Acute inpatient mental health wards and inpatient alternatives: a quantitative comparison of the care provided

Brynmor Lloyd-Evans

UCL

PhD Thesis

I, Brynmor Lloyd-Evans, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Abstract

Background

Literature indicates widespread dissatisfaction with UK acute psychiatric wards. Patients report boredom and insufficient time with staff. Residential alternatives to acute wards have been developed.

Aims

1) To review literature for the effectiveness and acceptability of alternatives
2) To identify or develop measures of content of care for acute inpatient and residential crisis services
3) To compare the content of care at alternatives and standard services and understand its relationship to patient satisfaction. Hypotheses tested were that alternatives provide greater total care, more social and psychological interventions but fewer physical and pharmacological interventions than standard wards.

Method

A systematic review of studies evaluating alternatives was conducted. Measures of content of care were reviewed. New measures were developed (CaSPAR, CaRICE and CCCQ-P) and their psychometrics explored.

Data were collected from 4 alternatives and 4 standard services using CaSPAR (n=224), CaRICE (1 recording week per service), CCCQ-P and CSQ (n=314). The relationship of service type, patient characteristics and CCCQ-P scores to patient satisfaction was explored.

Results

The limited current evidence does not contra-indicate alternatives and suggests patient satisfaction may be greater at community alternatives than standard wards.

No study hypotheses were corroborated. Sub-group analysis indicated community alternatives provided more psychological and less physical and pharmacological care than standard wards. All CCCQ-P variables were

significantly associated with patient satisfaction. Patient satisfaction was greatest at community alternatives. It remained significantly greater at alternatives than standard wards after adjusting for CCCQ-P variables.

Discussion

Community alternatives are a promising service model. Their greater acceptability than standard wards was not explained by measured differences in care. Intensity of care may influence patient satisfaction more than the types of intervention provided. Increasing staff-patient contact should be an aim for alternative and standard services. There may be differing, valid perspectives about what constitutes care: multi-methods assessment is required.

for Cindy

with love and thanks

Acknowledgements

Sonia Johnson from the Department of Mental Health Sciences, UCL and Mike Slade from the Health Services Population and Research Department at the Institute of Psychiatry supervised this thesis. I am extremely grateful for their frequent, helpful supervision.

Thanks also to Morven Leese from the Institute of Psychiatry for additional statistical advice and Paul Bebbington from UCL for commenting on a completed draft of this thesis. I am grateful to Martin Orrell from UCL and Tim Weaver from Imperial College London, who examined this thesis and whose comments have informed its final version.

I am very grateful for the help of staff and service users from services where data were collected for the quantitative investigation presented in Chapters 6 – 9 of this thesis. The quantitative investigation was undertaken as part of The Alternatives Study, a national research study funded by the National Institute for Health Research Service Delivery and Organisation Programme. The views expressed in this thesis are the author’s own and not necessarily those of the NHS, the NIHR or the Department of Health. The relationship between this thesis and The Alternatives Study is clarified in Section 1.5. Except where stated, this thesis represents the author’s own work.

Contents

List of Tables

Table **Page**

List of figures

Abbreviations used in this thesis

ACT	Assertive Community Treatment
ASI	Addiction Severity Index
BASIS-32	Behaviour and Symptom Identification Scale
BPRS	Brief Psychiatric Rating Scale
CAN	Camberwell Assessment of Need
CaRICE	Camden Record of Inpatient Care Events
CaSPAR	Camden Staff-Patient Activity Record
CCCQ-P	Camden Content of Care Questionnaire (patient version)
CCCQ-S	Camden Content of Care Questionnaire (staff version)
CENTRAL	Cochrane Library Central Register of Controlled Trials
CI	Confidence Intervals
CMHC	Community Mental Health Centre
CMHT	Community Mental Health Team
CRT	Crisis Resolution Team
CSIP	Care Services Improvement Partnership
CSQ	Client Satisfaction Questionnaire
CSRI	Client Service Receipt Inventory
DACTS	Dartmouth Assertive Community Treatment Scale
DARE	Cochrane Library Database of Abstracts of Reviews of Effects
EBP	Evidence Based Practices Program
ESMS	European Service Mapping Schedule
GAF	Global Assessment of Functioning
GAS	Global Assessment Scale
HoNOS	Health of the Nation Outcome Scale
HRSD	Hamilton Rating Scale for Depression
ICMHC	International Classification of Mental Health Care
IFACT	Index of Fidelity to Assertive Community Treatment
IMPS	Inpatient Multi-Dimensional Scale for Rating Psychotic Patients
ITT	intention-to-treat
LQLP	Lancashire Quality of Life Profile

MBU	Mother and Baby Unit
MeSH	Medical sub-heading
MMPI	Minnesota Multiphasic Personality Inventory
MSER	Mental State Examination Recoord
OR	odds ratio
PANSS	Positive and Negative Symptoms Scale
PBAS	Patients’ Behaviour Assessment Scale
PEF	Psychiatric Evaluation Form
POC	Perceptions of Care Questionnaire
PSE	Present State Examination
PSS	Psychiatric Status Schedule
QUIS	Quality of Interactions Schedule
QUOROM	Quality of Reporting of Meta-analyses Guidelines
RCT	randomised controlled trial
SBS	Social Behaviour Schedule
SD	standard deviation
SF-36	Health Survey – Short Form
SF-36(V)	Health Survey – Short Form (Veterans Version)
SMD	standardised mean difference
TES	Treatment Effectiveness Scale
VSSS	Verona Service Satisfaction Scale
WMD	weighted mean difference

Glossary

Term	*Definition*
Accredited accommodation	Placement of a patient with a private family (supported by mental health services) during a period of mental health crisis
Alternatives	Residential services for adults in acute mental health crisis which differ from standard wards by being any of: based in the community; implementing time-limited admission; using a distinctive therapeutic model; dedicated to a specific clinical group; dedicated to a specific socio-demographic group
The Alternatives Study	UK study involving a national survey of alternatives and comparison of representative alternatives and standard services
CaSPAR	Camden Staff-Patient Activity Record: a measure of intensity of care at services developed and used in this thesis
CaRICE	Camden Record of Inpatient Care Events: a measure of the intensity and nature of care at services developed and used in this thesis
CCCQ-P	Camden Content of Care Questionnaire (Patient version): a patient-report measure of the intensity and nature of care received by individual patients during admission, developed and used in this thesis

Term	*Definition*
CCCQ-S	Camden Content of Care Questionnaire (Staff version): a staff-report measure of the intensity and nature of care received by individual patients during admission, developed for this thesis
Clinical crisis house	A type of community alternative which is managed by the statutory sector and shares many characteristics with acute wards
Community alternatives	Alternatives based in non-hospital settings
Content of care	The intensity and nature of interventions delivered to patients at a service
Crisis house	Small community alternatives, typically based in a house in a residential street.
Crisis Team Beds	A type of community alternative characterised by having few beds, being short-stay and closely integrated with Crisis and Home Treatment Teams
Evidence Based Practices (EBP)	A programme run by the United States Department of Health and Human Services providing a process for defining then evaluating models of mental health services and practice, aimed at establishing effective practice
Event recording	Content of care measurement recording pre-defined events of interest at or near the time they occur

Term	*Definition*
Family sponsor homes	Placement of a patient with a private family (supported by mental health services) during a period of mental health crisis
General care organisation	A subscale used in CaRICE and CCCQ-P relating to assessment, care planning and care coordination
Hawthorne effect	The process where the conduct of research affects what is being researched
Incident recording	Content of care measurement recording pre-defined events of interest at the time they occur or retrospectively
Inpatient alternatives	Alternatives based in hospital settings
Intensity of care	How much care is provided at services, comprising duration and/or frequency of interventions: a domain of content of care
Mental Health Act status	Whether a patient is compulsorily detained under the Mental Health Act or not
Nature of care	The types of intervention provided to patients: a domain of the content of care
Non-clinical crisis house	A type of community alternative less integrated with statutory services and with fewer qualified staff than other types of alternative

Term	*Definition*
Physical and pharmacological interventions	A subscale used in CaRICE and CCCQ-P relating to observations, restraint, physical healthcare and help with medication
Process	What is done for patients at a service: the process of care can be distinguished from input factors (characteristics of patients or the settings of services) and patient outcomes
Programme implementation	Evaluation of the extent to which a service is meeting specific criteria or agreed objectives, relying on process of care measurement
Psychological interventions	A subscale used in CaRICE and CCCQ-P relating to help with past events, inter-personal problems and symptom coping strategies.
Qualified staff	Staff with a mental health professional qualification and registration with a mental health professional body, including psychiatrists, psychologists, social workers, nurses and occupational therapists
Reactivity	Synonymous with the Hawthorne effect
Refocusing model	A model of care developed for acute wards, characterised by minimal use of formal observations, constituting an alternative to standard care

Term	*Definition*
Social interventions	A subscale used in CaRICE and CCCQ-P relating to help with practical and occupational concerns and family support.
Soteria	A model of community alternative developed in California in the 1970s for people with first or second episode psychosis: care characterised by minimal use of medication
Standard services	General adult acute inpatient mental health wards providing standard care
Style of care	A process element distinct from content of care, relating to the atmosphere at services and style in which care is provided
Tidal Model	A model of care developed for acute wards, characterised by daily care planning and valuing patient experience of crisis, constituting an alternative to standard care
Time recording	Content of care measurement recording any activity in pre-defined recording periods
Unqualified staff	Staff without a mental health professional qualification, such as Healthcare Assistants and social care staff.

Chapter 1: Introduction

This thesis compares the care provided in acute inpatient mental health services with the care provided in residential services offering an alternative to admission to a standard acute ward. In this chapter, the development of acute inpatient care is described. The aims of the thesis are presented and this study is set in the broader context of mental health services research.

1.1 The history of mental health inpatient care

In the western, industrialised world, mental health inpatient services – large asylums - were first established at the beginning of the nineteenth century and grew in number until the mid twentieth century. Fakhoury and Priebe (2007) describe reasons why societies invested in asylums:

- A developing social welfare movement and assumption of the state's responsibility to care for vulnerable people
- Reduced ability of families to support a family member with mental health problems because of family fragmentation and reduced means resulting from urbanisation
- Increased visibility and intolerance of social deviance in densely populated urban environments

Bed numbers in inpatient units reached a peak in the UK in the 1950's, when 154,000 inpatient beds were available (Davidge 1993). Since the 1950's, a process of deinstitutionalisation has seen the closure of hospitals and a reduction in inpatient bed numbers across the industrialised world (Fakhoury and Priebe 2000), including the UK. In 2008/9, UK government statistics report that there were just under 17,500 psychiatric inpatient beds for adults age 16-65 in England, including 11,200 acute beds (Department of Health 2009). Drivers for deinstitutionalisation included:

- A general social movement emphasising the community as a positive helping resource (Hawks 1975)
- An increasing importance placed on patients' liberties and quality of life (Peele and Chodoff 1999)

- Increased public awareness of unacceptable standards of care in inpatient institutions (Martin 1984)
- The development of better drug treatments and psychosocial interventions to assist management of people with mental health problems in the community (Szmukler and Holloway 2001)
- Research evidence of the harmful effects of understimulating environments in large mental hospitals (Wing and Brown 1970) and the possibility of successfully resettling long-stay patients in the community (Leff et al. 1994)

The closure of psychiatric hospitals did not always lead to patients living independently with support from community mental health services. Priebe and colleagues (2005) describe a process of transinstitutionalisation in Europe: during the 1990s, in several countries including the UK, a reduction in psychiatric inpatient beds was mirrored by an increase in other forms of institutional care, including supported housing, forensic beds and prison numbers. In this decade however, there is some evidence that developments in community mental health care have genuinely reduced admissions to residential services for people with mental illness: an audit of initial stages of implementation in England of Crisis Resolution Teams, dedicated to providing short-term, intensive home treatment to avert hospital admissions, found their implementation was associated with a mean 10% reduction in inpatient admissions (Glover et al. 2006).

