6 - 8 High 3 - 5 Moderate 0 - 2 Low		
2	Arrested under the age of 16	YES	NO	OMIT			
3	Currently unemployed	YES	NO	OMIT			
4	Some criminal friends	YES	NO	OMIT			
5	Alcohol/ Drug problem: School/work	YES	NO	OMIT			
6	Psychological assessment indicated [‡]	YES	NO	OMIT			
7	Non-rewarding parental	0	1	2		3	OMIT
8	Attitudes/ Orientation: Supportive of crime	0	1	2		3	OMIT

© Multi-Health Systems, Inc. 1998.

[‡] Psychological assessment may be indicated due to: intellectual functioning; mood/affect difficulties; negative attitudes towards self; hostility, anger, aggression; poor impulse control, interpersonal skills or assertiveness; severe withdrawal or over-activity; delusions/hallucinations; disregards the feelings of others; lacks guilt/shame; commits bizarre criminal acts; appears irrational

Camberwell Assessment of Need - Forensic Short Version (CANFOR-S)

Rated By:

Date:

0 = No Problem 1 = Met Need 2 = Unmet Need 8 = Not Applicable ** 9 = Not Known

Circle who is interviewed (U = User, S = Staff)	User	Staff	Index Offence
1 Accommodation ** Do you have a place to live when you leave the hospital?			
2 Food Are you able to prepare your own meals and do your own shopping for food?			
3 Looking after the living environment Are you able to look after your room? Does anyone help you?			
4 Self-care Do you have any problems keeping yourself clean and tidy?			
5 Daytime Activities How do you spend your day? Do you have enough to do?			
6 Physical Health How well do you feel physically? What about side effects from medication?			
7 Psychotic symptoms Do you hear voices or have problems with your thoughts?			
8 Information about condition and treatment Have you been given clear information about current medication, treatment, rights			
9 Psychological distress Have you recently felt sad or low? Have you recently felt anxious or frightened?			
10 Safety to self Do you have thoughts of harming yourself? Do you put yourself in danger in any way			
11 Safety to others (excluding sexual offences / arson) Have you threatened other people or been violent? (e.g., lost your temper?)			
12 Alcohol Do you have a problem with alcohol?			
13 Drugs (including solvents) Do you have a problem with drugs?			
14 Company Are you happy with your social life? Do you wish you had more contact with others?			
15 Intimate relationships Do you have a partner? Do you have problems with your close relationships?			
16 Sexual expression How is your sex life? Are you experiencing difficulties with sexual matters?			
17 Child care ** Do you have any children under 18? Do you care for them? Do you have access?			
18 Basic education Do you have any difficulty in reading, writing, or understanding English?			
19 Telephone Do you know how to use a telephone? Is it easy to find one that you can use?			
20 Transport ** Do you have any problems using the bus, train or taxi? Do you get a free bus pass?			
21 Money Do you have any problems budgeting your money? Do you manage to pay your bills?			
22 Benefits Are you sure you are getting all of the benefits you are entitled to?			
23 Treatment Do you agree with the treatment (medical and/or psychological) prescribed?			
24 Sexual offences (where indicated) ** Do you think you might be at risk of committing a sexual offence?			
25 Arson (where indicated) ** Do you think you might be at risk of setting fires?			

A Met Needs	(count the number of 1s)			Total YES
B Unmet Needs	(count the number of 2s)			
C Total Needs	(add together A and B)			

APPENDIX I: COPYRIGHT NOTICE
(INTERNATIONAL JOURNAL OF FORENSIC MENTAL
HEALTH)

Permissions

T & F Reference Number: P012616-03

1/26/2016

Gregg Shinkfield
PhD candidate
Swinburne University
Melbourne, Australia
Gregg.Shinkfield@forensicare.vic.gov.au

Dear Mr. Shinkfield,

We are in receipt of your request to reproduce your article

Gregg Shinkfield & James Ogloff (2014)
A Review and Analysis of Routine Outcome Measures for Forensic Mental Health Services
International Journal of Forensic Mental Health 13 (3): 252-271.
DOI: 10.1080/14999013.2014.939788

For use in your dissertation

You retain the right as author to post your Accepted Manuscript on your departmental or personal website with the following acknowledgment: "This is an Accepted Manuscript of an article published in the *International Journal of Forensic Mental Health* online [August 30, 2014], available online: <http://www.tandfonline.com/doi/full/10.1080/14999013.2014.939788>

This permission is all for print and electronic editions.