The reduction of inpatient beds over the last sixty years has seen a corresponding change in the role of inpatient care, from the default care provision for people with acute mental illness to the treatment of last resort, used only in circumstances and for periods of time when support at home or in less restrictive residential settings cannot be managed (Department of Health 2005). Typical length of admission has decreased correspondingly (Szmukler and Holloway 2001), with lengths of stay in European acute inpatient services now lasting on average 1-5 weeks rather than months or years (McCrone and Lorusso 1999). Reductions in bed numbers and reduced length of stay have also influenced the profile of patients admitted to acute wards. Numbers of readmissions and the proportion of inpatients with previous experience of hospital care have risen: UK Census data indicates that 30% of inpatients in 2007 were also inpatients in 2006, and 20% of them had also been in hospital at the time of the 2005 census (Healthcare Commission 2007). As thresholds for admission have risen with the need for all less restrictive options to be exhausted, inpatients have been characterised by higher levels of morbidity (Patrick 1989) and the proportion of patients who are admitted compulsorily has risen (Wall et al. 1999).

Inpatient services’ change in role has been accompanied by change in the nature of care provision in acute inpatient wards. Traditional psychiatric hospitals established before the advent of deinstitutionalisation offered care in large asylums, typically situated outside main towns and cities, often with spacious living areas and grounds (Fakhoury and Priebe 2007). Patients’ needs for food, shelter, clothing, occupation, leisure activity, social interaction and a minimal income (Thornicroft and Bebbington 1989) were addressed entirely within the institution during long admissions. Goffman (1961) characterised mental health hospitals as “total institutions”: closed communities with their own hierarchies and customs, which were difficult to leave and isolated patients completely from the outside world. Quirk and colleagues (2006) contrast this with the “permeable institution” found in modern UK inpatient wards. Patients typically have periods of leave from hospital during a stay; community professionals and family and friends visit; contact with the outside world is maintained through media including television and mobile phones; institutional identities are blurred to the point where staff and patients on wards are not easily distinguished by visitors or new patients (Quirk et al. 2006).

1.2 A model of inpatient acute care

What is the function of acute inpatient care in the current UK acute care system? As described above, the purpose of acute admission is defined negatively as the care to be provided when less restrictive alternatives are not viable (Department of Health 2005). Minimising the duration of admission is one aim of inpatient care, with planning from an early stage of admission to facilitate prompt discharge recommended (Ramsay and Holloway 1998, Department of Health 2005). Flannigan and colleagues (1994) identified two main types of reason for admission to acute wards in inner London: challenging behaviours (harm to self or others) and social/preventive factors (prevention of psychosocial distress, removal from stressful environment or relief for carers). Ramsay and Holloway (1998) propose elements of effective management of an inpatient admission should include identification of the reason for admission, thorough assessment of a patient's problems and development of treatment aims and objectives. However, researchers have concluded that surprising little is known about the modes of operation in acute inpatient wards (Muijen 1999, Quirk and Lelliot 2001, Department of Health 2005) and the purpose of acute admission has also been insufficiently defined (DoH 2005, Bowers et al. 2009). Bowers (2005) notes the ideological confusion about the nature and purpose of inpatient psychiatry and concludes that *"over the last few decades, acute inpatient psychiatric care has rather lost its way" (Bowers 2005 p.231).*

In an attempt to address these problems, Bowers and colleagues (2009) developed a conceptual model of the aims and functions of acute inpatient psychiatry, drawing on a literature review of reasons for admission to acute inpatient services (Bowers 2005) and qualitative research with key inpatient staff (Bowers et al. 2005). This is provided in Figure 1.1. It describes the factors determining who gets admitted to acute inpatient services, the function of admission and the modes of operation in inpatient care.

Figure 1.1: A model of the aims and functions of acute inpatient psychiatry (Bowers et al 2009)

The model of the aims and functions of acute inpatient psychiatry by Bowers and colleagues (2009) is third party copyright material, for which permission for reproduction in a publicly accessible electronic copy of this thesis has not been obtained. The model is visually represented in:

Bowers,L.; Chaplin,R.; Quirk,A.; Lelliott,P. (2009) “A conceptual model of the aims and functions of acute inpatient psychiatry” *Journal of Mental Health* vol. 18(4) pp 316-325

Factors determining admission: Figure 1.1 shows four types of criterion influencing a decision to admit someone to hospital. First, the person must have a mental disorder and be experiencing acute illness. This *admission illness* in itself is insufficient reason to admit. Second, an additional *admission problem* must be identified. This will commonly concern risk, treatment refusal or failure of self care, but may also involve socially intolerable behaviour or uncertainty about the nature of a person's problems requiring careful assessment. Third, environmental and resource factors create an *admission filter*, influencing decision about whether admission is possible, desirable or necessary. These include consideration of a person's own support network and the availability of appropriate community services. Patient preference may also influence admission decisions where not otherwise clear cut. Finally, the attitudes and customs of the service or individual clinician making the *admission decision* add an idiosyncratic element to who gets admitted and for what reason.

The function of admission: The Bowers model identifies four purposes of admission. The *primary admission task* is to address the admission illness and problems which led to admission. This may involve any of the following tasks: psychiatric treatment, safety, assessment, basic care, rehabilitation, high tolerance accommodation, the resolution of personal or social stress (Bowers et al. 2009). Help with difficulties associated with the admission problem but which would not on their own have resulted in admission is described as an *admission bonus.* Examples of this include providing clean clothes or a bath for patients who have self-neglected or finding somewhere more appropriate to live for someone whose accommodation had broken down during their crisis. Admission may provide an opportunity to make incremental improvements in other long-standing problems not closely related to the person's admission, such as deficits in social network or activity: these are described as *secondary admission tasks.* Finally, a task of admission is to *prevent iatrogenesis*, such as institutionalisation, loneliness through disruption of social networks, intense contact with other people exacerbating acute psychosis.

Modes of operation: Five elements of inpatient services' operation are distinguished. *Management* in inpatient settings involves not only delivering interventions to address admission problems, but also providing basic living needs such as food and access to outside space, a daily routine and activities suitable for different patients. *Treatment* can involve more intensive or potentially risky interventions than are possible in the community because closer monitoring and more sustained staff presence are possible. Inpatient care provides a setting where coercion has a legal basis for detained patients and the social context of a hospital with rules and customs, staffed continually by nurses and doctors, helps encourage compliance through a *legitimate authority.* This can involve direct *containment* of patients, denying leave from hospital, using seclusion or continual staff presence to observe a patient or forcibly administering treatment. Finally, inpatient care provides continuous staff presence: tasks are less proscribed by staff availability and opportunities for interventions can be identified and capitalised on whenever they present. The close proximity of staff and patients in inpatient settings and spending time with patients not always in a goal orientated way allows relationships to develop which may potentially aid assessment and persuasion to accept treatment or be therapeutic in their own right. Bowers describes this as *presence+* (Bowers et al. 2009).

Although not included in the visual representation of the model in Figure 1.1, the Bowers model also helps understand the context and criteria for discharge. This will usually occur when the primary admission task is complete, whatever admission bonus or progress with secondary admission tasks has been achieved.

Bowers and colleagues (2009) stress that their model describes how inpatient wards do operate currently in the UK, rather than an idealised model of how they could or should function. Three ways in which it is particularly useful are:

- While elucidating reasons for admission, it acknowledges the complexity of the decision to admit, which will be influenced by local service resources and the practice of the admitting clinician as well patient variables. It shows how the characteristics and problems of patients at admission may vary considerably within and between inpatient services.
- It highlights that the function of admission goes beyond resolving a patient's primary admission problem. Inpatient care offers an opportunity

to address related or additional problems patients have. It also acknowledges the potential of inpatient care to create or exacerbate problems. The extent to which an admission bonus or help with secondary admission problems is achieved during admission and iatrogenesis is avoided may have a great impact on patients' experience of care and outcomes.

- It illustrates that the nature of inpatient care is distinct from care provided in community services. Containment and coercive interventions are an important part of inpatient care. Presence – staff contact with patients – is also common and important, even when not overtly directed to specific goals. High levels of staff-patient contact are required for the detailed, ongoing assessment provided in inpatient services and can have its own therapeutic benefit. An implication of the Bowers model is that assessment of inpatient care may be inadequate if focused only on the provision of specific interventions: it may also need to consider the amount of time staff spend with patients overall.

1.3 Effectiveness and patient experience: the evidence regarding acute inpatient wards

Despite the development of community care, an ongoing need for inpatient services has been acknowledged (Szmukler and Holloway 2001). Treatment at home is not always practical or desirable, particularly when a patient's risk to self or others is too great to allow long periods of time without supervision or where home environment is exacerbating a patient's crisis (Johnson et al. 2007). There is therefore a need to establish effective inpatient services; however, an ongoing lack of research evaluation of inpatient services has been acknowledged repeatedly this decade. Jepson et al. (2000) identified only one systematic review on a theme related to acute in-patient care: this found no differences in outcomes between routine admissions and planned short hospital stays (Johnstone and Zolese, 1999). Szmukler and Holloway (2001) characterise inpatient care as, in research terms, the Cinderella of contemporary mental health services. The UK Department of Health report a dearth of mental health research studies relevant to mental health inpatient care (Department of Health 2005) and Bowers et al.

(2009) concludes that research into the effectiveness of hospital care remains largely absent.