For the posting of the full article it must be in a secure, password-protected intranet site only.

We will be pleased to grant you permission free of charge on the condition that:

This permission is for non-exclusive English world rights. This permission does not cover any third party copyrighted work which may appear in the material requested.

Full acknowledgment must be included showing article title, author, and full Journal title, reprinted by permission of the International Association of Forensic Mental Health Services (<http://www.iafmhs.org>).

Thank you very much for your interest in Taylor & Francis publications. Should you have any questions or require further assistance, please feel free to contact me directly.

Sincerely,

Mary Ann Muller
Permissions Coordinator
Telephone: 215.606.4334
E-mail: maryann.muller@taylorandfrancis.com

APPENDIX J: COPYRIGHT NOTICE
(INTERNATIONAL JOURNAL OF MENTAL HEALTH
NURSING)

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Jan 18, 2016

This Agreement between Gregg Shinkfield ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	3792201166917
License date	Jan 18, 2016
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	International Journal of Mental Health Nursing
Licensed Content Title	Use and interpretation of routine outcome measures in forensic mental health
Licensed Content Author	Gregg Shinkfield,James Ogloff
Licensed Content Date	Sep 6, 2014
Pages	8
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Measuring the progress and outcome of patients at the Thomas Embling hospital
Expected completion date	Jan 2017
Expected size (number of pages)	300
Requestor Location	Gregg Shinkfield Centre for Forensic Behavioural Science 505 Hoddle St Clifton Hill, Australia 3068 Attn: Gregg Shinkfield
Billing Type	Invoice
Billing Address	Gregg Shinkfield Centre for Forensic Behavioural Science 505 Hoddle St Clifton Hill, Australia 3068 Attn: Gregg Shinkfield
Total	0.00 AUD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction

(along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://myaccount.copyright.com>).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order**, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. **For STM Signatory Publishers clearing permission under the terms of the [STM Permissions Guidelines](#) only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.
- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.
- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.
- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.
- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.
- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.
- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The [Creative Commons Attribution License \(CC-BY\)](#) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The [Creative Commons Attribution Non-Commercial \(CC-BY-NC\) License](#) permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The [Creative Commons Attribution Non-Commercial-NoDerivs License](#) (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is

properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library

<http://olabout.wiley.com/WileyCDA/Section/id-410895.html>

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

APPENDIX K: COPYRIGHT NOTICE
(JOURNAL OF FORENSIC PSYCHIATRY & PSYCHOLOGY)



RightsLink®

[Home](#)[Create Account](#)[Help](#)

Title: Comparison of HoNOS and HoNOS-Secure in a forensic mental health hospital

Author: Gregg Shinkfield, James Ogloff

Publication: Journal of Forensic Psychiatry & Psychology

Publisher: Taylor & Francis

Date: Oct 17, 2016

Copyright © 2016 Routledge

[LOGIN](#)

If you're a [copyright.com user](#), you can login to RightsLink using your copyright.com credentials. Already a [RightsLink user](#) or want to [learn more?](#)

Thesis/Dissertation Reuse Request

Taylor & Francis is pleased to offer reuses of its content for a thesis or dissertation free of charge contingent on resubmission of permission request if work is published.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2016 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at customercare@copyright.com

APPENDIX L: COMPARISON OF HoNOS AND HoNOS-SECURE WORDING

In the following appendix, the changes that were made to the wording of HoNOS items when adapted to the HoNOS-Secure are highlighted. All alterations to the original text are highlighted in grey.

HoNOS	HoNOS- Secure
-------	---------------

1. Overactive, aggressive, disruptive or agitated behaviour Include such behaviour due to any cause, eg, drugs, alcohol, dementia, psychosis, depression, etc. Do not include bizarre behaviour, rated at Scale 6. 0 No problems of this kind during the period rated. 1 Irritability, quarrels, restlessness etc. Not requiring action. 2 Includes aggressive gestures, pushing or pestering others; threats or verbal aggression; lesser damage to property (eg, broken cup or window); marked over-activity or agitation. 3 Physically aggressive to others or animals (short of rating 4); threatening manner; more serious over-activity or destruction of property. 4 At least one serious physical attack on others or on animals; destruction of property (e.g., fire-setting); serious intimidation or obscene behaviour.	1. Overactive, aggressive, disruptive or agitated behaviour <u>Include</u> behaviour due to any cause (drugs/alcohol/dementia/psychosis/depression), etc. <u>Do not</u> include bizarre behaviour, rated at Scale 6. Rate sexual behaviours at Scale 8 (I), but rate any violence/intimidation here. 0. No problems of this kind during the period rated. 1. Some irritability, quarrels, restlessness, <u>disruptive behaviour</u>, etc. 2. Includes <u>occasional</u> aggressive gestures, pushing, pestering or <u>provoking</u> others; threats or verbal aggression; lesser damage to property (e.g., broken cup or window, <u>cigarette burns</u>); marked over-activity or agitation. 3. Physically aggressive to others or animals (short of rating 4); <u>persistently</u> threatening manner; more serious over-activity or destruction of property (e.g., broken doors, <u>minor fire setting to ashtrays</u>, etc). 4. At least one serious physical attack on others or on animals; destructive of property (e.g., <u>dangerous</u> fire setting); <u>use of weapons</u>; <u>persistent</u> serious intimidation behaviour.
--	--

2 Non-accidental self-injury Do not include accidental self-injury (due eg, to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and the injury at Scale 5. Do not include illness or injury as a direct consequence of drug or alcohol use rated at Scale 3, (eg, cirrhosis of the liver or injury resulting from drunk driving are rated at Scale 5). 0 No problem of this kind during the period rated. 1 Fleeting thoughts about ending it all, but little risk during the period rated; no self-harm. 2 Mild risk during period; includes non-hazardous self-harm eg, wrist- scratching. 3 Moderate to serious risk of deliberate self-harm during the period rated; includes preparatory acts eg, collecting tablets. 4 Serious suicidal attempt or serious deliberate self-injury during the period rated.	2. Non-accidental self-injury <u>Do not</u> include <u>accidental</u> self-injury (due to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and injury at Scale 5. <u>Do not</u> include illness/injury as a direct consequence of drug/alcohol use rated at Scale 3 (e.g., cirrhosis of liver or injury resulting from drunk driving are rated at Scale 5). 0. No problem of this kind during the period rated. 1. Fleeting thoughts about <u>self-harm or suicide</u>, but little risk; no self-harm. 2. Mild risk during period; includes non-hazardous self-harm (e.g., wrist scratching, <u>not requiring physical treatment</u>); <u>persistent or worrying thoughts about self-harm</u>. 3. Moderate to serious risk of deliberate self-harm; includes preparatory acts (e.g., collecting tablets, <u>secreting razor blade, making nooses, suicide notes</u>). 4. Serious suicidal attempt and/or serious deliberate self-harm during period (i.e., person seriously <u>harmed self, or intended to, or risk death by their actions</u>).
--	--

3 Problem drinking or drug-taking Do not include aggressive or destructive behaviour due to alcohol or drug use, rated at Scale 1. Do not include physical illness or disability due to alcohol or drug use, rated at Scale 5. 0 No problem of this kind during the period rated. 1 Some over-indulgence, but within social norm. 2 Loss of control of drinking or drug-taking; but not seriously addicted. 3 Marked craving or dependence on alcohol or drugs with frequent loss of control, risk taking under the influence, etc. 4 Incapacitated by alcohol or drug problems.	3. Problem drinking or drug taking <u>Do not</u> include aggressive/destructive behaviour due to alcohol/drug use, rated at Scale 1. <u>Do not</u> include physical illness/disability due to alcohol or drug use, rated at Scale 5. 0. No problem of this kind during the period rated (e.g., minimal cannabis use, drinking within health guidelines). 1. Some over-indulgence but within the social norm (e.g., significant cannabis use, other low risk activity). 2. Loss of control of drinking or drug taking, but not seriously addicted (e.g., regular cannabis use, drinking above health guidelines); (in controlled settings - occasional positive urine tests, loss of leave or delayed discharge on account of attitude or behaviour towards drink and drugs). 3. Marked dependence on alcohol or drugs with frequent loss of control, drunk driving; (in controlled settings - drug debts, frequent attempts to obtain drugs; persistent pre-occupation with drink/drugs; repeated intoxication or positive urine tests). 4. Incapacitated by alcohol/drug problems.
---	--