More is known about patients’ experience of inpatient care from a number of qualitative investigations and surveys conducted in the last two decades. Boredom and difficulty in securing time to talk to staff have been consistently reported by patients (Higgins et al. 1999, Sainsbury Centre for Mental Health 1998, Rose 2001, Sainsbury Centre for Mental Health 2006). In a survey of the quality of care on acute wards conducted by the Sainsbury Centre for Mental Health (1998), 40% of patients reported taking part in no social or recreational activity during their admission; 30% in no structured activity of any sort. Baker (2000) surveyed 343 ex-patients from inpatient wards: 56% of respondents reported finding inpatient wards not conducive to recovery. 57% said they did not have enough contact with staff; of these, 82% reported spending 15 minutes or less with staff per day. Rose (2001) reported patients’ perceptions that nurses prefer the environment of their office to the rooms which the patients used, with too many spending their time chatting with other nurses or doing paperwork. Service users have described a limited range and quality of activities available on inpatient wards (Healthcare Commission 2003). Qualitative interviews with inpatients (Gilburt et al. 2008) identified the process of talking and feeling listened to by staff as of prime importance to patients and a necessary concomitant to the success and acceptability of specific interventions such as medication prescription. Patients reported that the unavailability of staff on acute wards was a major obstacle to effective communication and a factor limiting desired activity, such as spending escorted time off the ward.

This consistent picture from patient-report of limited activities and staff availability in inpatient wards has some support from surveys of inpatient staff and observational studies. Higgins and colleagues (1999) surveyed the views of nurses on acute inpatient wards. Staff reported feeling very restricted in the time they could spend in direct contact with patients. The main focus was on discharge planning rather than providing care to patients during an admission. Staff activity which did involve direct patient care tended to involve responding to the crises of a minority of severely unwell, demanding patients rather than providing proactive, planned care for all patients. Using an observational measure, the authors found that the more senior and experienced nursing staff were, the less time they spent with patients, mainly performing administrative or managerial duties instead. On average, patients spent just 4% of time in direct interaction with staff. Ford and colleagues (1998) described a similar situation in a survey of 119 UK acute inpatient wards conducted for the Mental Health Act Commission. During their visits, at the time of observation they found no nurses interacting with patients on 26% of wards. Where staff were visible on the wards, they were frequently engaged in activities which offered limited opportunity for meaningful therapeutic engagement with patients. On 60% of wards visited, staff were noted conducting frequent observations (checking and recording the whereabouts and safety of patients at regular intervals). On 11% of wards, staff were observed on door duty (sitting by the main entrance to the ward and signing patients and visitors in and out). Garcia and colleagues (2005) surveyed 303 managers of acute wards: ward managers on 36% of wards reported that no social or leisure activities were routinely available. Rising levels of acuteness and Mental Health Act detention has also been reported as leading to an increasingly disturbed and disturbing environment in inpatient wards (Patrick 1989). The Mental Health Act Commission (2008) concluded that acute wards appeared to be *“tougher and scarier places than they were a decade ago.”*

The studies above suggest that patient dissatisfaction with inpatient care is common and that the availability to patients of time with staff and organised activity is often limited and experienced as insufficient. They suggest patient satisfaction with inpatient admission may be strongly influenced by the level of staff-patient contact provided, although they do not provide empirical evidence of this relationship.

1.4 Residential alternatives to standard acute inpatient care

A range of innovative inpatient services and non-hospital residential acute services has been developed: these will be described in Chapter 2. They are of interest given the broad agreement that inpatient acute services need improvement (Johnson et al. 2007) and limited current knowledge about components or models of effective inpatient care. The term “alternatives” will be used throughout the thesis to describe residential services providing an alternative to standard acute inpatient admission. A definition of alternatives was developed for a UK study (Johnson et al. 2007). The criteria for alternatives were broadly defined in order to minimise the risk of missing important, innovative models. Services were classed as alternatives which were: residential services for adults aged 16-65 with acute mental disorders, which also meet at least one of the following criteria:

- Based in the community, i.e. non-hospital services
- Time-limited, meaning services offering admission with a time limit or planned maximum stay of 14 days or fewer
- Dedicated to a specific diagnostic group
- Dedicated to a specific socio-demographic group
- Implementing a specific therapeutic model involving changes to the working practices of more than one professional group.

1.5 Mental health service evaluation and content of care measurement

The measurement of patient outcome has risen to prominence over the last forty years (Donabedian 1966, Ellwood 1988). Three levels of evaluation can be differentiated (Burns and Priebe 1996): a) treatment level, concerning specific interventions; b) programme level, concerning combinations of treatment components provided by a service; and c) system level, concerning all programmes for a defined target group in a specific area. Outcomes can be considered at each level. Outcome measurement can serve numerous purposes including evaluating the clinical and cost effectiveness of interventions, clinical audit, service planning, quality improvement and as an aid to clinical decision making in routine clinical practice (Gilbody et al 1992). In the UK, a new emphasis on service evaluation and outcome measurement emerged in the 1990s, driven by factors including:

- the setting of public health targets for mental illness in “Our Healthier Nation” (Department of Health 1998)
- the synthesis of available evidence and setting of standards in the National Service Framework for Mental Health (Department of Health 1999)

- the move towards clinical guidelines provided by the National Institute of Health and Clinical Excellence (NICE) - for example, guidelines for the management and treatment of schizophrenia (National Collaborating Centre for Mental Health 2009)

Slade (2002) highlights challenges to outcome assessment. Different outcome measures may be needed to assess treatment, programme and system outcomes, while perspectives on what constitute important or positive outcomes may differ. The time period for outcome assessment is complicated by the possibility that different types of outcome may be desynchronous. Interpretation of outcome measurement is not straightforward in mental health settings when the effect of best quality care may at times be only to maintain current levels or slow decline. Slade (2002) concludes that the solution that has evolved in research studies has been to assess a wide range of treatment and programme level outcomes from multiple perspectives. Outcome measurement has developed from a narrow focus on levels of symptomatology and service use to a broader assessment of the impact of illness and treatment on an individual (Tansella and Thornicroft 2001). In addition to costs and service use, patient outcome measurement and service evaluation may now include the following domains: physical health, mental health, social functioning, role functioning, patient perception of health and well-being, needs and satisfaction with services (Gilbody et al. 2002). Perspectives which can be addressed in outcome measurement include those of patients, clinicians, carers and the tax payer (Clifford 1998).

The focus on outcome measurement in recent decades has been accompanied by acknowledgement that, by itself, it has limitations as useful service evaluation. Donabedian (1992) argues that both structure (the organisational properties of the settings where care is provided) and process (what is done for patients) will influence patient outcomes: to understand what influences the effectiveness of services and define service quality, process factors which are associated with good outcomes must be identified. Burns and Priebe (1996) and Mechanic (1996) highlighted that there is often variation in practice among purportedly similar services: detailed investigation of the content of service provision is necessary to understand variation in service outcomes.

Burns recently reiterated the need to characterise the interventions being assessed in mental health service evaluations (Burns 2009). To illustrate the importance of this, he uses the example of the evaluation of Assertive Community Treatment (ACT), a type of intensive community support for people with enduring mental health problems. Initially puzzling, discrepant results from studies evaluating the impact of ACT services in reducing inpatient admissions can be explained by variation in the content of the comparison service: ACT is only effective in reducing hospital bed use in studies where hospital use in the comparison service is relatively high (Burns et al. 2007). Burns (2009) concludes that undefined treatment as usual is common but inadequate as a comparison service in mental health services evaluation: services in both arms of a trial should be carefully characterised and their content assessed.

Understanding variation in service content to aid service evaluation is highly relevant for acute inpatient care which, as the model of Bowers and colleagues (2009) illustrates, exhibits complexity and potential variation in populations served, service function and modes of operation. The Medical Research Council has recently provided updated guidelines and a framework for evaluating complex interventions in healthcare (Craig et al. 2008). These guidelines (referred to hereafter as the MRC guidelines) propose four elements of the development and evaluation of complex interventions, described in Box 1.1.

Box 1.1 key elements in developing and evaluating complex interventions: the MRC Framework

Development:	Identifying the existing evidence base
	Identifying and developing theory
	Modelling process and outcomes
Feasibility and piloting:	Testing procedures
	Estimating recruitment and retention
	Determining sample size
Evaluation:	Assessing effectiveness
	Understanding change process
	Assessing cost—effectiveness
Implementation:	Dissemination
	Surveillance and monitoring
	Long-term follow-up

The MRC guidelines acknowledge that the process of developing and evaluating a complex intervention may not always follow the phases described in Box 1.1 in a linear or cyclical fashion; rather, there may be an iterative process of clarifying and refining the intervention being assessed and methods of assessment, leading to an increasingly definitive evaluation.

Three challenges to evaluating complex interventions highlighted in the MRC guidelines are relevant to evaluation of inpatient services:

- evaluation may follow widespread implementation of an intervention rather than precede it, limiting how far a researcher can modify it or affect its implementation
- Ideal evaluation designs are not always practicable; observational designs may be justifiable: a judgement is needed regarding the trade off between the importance of the intervention and the value of the evidence that can be gathered
- Strict standardisation of an intervention may not always be desirable: a level of adaptation to local circumstances may be appropriate

The MRC guidelines stress that process evaluation is not a substitute for evaluation of outcomes, but has a place in the evaluation of complex interventions to: a) inform the process of modelling interventions; b) assess fidelity and variation in implementation of an intervention; and c) help understand outcomes from an evaluation and causal mechanisms for an intervention's success or failure. All of these are relevant to evaluation of acute inpatient services.

1.6 Aims of this thesis

This thesis has three aims:

- to provide an overview of types of residential acute care which have been developed as alternatives to admission to a standard acute inpatient mental health ward, then to systematically review the evidence for their effectiveness and acceptability and how their content of care has been measured.
- to identify, or develop if necessary, measures appropriate to assess the content of care provided in acute residential and inpatient mental health services.
- to provide a quantitative assessment of the content of care provided at four standard UK inpatient acute wards and four alternative residential acute services and explore how care received may affect patients' satisfaction with services.

The quantitative investigation of content of care will focus on three research questions:

a) Is there more staff-patient contact at alternatives than at standard services?
b) Do the types of care provided differ between alternative and standard services?
c) Can differences between alternatives and standard services in patient satisfaction be explained by the content of care provided?

The quantitative investigation will provide a description and comparison of care provided at acute inpatient services and alternative types of crisis residential care.

It uses a hypothetico-deductive approach (Popper 1963) common in health services research. Six hypotheses will be tested:

1) The proportion of patients observed in contact with staff is greater at alternatives than standard services.
2) Staff-reported intensity of care is greater at alternatives than at standard services.
3) Patient-reported intensity of care is greater at alternatives than standard services.
4) The staff-reported and patient-reported intensity of social interventions is greater in alternatives than standard services.
5) The staff-reported and patient-reported intensity of psychological interventions is greater in alternatives than standard services.
6) The staff-reported and patient-reported intensity of physical and pharmacological interventions is greater in standard services than alternatives.

Three reasons why a focus on content of care and level of direct staff-patient contact at services in particular is of interest include: a) the availability of and contact with staff has been identified as important to patient experience (Rose 2001, Gilburt et al. 2008); b) maximising the time inpatient staff spend engaged with patients is an aim of UK health policy (Department of Health 2005); and c) “Presence+” (Bowers et al. 2009) is an important element of inpatient services’ mode of operation and need to be assessed in description of service content.

This thesis does not assess outcomes at inpatient services and alternatives. It does not provide an evaluation of the effectiveness of services. It does contribute to the development and evaluation of residential and inpatient acute services in accordance with the elements required for evaluation of complex interventions described in the MRC guidelines (Craig et al. 2008). A review of literature regarding the effectiveness of alternatives can identify their existing evidence base. The development of measures of content of care contributes to the feasibility and piloting of procedures which can be used in process evaluation of inpatient services and alternatives. The description and comparison of care at alternatives and inpatient services can help refine definition of inpatient and residential acute service models for evaluation in future research and aid understanding of one measured outcome from The Alternatives Study – patient satisfaction – with which to generate hypotheses and inform procedures for future more definitive trials.

1.7 Relationship of this thesis to The Alternatives Study

Data for this thesis were collected as part of The Alternatives Study (Johnson et al. 2007), a national research study funded by the NHS Service Delivery and Organisation Programme. The Alternatives Study evaluated residential alternatives to standard acute inpatient mental health services. The following components of this thesis were guided by decisions already made concerning the structure of The Alternatives Study:

- Results from UK national service mapping of residential alternatives, reported in Chapter 2.
- Inclusion criteria for residential alternative services (Literature review, Chapter 3).
- Choice of services participating in the quantitative study in this thesis (Chapters 6-9).
- Choice of measure of satisfaction (Client Satisfaction Questionnaire) used in the quantitative study in this thesis.
- Sampling frame for Client Satisfaction Questionnaire and patient report content of care measure (CCCQ-P) used in the quantitative study in this thesis.

Except where acknowledged, all other elements of this thesis represent the author’s own work.

Chapter 2

Alternative service models: an overview

2.1 Introduction

The problematic nature of acute inpatient mental healthcare in the UK and the desirability of developing effective alternatives to standard inpatient care were discussed in Chapter 1. This chapter provides a narrative account of the models of residential alternative to standard acute psychiatric inpatient wards which have been developed. Evidence for the effectiveness and acceptability of alternatives will be systematically reviewed in Chapter 3.

Clear descriptions and operational definitions of the different types of residential service that offer an alternative to admission are few. Braun and colleagues (1981) distinguished between alternatives to hospital admission (i.e. non-hospital, community-based services) and modifications of conventional hospitalisation (i.e. inpatient alternatives to standard wards). Stroul (1988) subdivided community-based alternatives into individual-based approaches such as short-term family placement and group-based approaches such as residential crisis houses. The Alternatives Study (Johnson et al. 2009) identified community-based residential crisis services as an alternative to standard acute wards and proposed four criteria by which a hospital inpatient service might constitute an alternative to a standard ward:

- Dedicated to a specific diagnostic group
- Dedicated to a specific socio-demographic group
- Operating a fixed maximum length of stay
- Implementing a specific therapeutic model involving changes in the practice of more than one profession within the service.

2.2 Community-based alternatives

Stroul's dichotomy (Stroul 1988) of individual-based approaches and group based approaches will be used to distinguish different models of community-based alternatives.

2.2.1 Individual-based approaches: Stroul (1988) surveyed community-based residential crisis facilities in the USA and reported that short-term housing and support at the homes of carefully selected families was the most widely available form of residential crisis care. Polak and colleagues evaluated family sponsor homes developed in the 1970's in Denver Colorado specifically to divert people from acute hospital admission (Polak et al. 1979). Host families supported one or two acutely ill patients, with training and round-the-clock access to advice or assistance from local crisis services. Support included involving patients in a normal family environment with participation in meals and other domestic tasks and activities. This service model has been replicated in Wisconsin, USA by Stein and colleagues (Stein 1991; Bennett, 2002), where it has run for over twenty years. An "accredited accommodation" scheme established in Powys, Wales in the last decade (Readhead et al. 2002) also sought to place acutely ill patients with individual families, although authors report the scheme in practice was frequently used for respite care or sub-acute problems. Hoult and colleagues (1983) describe the use in Australia of boarding houses to accommodate patients in crisis or for respite, with support from the local home treatment team.

2.2.2 Group-based approaches

One model of non-hospital crisis alternative is to provide beds alongside community mental health services. This model has been most extensively developed in Trieste, where crisis beds are provided within large Community Mental Health Resource Centres (Mezzina and Vidoni 1995). Similar services have been described in France (Katschnig et al. 1993) and the UK (Boardman et al. 1999), where in North Staffordshire, a number of small local resource centres with residential units have been used for over a decade to prevent hospital admission whenever possible. Wesson and Walmsley (2001) have described a community-based unit in Southport that combines day care and crisis admission beds. These beds have a maximum stay of three days and are used for a variety of purposes, including as an alternative to hospital, for early discharge from hospital and for patients needing supervision when starting on new medications.

Residential units in the community offering short-term emergency admission, sometimes known as crisis houses, have also been set up independently of larger community mental health units, with different levels of association with statutory mental health services. Crisis houses are smaller than traditional inpatient services, typically with 6-10 beds, and situated in buildings in residential streets, often not immediately identifiable as mental health units. Perhaps the best known type of crisis house is the Soteria service (Mosher et al. 1975; Mosher 1999). Established in California, USA by Mosher and colleagues, the original Soteria houses operated from 1971 to 1983. They provided care to people with first or second episode psychosis in informal settings. Primary staff were not clinically trained and worked long shifts (36-48 hours), designed to help them to attune to and engage with residents. Staff and residents shared responsibility for household tasks. There was minimal reliance on anti-psychotic medication: Bola and colleagues (2003) report results from two studies that 43% of Soteria residents were not medicated at all during admission or two year follow-up. The model has been replicated more recently in a number of European countries (Switzerland, Germany, Sweden, Hungary, Finland), publicised and evaluated primarily by Ciompi and colleagues in Bern, Switzerland (Ciompi and Hoffman 2004).

McCabe and colleagues (2004) reviewed residential crisis services and identified a number of other descriptions or evaluations of crisis house services developed in North America, although the model has not become a common part of acute service provision. Description of treatment provided or philosophy of care is often limited; however most services appear to adhere to traditional working patterns and clinical practices more closely than the Soteria model. Crisis houses have been established specifically for veterans (Hawthorne 2005) and dual diagnosis patients (Timko 2006) as well as general acutely ill populations.

Davies and colleagues (1994) suggest than crisis houses are the model of residential alternative to admission which has been most publicised in the UK. A report for The Mental Health Foundation (Faulkner et al. 2002) described eight residential crisis services, highlighting potential benefits of crisis houses in improving service user choice, reducing stigma and alleviating pressure on inpatient beds. This report included description of a user-led service in Birmingham called Anam Cara, which placed particular emphasis the value of peer support and using people’s own resources in coping with a crisis. Description and qualitative evaluation have also been provided for a women’s crisis house in North London called Drayton Park (Killaspy et al. 2000; Johnson et al. 2004). Drayton Park has run for more than a decade. It has close links with statutory mental health services, a majority of its residents have a history of previous hospital admissions and it is reported as highly valued by service users for providing a safe, calm and supportive environment.

2.3 Inpatient alternatives

Time-limited services: Brief stay inpatient services were first developed over forty years ago. Two reviews from the 1980s (Mattes 1982, Braun et al. 1981) identified a number of British and American studies of services aiming to reduce length of acute stay by providing intensive treatment including frequent medical review and planning discharge and aftercare from an early stage of admission. Some older studies evaluate services which might not now be considered brief-stay. Glick and colleagues (1975), for example, compared a brief-stay ward of 21-28 days with standard care of 90-120 days’ admission. The mainly positive results of such studies may have influenced the implementation of briefer stays as standard in contemporary acute inpatient care, where admissions of a month or less are the norm. Other studies, some also more than thirty years old, describe services with much briefer planned admissions, such as one week (Herz 1975; Mendel 1966) or three days (Voineskos 1974). Despite the trend towards provision of more intensive community acute care, such as home treatment, which might increase the feasibility of short-stay wards by facilitating early discharge, these very brief admission services do not appear to have become well established in practice or researched further. A recent Cochrane review of brief admission services (Alwan et al. 2008) identified no randomised controlled trials more recent than 1980 and no studies at all from after 1993.

Services with a distinct therapeutic model: Two nursing-led inpatient initiatives have been described in the literature: the Tidal Model (Barker 2001, Stevenson et al. 2002) and the Refocusing Model (Dodds and Bowles 2001). These are both aimed at changing the overall milieu on inpatient wards, are broader in scope than a single specific intervention and are intended to affect the working of the whole staff team. They can therefore be described as distinctive therapeutic models.

The Tidal Model was developed in Newcastle UK in the late 1990s. Stevenson and colleagues identify four main goals of the model (Stevenson et al. 2002):

- increased collaboration between staff and patients in care planning
- greater use of 1:1 time and group work as therapeutic interventions to promote well-being
- greater emphasis on the patients' own narrative experiences of illness and health
- an increased role in nursing-led care and formulations of problems within the inpatient staff team

The Tidal Model involves named nurses completing daily written care plans with patients, using patients' own language wherever possible. Gordon and colleagues (2005) report two UK evaluations of the implementation of the model. These papers reported multiple outcomes but indicated a possible association between introduction of the Tidal Model and reduced levels of conflict and untoward incidents on the wards. The Tidal Model has also been exported to inpatient services in a number of countries including Finland (Virtanen 2003), New Zealand (Cook et al. 2005) and Canada (Berger et al. 2006).

The Refocusing Model, developed in Bradford, UK shares with the Tidal Model an aim of increasing the amount and quality of staff-patient contact and offering patients "the gift of time" (Dodds and Bowles 2001). It is characterised by the devolution of responsibility for practical decisions about patient care from medical to nursing staff, the reduction and eventual cessation of formal observations on wards, with the substitution of increased 1:1 time and structured group activities. Regular community meetings for staff and patients are used to aid communication and collaboration. Dodds and Bowles reported that the implementation of the model in Bradford was associated with reductions in rates of absconding, self-harm and violence on the wards and increased ability of patients to name their allocated primary nurse (Dodds and Bowles 2001).

Socio-demographic and diagnosis-specific services: Mother and Baby Units (MBUs), typically small inpatient units where a mother may stay with a baby of up to one year old during a period of acute illness, were reported as a part of UK NHS acute care over fifteen years ago (Royal College of Psychiatrists 1992). Mental health care may be augmented by input from nursery nurses and other child health and welfare specialists. MBUs are recommended in the UK as the optimal treatment setting for treating perinatal psychosis (Royal College of Psychiatrists 2001).

Mathews and colleagues (2002) described the organisation of inpatient services in San Francisco, USA into ethnically-focused services for black, Asian, Latino and white groups. While not dedicated exclusively to specific ethnic populations, these services sought to provide a more culturally sensitive service than typical acute wards, e.g. by recruiting staff with appropriate language skills. Matthews and colleagues report more aftercare was arranged by staff for patients to whom they were ethnically matched, which raises the possibility that a patient's needs might be more appreciated by staff from a culturally similar background.