4 Cognitive problems Include problems of memory, orientation and understanding associated with any disorder: learning disability, dementia, schizophrenia, etc. Do not include temporary problems (eg, hangovers) resulting from drug or alcohol use, rated at Scale 3. 0 No problem of this kind during the period rated. 1 Minor problems with memory or understanding eg, forgets names occasionally. 2 Mild but definite problems, eg, has lost way in a familiar place or failed to recognise a familiar person; sometimes mixed up about simple decisions. 3 Marked disorientation in time, place or person, bewildered by everyday events; speech is sometimes incoherent, mental slowing. 4 Severe disorientation, eg, unable to recognise relatives, at risk of accidents, speech incomprehensible, clouding or stupor.	4. Cognitive problems <u>Include</u> problems of memory, orientation and understanding associated with any disorder: learning disability, dementia, schizophrenia, etc. <u>Do not</u> include temporary problems (e.g., hangovers) resulting from drug/alcohol use, rated at Scale 3. 0. No problem of this kind during the period rated. 1. Minor problems with memory and understanding (e.g., forgets names occasionally). 2. Mild but definite problems (e.g., has lost the way in a familiar place or failed to recognise a familiar person); sometimes mixed up about simple decisions; major impairment of long term memory. 3. Marked disorientation in time, place or person; bewildered by everyday events; speech is sometimes incoherent; mental slowing. 4. Severe disorientation (e.g., unable to recognise relatives, at risk of accidents, speech incomprehensible); clouding or stupor.
---	---

5 Physical illness or disability problems Include illness or disability from any cause that limits or prevents movement, or impairs sight or hearing, or otherwise interferes with personal functioning. Include side-effects from medication; effects of drug/alcohol use; physical disabilities resulting from accidents or self-harm associated with cognitive problems, drunk driving etc. Do not include mental or behavioural problems rated at Scale 4. 0 No physical health problem during the period rated. 1 Minor health problem during the period (eg, cold, non-serious fall, etc). 2 Physical health problem imposes mild restriction on mobility and activity. 3 Moderate degree of restriction on activity due to physical health problem. 4 Severe or complete incapacity due to physical health problem.	5. Physical illness or disability problems <u>Include</u> illness or disability from any cause that limits or prevents movement, or impairs sight or hearing, or otherwise interferes with personal functioning (e.g., pain). <u>Include</u> side effects from medication; effects of drug/alcohol use; physical disabilities resulting from accidents or self-injury associated with cognitive problems, drink driving, etc. <u>Do not</u> include mental or behavioural problems rated at Scale 4. 0. No physical health problem during the period rated. 1. Minor health problem during the period rated (e.g., cold, non-serious fall). 2. Physical health problem imposes mild restriction on mobility and activity (e.g., sprained ankle, breathlessness). 3. Moderate degree of restriction on activity due to physical health problem (e.g., has to give up work or leisure activities). 4. Severe or complete incapacity due to physical health problems.
--	--

6 Problems associated with hallucinations and delusions Include hallucinations and delusions irrespective of diagnosis. Include odd and bizarre behaviour associated with hallucinations or delusions. Do not include aggressive, destructive or overactive behaviours attributed to hallucinations or delusions, rated at Scale 1. 0 No evidence of hallucinations or delusions during the period rated. 1 Somewhat odd or eccentric beliefs not in keeping with cultural norms. 2 Delusions or hallucinations (eg, voices, visions) are present, but there is little distress to patient or manifestation in bizarre behaviour, that is, moderately severe clinical problem. 3 Marked preoccupation with delusions or hallucinations, causing much distress and/or manifested in obviously bizarre behaviour, that is, moderately severe clinical problem. 4 Mental state and behaviour is seriously and adversely affected by delusions or hallucinations, with severe impact on patient.	6. Problems associated with hallucinations and delusions <u>Include</u> hallucinations and delusions irrespective of diagnosis. <u>Include</u> odd and bizarre behaviour associated with hallucinations or delusions, such as thought disorder. <u>Do not</u> include aggressive, destructive or overactive behaviours attributed to hallucinations or delusions, rated at Scale 1. 0. No evidence of hallucinations/delusions during period rated. 1. Somewhat odd or eccentric beliefs not in keeping with cultural norms. 2. Delusions or hallucinations (e.g., voices, visions) present, but little distress to patient or manifestation in bizarre behaviour (i.e., clinically present but mild). 3. Marked preoccupation with delusions or hallucinations, causing much distress and/or manifested in obviously bizarre behaviour (i.e., moderately severe clinical problem). 4. Mental state and behaviour is seriously and adversely affected by delusions or hallucinations, with severe impact on patient/others.
---	---

7 Problems with depressed mood Do not include over-activity or agitation, rated at Scale 1. Do not include suicidal ideation or attempts, rated at Scale 2. Do not include delusions or hallucinations, rated at Scale 6. 0 No problems associated with depressed mood during the period rated. 1 Gloomy; or minor changes in mood. 2 Mild but definite depression and distress: e.g., feelings of guilt; loss of self-esteem. 3 Depression with inappropriate self-blame, preoccupied with feelings of guilt. 4 Severe or very severe depression, with guilt or self-accusation.	7. Problems with depressed mood <u>Do not</u> include over-activity or agitation, rated at Scale 1. <u>Do not</u> include suicidal ideation or attempts, rated at Scale 2. <u>Do not</u> include delusions or hallucinations, rated at Scale 6. 0. No problems associated with depressed mood during period rated. 1. Gloomy or minor changes in mood (not regarded as “depression”). 2. Mild but definite depression and distress (e.g., feelings of guilt; loss of self-esteem, but not amounting to a clinical episode of depression); troublesome mood swings. 3. Depression with inappropriate self-blame, preoccupied with feelings of guilt, at a level likely to attract diagnosis and treatment; clinically problematic swings of mood. 4. Severe or very severe depression, with guilt or self-accusation.
--	---

8 Other mental and behavioural problems Rate only the most severe clinical problem not considered at items 6 and 7 as follows: specify the type of problem by entering the appropriate letter: A phobic; B anxiety; C obsessive-compulsive; D stress; E dissociative; F somatoform; G eating; H sleep; I sexual; J other, specify. 0 No evidence of any of these problems during period rated. 1 Minor non-clinical problems. 2 A problem is clinically present at a mild level, eg, patient/client has a degree of control. 3 Occasional severe attack or distress, with loss of control eg, has to avoid anxiety provoking situations altogether, call in a neighbour to help, etc., that is, a moderately severe level of problem. 4 Severe problem dominates most activities.	8. Other mental and behavioural problems <u>Rate</u> only the most severe clinical problem <u>not</u> considered at items 6 and 7. Specify type of problem by entering the appropriate letter: <u>A</u> phobic; <u>B</u> anxiety; <u>C</u> obsessive compulsive; <u>D</u> stress; <u>E</u> dissociative; <u>F</u> somatoform; <u>G</u> eating; <u>H</u> sleep; <u>I</u> sexual (for sexual behaviour problem, see guidance in brackets); <u>J</u> other, specify. 0. No evidence of any of these problems during period rated. 1. Minor non-clinical problems; (impolite sexual talk/gestures). 2. A problem is clinically present, but there are relatively symptom-free intervals and patient/client has degree of control, i.e., mild level; (excessively tactile or non-contact sexual offence or very provocative, e.g., exposes self; walks around semi-naked, peeping into bedrooms, etc.). 3. Constant preoccupation with problem; occasional severe attack or distress, with loss of control, e.g., avoids anxiety provoking situations, calls neighbour to help, etc.; moderately severe level of problem; (sexual assault, e.g., touching breast/buttock/genitals over clothing). 4. Severe, persistent problem dominates most activities; (more serious sexual assault, i.e., genital contact, sexual touching under clothing).
--	--