Bonsack et al. (2001) reported a qualitative study of inpatient services in Lausanne, Switzerland following reorganisation into diagnosis-specific wards for people with psychosis, affective disorders, personality disorders and dependencies. They noted increased clarity in the process of care but additional problems in referring patients to the appropriate service.

2.4 Alternatives in England: results from The Alternatives Study

The reporting of alternative services internationally has been piecemeal and it is not clear how many such alternatives have been developed in different countries. Phase 1 of the Alternatives Study (Johnson et al. 2009) provided some epidemiological data about the prevalence of alternatives in the acute care system in England. It involved a national survey in 2005/6 to identify the extent of provision of alternatives in England and to develop a typology of residential alternatives to standard acute wards for adults of working age and describe their distribution and clinical populations.

The criteria for defining alternatives which are used in this thesis derive from those used in the Alternatives Study. Services were considered alternatives if they: a) served adults aged 18 to 65 years who would otherwise be admitted to an acute ward; and b) involved patients staying overnight at the service; and c) met at least one of the criteria referred to in Section 1.2 (community-based, for a specific diagnostic group or for a specific socio-demographic group, implementing time-limited, brief admission or using a distinctive therapeutic model).

131 alternative services were identified in the survey. 94 were within the National Health Service, 16 run by voluntary sector and 21 by private sector organisations. Services meeting each of the criteria for what constitutes an alternative were identified: numbers of services of each type are shown in Table 2.1.

Table 2.1 Alternatives identified by The Alternatives Study

Inclusion criterion met	**No. of alternative units**
Disorder specific	7
Socio-demographic specific	28
Specific therapeutic orientation	50
Time-limited admission	5
Community-based	41
Total	131

Telephone questionnaires, including a one-night census of the number and characteristics of service users, were completed with 109 alternative service managers (83% of those surveyed). These data indicated that alternative services provide about 1,300 acute beds, a majority of which were in NHS services. 250 of these beds were in community-based services. These figures compared with a total of 12,400 acute inpatient beds available in statutory services in England (Department of Health 2006). All alternative services would accept referrals from statutory mental health services. Alternatives were found to be serving a population including the severely mentally ill (substantial numbers of people with psychosis were represented in all types of service) although only 5 community-based services could accept detained patients directly from the community. The ethnicity of patients in alternative services corresponded with what would be expected from national census data matched to service location. Men were under-represented in alternatives, constituting only 40% of patients.

Some types of care reported to be available within the hospital services were less likely to be provided by the community alternatives. Medical interventions such as medication reviews and investigation and treatment of physical health problems were less frequently provided in alternatives. Psychological treatment and provision of structured groups and activity were also less common. It was unclear how or whether community-based alternatives compensated for some of these apparent gaps (for example, crisis resolution teams may be major providers of interventions such as medication reviews and supervision in some services). It is also possible that other forms of help which may have been provided by alternatives were not measured in the survey. However, clear indications of different models of care from standard inpatient services in alternatives were difficult to determine.

A cluster analysis of questionnaire data from The Alternatives Study survey derived a typology of alternatives containing eight categories of service. The service types are described below.

Community-based service types

1. **Clinical crisis houses (n=13):** These community-based services shared many characteristics with hospital wards. A high proportion (42%) of staff was nurses. Care Programme Approach meetings were held at all services and most had waking night staff. All were managed as well as funded by the statutory sector.
2. **Specialist crisis houses (n=5):** Most similar to clinical crisis houses, these services were targeted at specific client groups, e.g. women or people with first episode psychosis.
3. **Crisis team beds (n=13):** These services were typically very small and had the shortest lengths of stay. Staff tended not to be qualified mental health

professionals. Services were closely integrated with crisis/home treatment teams, which could provide additional support to clients when required. Service users of crisis team beds presented with lower severity of illness than users of all other service types.

4. **Non-clinical alternatives (n=11)**: Most services of this type were managed by voluntary sector organisations and reported less integration with statutory mental health services than other types of community alternative. Staff tended not to be qualified mental health professionals.

Hospital-based service types

5. **General therapeutic wards (n=35):** Services of this type all employed a specific therapeutic model. Private hospitals (n=11) typically provided a flexible programme of group and individual therapy, most usually based on based on cognitive behavioural interventions. NHS acute admission wards (n=24) reported using nursing-led models of care, most frequently The Tidal Model or the Refocusing Model. This service type had the highest bed numbers, and served a severely ill patient group.
6. **Wards for specific demographic groups (n=20)**: All but one of these services were Mother and Baby Units. One was a service for deaf people. They tended to be fairly small, relatively long-stay services.
7. **Therapeutic wards for specific groups (n=4):** These services could provide interventions for specific diagnostic groups, e.g. a dialectical behaviour therapy ward for people with borderline personality disorder. Three of the four were in the voluntary or private sector, the exception being an NHS unit in Lambeth for people with early psychosis.
8. **Short-stay wards and general wards for specific groups (n=9):** This group of hospital services included wards with a fixed brief length of stay (typically less than a week) and wards for a specific diagnostic group which did not describe provision of specific interventions or a tailored model of care for that group.

2.5 Conclusion

Studies over several decades suggest that a variety of alternative service models may be viable ways to provide residential acute mental health care. The

Alternatives Study survey indicates that alternative services, in hospital and community settings and with different service models, constitute a significant part of the acute inpatient mental health care system in England. Services can be found in the UK acute care system representative of each criterion proposed for defining an alternative, encompassing eight groups of services with distinct characteristics.

Several of the services described in this review have been short-lived. Similar services have been established in more than one setting or at different times without becoming fully embedded in any national acute care service system. This provokes some doubt about the sustainability or usefulness of alternatives, although The Alternatives Study survey did identify a number of community-based crisis houses more than a decade old (Johnson et al. 2009). The lack of a clearly established evidence base for their effectiveness may partly explain the reticence of administrators in expanding the provision of alternatives. With some exceptions, such as studies of the Soteria services (Mosher 1999, Ciompi 2004), the literature also fails to provide detailed description or measurement of the content of interventions provided by alternatives. There is therefore uncertainty about the nature or extent of innovation from standard care that alternatives provide.

The interest in alternative models of crisis care, given the shortcomings of traditional acute wards described in Chapter 1, the range of alternative services which have been established and their not insignificant role in the current UK acute care system all point to a need for clear evidence about alternatives' effectiveness. This evidence is currently lacking. A systematic review of the effectiveness and acceptability of residential alternatives to standard acute care and a collation of information provided about the content of care they offer will therefore be provided in Chapter 3.

Chapter 3

Residential alternatives to acute psychiatric hospital admission: a systematic review

3.1. Reason for the review

Reviews have compared residential alternatives and standard acute inpatient care. Joy and Saylan (2007) conducted a systematic review of one type of socio-demographic specific alternative, Mother and Baby Units, finding no studies for inclusion. Alwan and colleagues (2008) systematically reviewed brief-stay inpatient wards. Their definition of brief-stay wards was very broad however: included as alternatives in their analyses were services offering up to 28-day admission – arguably closer to current standard care than an alternative. Grawe and colleagues (2005) included some residential alternatives in a Norwegian language, broader review without meta-analysis of alternatives to acute admission, which also included home treatment. They searched only two electronic databases, included only randomised trials in the review, identified few relevant studies of residential alternatives and concluded there was insufficient evidence to comment on their effectiveness. Their conclusions suggest a comprehensive search strategy and inclusion of a broader range of study types may be required in a review to establish existing evidence about the effectiveness of residential alternatives. McCabe and colleagues (2004) reviewed residential crisis services and identified numerous relevant studies: this review though included any evaluation of residential alternatives, not just comparisons with standard inpatient care, is based on a search of only one electronic database (Medline) and does not meta-analyse results from different studies. Calton et al. (2008) reviewed studies of Soteria model services (crisis hostels for people with first onset psychosis, using minimal pharmacological treatment) but did not meta-analyse results. There are also a number of older, narrative reviews, focusing on North American literature, of time-limited or community alternatives, which identified some relevant studies (Arce and Vergare 1985), (Braun et al. 1979), (Mattes 1982), (Stroul 1988).

The studies included in the reviews above typically report outcomes for alternatives comparable to standard acute wards or favourable to the alternative.

However, these reviews provide limited information about the overall evidence regarding the effectiveness of residential alternatives because none uses a comprehensive search strategy, includes evidence from all relevant study designs and, where appropriate, meta-analyses data from included studies. A systematic review of all relevant studies, meta-analysing outcomes data from different studies where appropriate, thus has potential to clarify what is known about the effectiveness and acceptability of residential alternatives, adding to existing knowledge.

3.2. Objectives

This review used the criteria proposed by Johnson and colleagues (2009) to define a service as a residential alternative to standard acute inpatient care. The evidence for each type of alternative was reviewed. A systematic review of the literature was undertaken to address the question: are residential alternatives more effective, cost effective and acceptable to patients than standard acute inpatient services. A second focus of the review was to identify whether and to what extent studies comparing alternative and standard inpatient services measured the content of care provided.

3.3 Method

3.3.1 Inclusion criteria

Types of study

All studies reporting a specific quantitative comparison of the effectiveness or acceptability of residential alternatives and standard acute inpatient services were included in the review, i.e.:

- Randomised controlled trials
- Two group non randomised cohort studies (prospective quasi-experimental studies or retrospective cohort studies)
- Interrupted time series studies

Studies were excluded if they involved residential alternatives as one element of a broader package of care (e.g. day hospital + crisis admission if required), as such a design does not allow a clear comparison of the effectiveness of residential alternatives and standard care.

Types of participants

Adults aged 16-65 assessed by mental health professionals as needing acute inpatient admission.

Types of intervention

a) Standard acute inpatient mental health services

b) Residential and inpatient acute mental health services which offer an alternative to standard acute psychiatric wards in one of the following ways:

i) Services for a specific clinical subgroup (e.g. a specific diagnosis)

ii) Services for a specific socio-demographic population (e.g. a defined ethnic group). (In order to maintain consistency with the criteria for alternatives proposed by Johnson and colleagues (2009), single sex wards were excluded from this review: although catering for a specific socio-demographic group, they are increasingly a feature of standard UK acute inpatient care.)

iii) Services offering admission with a time limit or planned maximum stay of 14 days or fewer

iv) Services using a specific therapeutic model, distinctive from current standard acute inpatient care and affecting the normal working of at least two professional groups

v) Services which are community-based (i.e. non-hospital services)

Types of outcome measures

Studies investigating the following outcomes were included in this review:

1) Improvement (clinical and social outcome)
Clinical improvement, general or specific
Social functioning, including life skills
Relapses
Untoward incidents (including self-harm, aggression to others)
Quality of life

Discharged home/ Living independently
Employed

2) Service use
Number of days in hospital or residential alternatives
Readmission during study period

3) Satisfaction
Patient satisfaction with services

4) Cost
Total public or health care costs

3.3.2 Search strategy

The following methods were used to identify studies for the review:

a) Search of electronic databases

A systematic search of seven electronic databases was undertaken. Medline, PsycInfo, Web of Science, Cinahl, Embase, Cochrane Library (CENTRAL and DARE databases) and NHS Economic Evaluation Database were searched for the period January 1966 – February 2008. No language restrictions were applied.