9 Problems with relationships Rate the patient's most severe problem associated with active or passive withdrawal from social relationships, and/or non-supportive, destructive or self-damaging relationships. 0 No significant problems during the period. 1 Minor non-clinical problems. 2 Definite problems in making or sustaining supportive relationships: patient complains and/or problems are evident to others. 3 Persisting major problems due to active or passive withdrawal from social relationships, and/or to relationships that provide little or no comfort or support. 4 Severe and distressing social isolation due to inability to communicate socially and/or withdrawal from social relationships.	9. Problems with relationships <u>Rate</u> the patient's most severe problem associated with active or passive withdrawal from social relationships, and/or non-supportive, destructive or self-damaging relationships. <u>Take into account limited access to outside relationships in secure settings, include patients/ inmates/staff relationships.</u> 0. No significant problems during the period. 1. Minor non-clinical problem. 2. Definite problems in making or sustaining supportive relationships; patient complains and/or problems are evident to others. 3. Persisting major problems due to active or passive withdrawal from social relationships, and/or to relationships that provide little or no comfort or support. 4. Severe and distressing social isolation due to inability to communicate socially and/or withdrawal from social relationships.
--	--

10 Problems with activities of daily living Rate the overall level of functioning in activities of daily living (ADL): eg, problems with basic activities of self-care such as eating, washing, dressing, toilet; also complex skills such as budgeting, organising where to live, occupation and recreation, mobility and use of transport, shopping, self-development, etc. Include any lack of motivation for using self-help opportunities, since this contributes to a lower overall level of functioning. Do not include lack of opportunities for exercising intact abilities and skills, rated at Scale 11 and Scale 12. 0 No problems during period rated; good ability to function in all areas. 1 Minor problems only eg, untidy, disorganised. 2 Self-care adequate, but major lack of performance of one or more complex skills (see above). 3 Major problems in one or more areas of self-care (eating, washing, dressing, toilet) as well as major inability to perform several complex skills. 4 Severe disability or incapacity in all or nearly all areas of self-care and complex skills.	10. Problems with activities of daily living <u>Rate</u> the overall level of functioning in activities of daily living (ADL) (e.g., problems with basic <u>activities of self-care</u>; eating, washing, toilet), also <u>complex skills</u>; budgeting, organising where to live, recreation, mobility, use of transport, self-development, etc. <u>Include</u> any lack of motivation for using self-help opportunities, as this contributes to a lower overall level of functioning. Do <u>not</u> include lack of opportunities for exercising intact abilities and skills (<u>e.g., in secure settings</u>), rated at levels 11 and 12. 0. No problems during period rated; good ability to function in all areas. 1. Minor problems only (e.g., untidy, disorganised). 2. Self-care adequate, but major lack of performance of one or more complex skills (see above); <u>needs occasional prompting</u>. 3. Major problems in one or more area of self-care (eating, washing, dressing, toilet, etc.); has a major inability to perform several complex skills; <u>needs constant prompting or supervision</u>. 4. Severe disability/incapacity in all or nearly all areas of self-care and complex skills.
---	--

11 Problems with living conditions

Rate the overall severity of problems with the quality of living conditions and daily domestic routine. Are the basic necessities met (heat, light, hygiene)? If so, is there help to cope with disabilities and a choice of opportunities to use skills and develop new ones? Do not rate the level of functional disability itself, rated at Scale 10.

NB: Rate patient's usual accommodation. If in acute ward, rate the home accommodation. If information not obtainable, rate 9.

- 0 Accommodation and living conditions are acceptable; helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help.
- 1 Accommodation is reasonably acceptable although there are minor or transient problems (eg, not ideal location, not preferred option, doesn't like food, etc).
- 2 Significant problems with one or more aspects of the accommodation and/or regime (eg, restricted choice; staff or household have little understanding of how to limit disability, or how to help develop new or intact skills).
- 3 Distressing multiple problems with accommodation (eg, some basic necessities absent); housing environment has minimal or no facilities to improve patient's independence.
- 4 Accommodation is unacceptable (eg, lack of basic necessities, patient is at risk of eviction, or 'roofless', or living conditions are otherwise intolerable making patient's problems worse).