Search terms for acute, residential mental health services were combined with terms for different types of alternative. The search terms used in the Medline search are presented in Table 3.1 below. Searches were conducted using Medical subheadings (MeSH terms) and within title and abstract. Search terms were modified as necessary to search other databases.

Table 3.1 Alternatives review: search terms

Default search limits = title and abstract (except where otherwise stated), Years 1966- February 2008		
#	**Search term**	**Description**
1	mental disorders[MeSH].exp	
2	mental	
3	psychiatr*	
4	#1 OR #2 OR #3	Mental health services
5	crisis intervention[MeSH].exp	
6	crisis	
7	acute	
8	emergency	
9	#5 OR #6 OR #7 OR #8	Acute services
10	residential treatment[MeSH].exp	
11	residential	
12	hospital*	
13	inpatient*	
14	"crisis house"	
15	"community beds"	
16	"crisis beds"	
17	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16	Residential services
18	#4 AND #9 AND #17	Acute, residential, mental health services
19	"disorder specific" OR "disorder-specific"	
20	specialist	
21	#19 OR #20	Disorder-specific services
22	"socio-demographic" OR "sociodemographic"	
23	ethnic*	
24	disabled.tw OR disability	
25	"visual* impair*" or blindness	
26	deaf	
27	"mother and baby"	
28	"parent and child"	
29	#22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28	Socio-demographic specific services
30	"therapeutic model"	
31	"model of care"	
32	alternative	
33	innovative	
34	#30 OR #31 OR #32 OR #33	Specific therapeutic model services
35	"brief admission"	
36	"brief stay"	
37	"short stay"	
38	"time limited"	
39	#35 OR #36 OR #37 OR #38	Time-limited services
40	community	
41	"non-hospital"	
42	#40 OR #41	Non-hospital services
43	#21 OR #29 OR #34 OR #39 OR #42	Alternative services
44	#18 AND #43	Alternative, acute, residential, mental health services
45	"sponsor homes"	
46	"accredited accommodation"	
47	Soteria	
48	"crisis intervention" AND Trieste	
49	"Tidal Model"	
50	refocusing AND Dodds.au	
51	#45 OR #46 OR #47 OR #48 OR #49 OR #50	Terms identifying well-known alternative services
52	#44 OR #51	Alternative, acute, residential, mental health services

b) Reference searching

Reference lists of studies identified through the electronic search for inclusion in the review and of review articles were handsearched for relevant studies.

c) Grey literature searching

Relevant studies reported in dissertations, conference reports or other sources other than published journals were sought from the following two sources:

i) British Library Directory of Published Proceedings (a directory of conference reports): a search using British Library electronic resources by title word and keyword was undertaken, using equivalent search terms to those listed for the Medline search in Table 3.1. Conference proceedings were searched by title, then content if potentially relevant.

ii) www.osti.gov/graylit (American database of dissertations and conference reports): Defense and Environment Department collections were searched using equivalent terms to those listed in Table 3.1: all listings were searched by title, then abstract if potentially relevant.

d) Personal contact

A sample of experts was contacted and asked to identify any additional relevant studies. This comprised 12 members of the UK Mental Health Research Network Acute Care Group and 10 non-UK researchers from Europe, USA and Australia with a leading role in the development of alternatives and/or health services research in their country.

3.3.3 Study Selection

Selection of studies for inclusion in the review was conducted by the author of this thesis using the following process. Titles of all identified studies were read. The abstracts of potentially relevant studies were retrieved and read; the full text of studies still considered potentially relevant was then retrieved and read. Studies where the first assessor was uncertain were assessed by a second researcher and a decision on inclusion was reached by discussion.

In order to investigate reliability of selection, 20 studies were assessed by a second researcher, who was blinded to the author and journal title. 10 of these were studies selected by the first assessor for inclusion in the review; 10 had been rejected by the first assessor after reading the article's full text.

3.3.4 Data Abstraction and Quality Assessment

The following data were extracted from all included studies:

- Type of alternative (specific clinical population, specific socio-demographic population, time-limited service, distinctive therapeutic model service, community-based service)
- Name of alternative service
- Country in which study was set
- Type of study:
  1 = Randomised controlled trial
  2 = Prospective non-randomised two-group study
  3 = Retrospective non-randomised two group study
  4 = One group interrupted time series study
- Inclusion criteria for study participants
- Study outcomes
- Outcome measures
- Results: for each type of outcome, whether evidence favoured the alternative, the standard service or indicated no significant difference between services' outcomes.
- Was content of care measured?
  0 = No measurement of content of care
  1 = Specific element(s) only of content of care measured at alternative only
  2 = Specific element(s) only of content of care measured at both services
  3 = A measure of all the care provided at the alternative service only
  4 = A measure of all the care provided at both services

The methodological quality of each study included in the review was assessed using a standard form adapted from the quality assessment tool of Thomas (2003). Studies were rated as strong, moderate or weak regarding potential selection bias, allocation bias, accounting for confounders, blinding, data collection methods and withdrawals, according to the criteria set out in the

Dictionary accompanying the quality assessment tool (Thomas 2003). The analysis strategy and intervention integrity of studies were also noted.

The criteria for assessing study quality using the Thomas (2003) measure were clarified for this review as follows:

Selection bias if study criteria for participants' inclusion are different from service's usual admission criteria, consider it somewhat likely that participants are not typical of the target population.

Confounders variables relating to severity of illness at admission constituted important confounders in this review.

Data collection methods

i) Studies are rated as strong if at least one scale was used previously in a published study and reported to have adequate psychometric properties.

ii) Studies using audit data on death or bed use (length of stay or number of readmissions) are rated as strong for data collection.

iii) Studies are rated as moderate if at least one of their outcome rating scales has been previously used in a published study

iv) Studies using other audit data, e.g. employment status, incident rates on wards, are rated as moderate.

Analysis the following features were recorded:

SS – did a sample size/power calculation inform the size of study groups?

ITT – was it stated that analysis was performed on an intention-to-treat basis?

Integrity the following features were recorded:

E – Exposure: was it stated that at least 80% of subjects received the planned intervention?

C – Consistency: was there any measure of the consistency of the intervention?

Operational criteria were created from quality ratings to distinguish studies of high, moderate and low quality overall. Studies were rated as high quality if they reported allocation concealment during randomisation, analysed data based on an intention-to-treat principle and were rated strong in all domains in the Thomas tool bar blinding. All other randomised controlled trials were rated as moderate quality, as were non-randomised studies which demonstrated: i) no significant difference between experimental and control groups for the confounder of severity of illness at admission; or ii) adjusted in analyses for difference in severity of illness between groups; and iii) rated at least moderate for all Thomas criteria assessed except blinding. (The nature of the intervention made reliable blinding of participants and raters impossible.)

3.3.5 Analysis

All studies of moderate or high quality were eligible for inclusion in meta-analysis. Meta-analyses were conducted using Revman 4.2 software (Cochrane Collaboration 2003).

Analyses were conducted separately for each type of alternative service and each outcome listed in Section 3.3.1. Short, medium and long-term outcomes were analysed separately and defined as follows:

Short term: outcomes at discharge
Medium term; outcomes post discharge up to one year follow up
Long term: outcomes beyond one year follow up

Usable outcomes

Data from eligible studies were excluded from meta-analyses in cases of:

a) inadequate reporting: data could not be analysed if insufficient sample size or spread information were reported. For analysis to be possible, the number experiencing an outcome and total number of participants was required for each arm of the study for binary data; the number, mean score and standard deviation for each arm for continuous data.

b) unstandardised measures: data from rating scales were only included if the scale had previously been described in a peer-reviewed journal. Unpublished instruments are more likely to report statistically significant findings than those that have been peer-reviewed and published (Marshall 2000).

c) high dropout rates: for study outcomes where more than 40% of participants at baseline were reported as lost to follow up, data were excluded from analyses

d) skewed data: To avoid applying parametric tests to non-parametric data, data with a high probability of skew were not meta-analysed. Data were considered to be likely to be skewed for continuous data where the standard deviation multiplied by two was more than the mean (Altman and Bland 1996). Data with a high probability of skew were not meta-analysed but presented individually. For outcomes where data from some studies were not skewed, these data were analysed and results presented.

A summary of findings from studies which were eligible for meta-analyses but did not provide usable data was reported in the results (Table 3.6).

Where possible, endpoint data were analysed. If both endpoint and change data were available, only the former were presented. Where intention-to-treat data were not provided by studies, data from completers were used in analyses as reported. For binary outcomes the random effects odds ratio (OR) and its 95% confidence interval (CI) were calculated. For continuous outcomes, in analyses where all studies used the same outcome measure, a weighted mean difference (WMD) random effects model with 95% confidence intervals was calculated. Where studies used different outcome measures, a standardised mean difference (SMD) random effects model with 95% confidence intervals was calculated.

Investigation of heterogeneity

The presence of clinical and methodological heterogeneity among studies included in this review (due to variation in the services being studied and study design) was likely. Random effects meta-analyses were therefore conducted to incorporate heterogeneity among trials in analyses.

Statistical heterogeneity between studies in analyses was investigated using Chi^2 and I^2 tests. Guidelines of the Cochrane Collaboration Handbook for Reviewers (Cochrane Collaboration 2006) were used as a basis to consider subgroup analyses of significantly heterogeneous studies. (Chi^2 test p value of <0.10 or I^2 test value of >50% indicate substantial heterogeneity.)

Sensitivity analyses

The following sensitivity analyses were carried out for all outcomes where possible.

i) randomisation: only randomised controlled trials were included in analyses.

ii) higher quality studies: only studies assessed as high quality were included in analyses.

iii) Skewed data: For outcomes where skewed data had been excluded and only non-skewed data presented in the main analysis, skewed data were included in sensitivity analyses.

3.4 Results

3.4.1 Study Selection

Reliability of selection: In the test of reliability of inclusion of studies, requiring a second rater to screen a sample of 20 studies for inclusion or exclusion, initial agreement with the first rater was obtained for 19 (95%) of studies. The remaining study had been included by the first rater but not the second, but inclusion was agreed following discussion.

Study inclusion: Twenty seven studies were identified for inclusion in the review. Eighteen of these did not meet the quality criteria for inclusion in meta-analyses. Of nine eligible studies, six provided no usable data, leaving three studies included in meta-analyses. The results of the electronic literature search are presented in Table 3.2.