11. Problems with living conditions

Rate overall severity of problems with quality of living conditions and daily domestic routine. Are basic necessities met (heat, light, hygiene)? If so, is there help to cope with disabilities and a choice of opportunities to use skills and develop new ones? Do not rate the level of functional disability itself, rated at Scale 10.

N.B. Rate patient's usual accommodation whether community, open or secure setting (hospital or prison). If in acute ward/other temporary care, rate home accommodation.

0. Accommodation and living conditions acceptable; help to keep disability at Scale 10 to lowest level possible, supportive of self-help.
1. Accommodation reasonably acceptable although there are minor or transient problems (e.g., not ideal location, not preferred option, doesn't like the food, etc.).
2. Significant problems with one or more aspects of the accommodation/regime (e.g., restricted choice; **inflexible programme**; staff or household have little understanding of how to **limit disability**, or how to help use or develop new or intact skills).
3. Distressing multiple problems with accommodation/**regime** (e.g., some basic necessities absent, environment has minimal/no facilities to improve patient's independence); **unnecessarily restrictive physical security** (e.g., no access to outdoors, awaiting transfer to less secure facilities).
4. **Environment** unacceptable (e.g., lack of basic necessities or patient at risk of eviction/**arbitrary transfer**); **'roofless' or highly restrictive living conditions** otherwise intolerable making patient's problems worse; severe physical confinement (e.g., much of daytime locked in room/cell, confined **unnecessarily in seclusion or unfurnished room**)

12 Problems with occupation and activities

Rate the overall level of problems with quality of day-time environment. Is there help to cope with disabilities, and opportunities for maintaining or improving occupational and recreational skills and activities? Consider factors such as stigma, lack of qualified staff, access to supportive facilities, eg, staffing and equipment of day centres, workshops, social clubs, etc. Do not rate the level of functional disability itself, rated at Scale 10. NB: Rate the patient's usual situation. If in acute ward, rate activities during period before admission. If information not available, rate 9.

12. Problems with occupation and activities

Rate overall level of problems with quality of day-time environment. Is there help to cope with disabilities, opportunities for maintaining or improving occupational and recreational skills and activities? Consider stigma, lack of appropriate Qualified Staff, access to supportive facilities (e.g., staffing/equipment at Day Centres, workshops, social clubs). Do not rate level of functional disability itself, rated at Scale 10.

N.B. Rate patient's usual situation, whether in community, open or secure setting (hospital or prison). If in an acute ward/temporary care, rate activities during period before admission.

0 Patient's day-time environment is acceptable; helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help. 1 Minor or temporary problems, eg, late pension cheques, reasonable facilities available but not always at desired times etc. 2 Limited choice of activities, eg, there is a lack of reasonable tolerance (eg, unfairly refused entry to public library or baths etc.); or handicapped by lack of a permanent address; or insufficient carer or professional support; or helpful day setting available but for very limited hours. 3 Marked deficiency in skilled services available to help minimise level of existing disability; no opportunities to use intact skills or add new ones; unskilled care difficult to access. 4 Lack of any opportunity for daytime activities makes patient's problem worse.	0. Patient's day time environment acceptable; helps to keep disability rated at Scale 10 to lowest level possible; supportive of self-help. 1. Minor or temporary problems (e.g., late giro cheques; reasonable facilities available but not always at desired and appropriate times, etc.). 2. Limited choice of activities; lack of reasonable tolerance (e.g., unfairly refused entry to public library/baths; lack of day areas); lack of facilities in large establishment; handicapped by lack of permanent address; insufficient carer/professional support; or helpful day setting available but for very limited hours. 3. Marked deficiency in skilled services available to help minimise level of existing disability; no opportunities to use intact skills or develop new ones; unskilled care difficult to access; no activity areas available; leave withheld from small establishment causes restriction. 4. Lack of opportunity for daytime activities makes problem worse; long periods of enforced inactivity each day (e.g., prison cell).
--	--