Table 3.2: Alternatives review: electronic literature search

Database	**Search engine**	**Date of search**	**Articles generated**	**Whole article retrieved** *(additional studies identified by 2nd researcher)*
Medline	Pubmed www.ncbi.nlm.nih.gov	22.02.08	1854	68 (7)
Cinahl	EBSCOhost	22.02.08	403	12 (0)
Embase	WebSPIRS 5.12	22.02.08	1111	24 (3)
PsycInfo	WebSPIRS 5.12	22.02.08	1458	40 (10)
Cochrane (CENTRAL and DARE)	Wiley InterScience	22.02.08	139	10 (3)
Web of Science	ISI Web of Knowledge v.4.1	22.02.08	2679	32 (9)
NHS Economic Evaluation Database	Centre for Reviews and Dissemination www.crd.york.ac.uk/CRDweb	26.02.08	146	3 (0)
Database Total			7790	189
Directory of Published Proceedings	British Library Online Catalogue	29.06.07	397	0
Graylit Network	www.osti.gov/graylit	09.07.07	1082	1(0)
Grey literature Total			1479	1

Four additional studies were included in the review following handsearching of reference lists from studies identified through the electronic search. No additional studies were yielded from consultation with experts. The study flow of assessment for inclusion in the review is summarised in Figure 3.1 below.

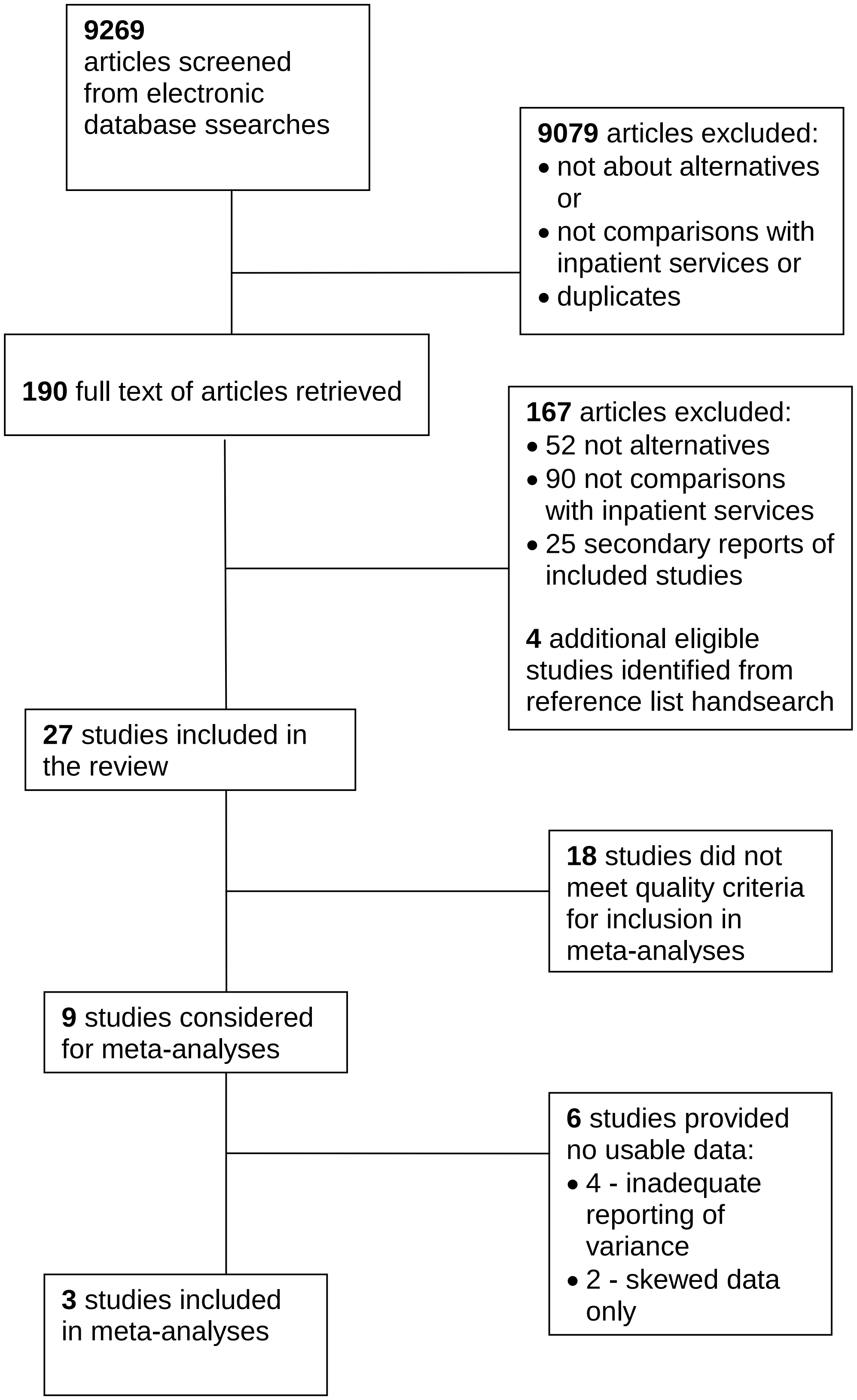


Figure 3.1: Selection of studies: alternatives review

3.4.2 Types of service

The review identified studies of community-based and time-limited services and services with a specific therapeutic model. No studies of inpatient services for specific diagnostic or socio-demographic groups were found, although some studies of community-based services also included clinical or socio-demographic inclusion criteria for participants.

Community-based services

Fifteen studies of community-based services published from 1969-2006 were identified. Eleven of these were of American services. A range of non-hospital service models has been evaluated. Four studies concern Soteria houses, described in Section 2.2.2. Other crisis hostels described in studies in this review also provide care in small, homely settings, typically about eight to twelve-bedded. Services varied in closeness of links with statutory services. Boardman and colleagues (1999) described a residential unit embedded within a Community Mental Health Resource Centre, staffed by clinically qualified mental health professionals including supervision by psychiatrists; Timko and colleagues (2006) by contrast studied independently run services contracted by statutory agencies. Two studies in this review (Timko et al. 2006; Hawthorne et al. 2005) described services funded by American Veterans Associations, providing care mainly to veterans. Apart from Soteria hostels, crisis houses in studies in this review were not reported as guided by a clear, manualised model of care: types of care provided were not described in detail. Placement with individual families was the other model of community-based alternative to hospital identified in this review. A study of family sponsor homes (described in Section 2.2.1) was included (Polak and Kirby 1976).

Time-limited services: Seven studies evaluated time-limited hospital inpatient services ranging from one day to eight days maximum planned admission. Studies dated from 1966-1996 and evaluated services in North America and the UK. All the brief-stay wards in studies in this review were located within larger hospital complexes and accepted general acute admissions. Studies of brief stay services reported a service aim to increase the intensity with which care was provided compared to standard acute wards (e.g. providing assessment, medication review, help solving psychosocial problems, aftercare planning within the period of a brief admission), but distinctly different interventions were not described.

Services with a distinctive therapeutic model: Five studies of services with a specific therapeutic model involved one of two UK nursing-led models of care developed in the last decade, the Tidal Model or the Refocusing Model (described in Section 2.3).

The studies included in the review are listed in Table 3.3.

Table 3.3: Alternatives review: included studies

Type of Service	No. of included studies	References
Community-based: non-hospital services	15	**1) Included in meta-analysis** Boardman et al. (1999) Fenton et al. (1998) Hawthorne et al. (2005) **2) Eligible for meta-analysis but no usable data** Mosher et al. (1995) Polak and Kirby (1976) Timko et al. (2006) **3) Not eligible for meta-analysis** Bittle (1986) Brook (1973) Ciompi et al. (1992) Ciompi et al. (1993) Goveia and Tutko (1969) Hawthorne et al. (1999) Mosher and Menn (1978) Rappaport et al. (1987) Readhead et al. (2002)
Time-limited services: maximum planned stay 14 days or fewer	7	**1) Eligible for meta-analysis but no usable data** Herz et al. (1975) Hirsch et al. (1979) Olfson et al. (1990) **2) Not eligible for meta analysis** Ianzito et al. (1978) Mendel (1966) Schneider and Ross (1996) Voineskos et al. (1974)
Services dedicated to a specific clinical group	0	
Services dedicated to a specific socio-demographic group	0	
Services with a specific therapeutic model	5	**1) Not eligible for meta-analysis** Dodds and Bowles (2001) Berger (2006) Gordon et al. (2005) Lafferty and Davidson (2006) Stevenson et al. (2002)

3.4.3 Study characteristics

The twenty seven studies included in this review comprised seven randomised controlled trials, seven non-randomised prospective quasi-experimental studies, eight retrospective two-group studies and five interrupted time series studies. Most were of small or medium size, only four studies having more than 250 participants (Bittle (2006); Hawthorne et al. (1999); Mendel (1966); Voineskos et al. (1974)). Duration of studies ranged from the period of admission only, up to forty months follow-up. Details of the characteristics and results from all studies are provided in Table 3.4.

Table 3.4: Alternatives review: included studies - characteristics and results

Study Reference	**Description** **Alternative types:** **1 = community-based** **2 = time-limited** **3 = clinically specific** **4 = sociodemographic specific** **5 = therapeutically specific**	**Design**	**Participants** **total n (alternative n /comparison n)**	**Outcomes*** **Outcome domains:** **1 = clinical improvement** **2 = service use** **3 = satisfaction** **4 = costs**	**Results**
1. Studies included in meta-analyses					
Boardman et al. (1999)	CMHC beds, UK	Prospective non-randomised quasi-experiment	Adults admitted to either service but : no acute admissions in last 12 months english-speaking no primary diagnosis other than mental illness n = 177 (110/67)	1.Improvement - GAF, HoNOS, PSE, CAN, HRSD, SBS, LQLP 2. Service use - Bed use and % patients readmitted at 12 month follow up 3. Satisfaction - VSSS 4. cost	3: favours alternative 1,2 & 4: no significant difference
Fenton et al. (1998)	Crisis hostel, Maryland, USA	RCT	Adults assessed as requiring acute admission who are: Voluntary, insured, consenting to participate n = 119 (69/50)	1.improvement: PANSS 2. service use: % patients readmitted at 6 month follow up 3. satisfaction: Unpublished 7 point scale 4. cost	1 & 2 & 3: no significant difference 4. favours alternative
Hawthorne et al. (2005)	6 crisis hostels, San Diego, USA	RCT	Veterans age 18-59 who: Have diagnosis of affective disorder, bi-polar disorder or psychosis, Are voluntary patients, Consent to participate in study n = 99 (52/47)	1.improvement: PANSS, SF-36V 2. service use: no. of readmissions at 2 month follow-up 3. satisfaction: POC 4. cost	1&2&3: no significant difference 4: favours alternative
2. Studies eligible for meta-analyses but with no usable data					
Timko et al. (2006)	Veterans crisis hostels, USA	RCT	Adults assessed as requiring acute admission with: Dual diagnosis No immediate risk to self or others n.b. sample mostly veterans n = 230 (57/173)	2. % patients readmitted over 30 day follow up	2. No significant difference
Mosher et al. (1995) (Soteria study 2)	1. Soteria crisis hostel, California, USA	RCT	Adults age 16-30 requiring acute admission who: have diagnosis of schizophrenia. no more than one previous brief admission are unmarried n = 100 (45/55)	1. 7-point measure of global improvement at 6 week follow-up	1. No significant difference

Study Reference	**Description** **Alternative types:** **1 = community-based** **2 = time-limited** **3 = clinically specific** **4 = sociodemographic specific** **5 = therapeutically specific**	**Design**	**Participants** **total n (alternative n /comparison n)**	**Outcomes*** **Outcome domains:** **1 = clinical improvement** **2 = service use** **3 = satisfaction** **4 = costs**	**Results**
Polak and Kirby (1976)	1. Adult family placement, Colrado, USA	RCT	Adults assessed as requiring acute admission n = 85 (37/38) (10 further patients at crisis hostel excluded for clinical reasons)	1. Goal attainment system, Unspecified community adjustment scale 3. TES 18 month follow-up	1. no significant difference 3. favours alternative
Hirsch et al. (1979)	2. 8-day target hospital admission ward, UK	RCT	Adults age 16+ requiring acute admission No diagnosis of brain injury or major physical health problem n = 224 (115,109)	1. PSE and PBAS at 3 month follow up 2. % patients readmitted and bed use over 1 year follow up.	1&2: no significant differences
Herz et al. (1975)	2. 1-week target hospital admission ward, New York, USA	RCT	Adults age 16+ requiring acute admission who live with a responsible adult, ave a diasgnosis of mental illness. Limitations on co-morbidity n = 175 (51: brief hospital; 61: brief hospital and day care; 63: standard hospital)	1. PSS, GAS 2. no. of patients readmitted over 7 month follow up	1&2: no significant differences
Olfson (1990)	2. 5-day time-limited crisis admission ward, New England USA	Prospective non-randomised quasi-experiment	Adults assessed as requiring acute admission: -diagnosis of schizophrenia -one or more previous admissions -stable housing -no current substance abuse or major medical problems N = 26 (8,18)	1. BPRS and GAS scores at 3 month follow up 2.: bed days over 3 month follow up	1.&2. no significant difference
3. Studies which did not meet quality criteria for inclusion in meta-analyses					
Hawthorne et al. (1999)	1. 5 crisis hostels, San Diego, USA	Prospective non-randomised quasi-experiment	Adults requiring acute admission with diagnosis of depression, psychosis or bi-polar disorder n = 554 (368/186)	1. BASIS-32, SF-36 2. No. of readmissions at 4-month follow-up 3. CSQ	1&2&3: no significant difference

Study Reference	Description **Alternative types:** **1 = community-based** **2 = time-limited** **3 = clinically specific** **4 = sociodemographic specific** **5 = therapeutically specific**	Design	Participants **total n (alternative n /comparison n)**	Outcomes* **Outcome domains:** **1 = clinical improvement** **2 = service use** **3 = satisfaction** **4 = costs**	Results
Goveia and Tutko (1969)	1. Crisis hostel, California USA	Prospective non-randomised quasi-experiment (some but not all subjects randomised)	Adults assessed as requiring acute admission who are: compliant with treatment, not very acutely ill, ill due to reaction to environmental stressors, consenting to participate n = 98 (62,36)	1. MMPI, Rorschach, employment status at discharge 2. Length of initial stay, % patients readmitted in 1 year follow up 3. Semantic Differential Test	1. employment status favours alternative 2. Length of initial stay briefer for alternative No significant difference for other measures
Mosher and Menn(1978) (Soteria study 1)	1. Soteria crisis hostel, California, USA	Prospective non-randomised (pseudo-randomised) quasi-experiment	Adults age 16-30 requiring acute admission who: have diagnosis of schizophrenia. no more than one previous brief admission are unmarried n = 79 (37/42)	1. IMPS, untitled short scale (Venables and O'Connor), Work status 2. no. of patients readmitted at 2 year follow-up	1&2: no significant difference
Ciompi et al. (1993)	1. Soteria crisis hostel, Switzerland	Prospective non-randomised quasi-experiment	Adults age 17-35 Recent onset (1 year) of DSMIII diagnosis of schizophrenia or similar Acutely ill Not drug or alcohol dependent Compliant with treatment n= 44(22/22)	1. BPRS,Housing status Job status,Composite global measure 2. % patients readmitted 4. cost over 2 year follow-up	1&2: no significant difference 4: favours comparison service
Bittle et al. (1986)	1. 2 crisis hostels, Illinois, USA	Retrospective non-randomised cohort study	Adults requiring acute admission: Exclusion criteria re previous admissions, high risk, co-morbidity n = 4305 (594/3711)	2. no. of readmissions over 40 month study period	No significant difference
Rappaport et al. (1987)	1. 45 bed crisis hostel, California, USA	Retrospective non-randomised cohort study	Adults assessed as requiring acute admission n = 203 (134,69) (clinically similar groups drawn from larger cohort)	1. PEF score at discharge	1. Favours comparison service

Study Reference	**Description** **Alternative types:** **1 = community-based** **2 = time-limited** **3 = clinically specific** **4 = sociodemographic specific** **5 = therapeutically specific**	**Design**	**Participants** **total n (alternative n /comparison n)**	**Outcomes*** **Outcome domains:** **1 = clinical improvement** **2 = service use** **3 = satisfaction** **4 = costs**	**Results**
Brook (1973)	1. Crisis hostel, Denver, USA (time-limited to 7 days)	Non-randomised cohort study (not specified if retrospective)	All adults requiring acute admission n = 98 (49/49)	1. unspecified measure 2. No. of readmissions at 6 month follow up	1. favours comparison service 2. favours alternative
Ciompi et al. (1992)	1. Soteria crisis hostel, Switzerland	non-randomised quasi-experiment (not stated if retrospective)	Adults age 17-35 Recent onset (1 year) of DSMIII diagnosis of schizophrenia or similar Acutely ill Not drug or alcohol dependent Compliant with treatment n = 28 (14/14): unclear whether these form part of larger cohort subsequently reported [23]	1. BPRS, Housing status, Job status, Composite global measure (all at 6 week follow up) 4. cost of index admssion	1 & 4: no significant difference
Readhead et al. (2002)	1. Adult family placements, UK	Interrupted time series study	Adults age 18-64 assessed as requiring acute admission with: No immediate high risk to self or others No need for treatment change n not stated	2. bed use 4. cost over 1 year follow up compared to previous years	2. favours alternative
Mendel (1966)	2. 7-day time limited admission ward, California USA	Prospective non-randomised quasi-experiment	Adults 18+ assessed as requiring acute admission: -diagnosis of schizophrenia -admission from community -voluntary or on 72-hr Section n = 443 (114/329)	1. social functioning (measure not specified); % patients financially self-sufficient at 18-month follow up 2. %patients discharged to community from ward within planned time; % patients readmitted within 18 month follow-up	1. Favours alternative (measure of functioning) 2. No significant difference
Schneider and Ross (1996)	2. 3 day crisis admission ward, Connecticut, USA	Retrospective non-randomised cohort	Adults assessed as requiring acute admission n = 1370 (590/780)	2. no. of patients readmitted over 30-day follow-up	2: No significant difference

Study Reference	**Description** **Alternative types:** **1 = community-based** **2 = time-limited** **3 = clinically specific** **4 = sociodemographic specific** **5 = therapeutically specific**	**Design**	**Participants** **total n (alternative n /comparison n)**	**Outcomes*** **Outcome domains:** **1 = clinical improvement** **2 = service use** **3 = satisfaction** **4 = costs**	**Results**
Voineskos et al. (1972)	2. 3 day crisis admission ward, Canada	Retrospective non-randomised cohort	Adults assessed as requiring acute admission n = 868 (439/429)	2. length of initial admission	1. Favours alternative
Ianzito et al. (1978)	2. 24 hour admission ward, Massachusetts, USA	Retrospective non-randomised cohort study	Adults assessed as requiring acute admission n = 184 (83/101) (also 193 non-admitted patients evaluated)	1. unspecified measure of global improvement 3. unspecified measure of compliance All at 2 week follow-up	1&3: no significant differences
Gordon et al. (2005)	5. Tidal model ward, Birmingham, UK	Retrospective non-randomised cohort study	Adults on an acute ward n not stated (service level data only collected)	1/3. composite of number of untoward incidents 1 year pre and post introduction of Tidal Model	1&3. Favours alternative
Stevenson et al. (2002)	5. Tidal Model ward, Newcastle, UK	Interrupted time series study	Adults on an acute ward n = 150 (81/69)	1. rates of self-harm, suicide or violence during admission, compared over 6 month pre and post Tidal Model introduction	1. Favours alternative
Berger et al. (2006)	5. Tidal model ward, Canada	Interrupted time series study	Adults on an acute admission ward who consent to participate n = 46 (not stated)	3. POC completed during admission over 6 month study period, compared to previous 6 months	1. Favours alternative
Dodds and Bowles (2001)	5. Refocusing model ward, UK	Interrupted time series study	Adults on an acute ward n not stated (service level data only collected)	1. rates of self-harm, suicide or violence during admission, compared over 6 month pre and post Refocusing Model introduction	1. Favours alternative
Lafferty and Davidson (2006)	5. Tidal Model ward, Glasgow, UK	Interrupted time series study	Adults on an acute ward n not stated (service level data only collected)	1/3. number of untoward incidents 1 year pre and post introduction of Tidal Model (10 indicators measured)	1&3. Favours alternative

* Acronyms of outcome measures are described in the list of Abbreviations and references are provided in Appendix 1

No studies identified by this review were assessed as high quality. Nine studies were rated as moderate quality, including two well-designed quasi-experimental studies in addition to seven randomised controlled trials. Three of these studies were of brief-stay wards (Olfson et al. (1990); Hirsch et al. (1979); Herz et al. (1975)) and six of community-based services, residential crisis beds (Timko et al. (2006); Hawthorne et al. (2005); Boardman et al. (1999); Fenton et al. (1998); Mosher et al. (1995)) or family placement (Polak and Kirby 1976). None rated strong on all criteria of the Thomas assessment tool (Thomas 2003). Only one study (Fenton et al. 1998) clearly described allocation concealment procedures during randomisation. Where participants were lost to follow up, no studies based analysis on intention-to-treat; all provided completer data only. Reported levels of loss of participants during the course of studies varied substantially for studies of moderate quality. Only two (Polak and Kirby (1976); Herz et al. (1975)) included in outcomes data all potential subjects assessed as eligible. Reported overall attrition rates on individual outcomes in other moderate quality studies range from 3 - 52% of potential participants. Service use data, gathered from routinely collected records, was generally more comprehensive than assessment of patients' functioning or satisfaction.

Of the eighteen remaining studies assessed as low quality, fifteen did not adequately measure or adjust for confounders. Five had high risk of selection bias (more than 40% of those eligible declining to participate); two had unacceptably high withdrawal rates (more than 40%) for all outcomes and one used only unpublished outcome measures.

Only two of 27 studies measured and fully reported the care provided at alternatives and standard services (Olfson et al. 1990, Timko et al. 2006). There is therefore limited information available about the content of care in alternative services. This hampers consideration of the results from this review, as knowledge about the content of service interventions is necessary to understand differences in service outcomes (Mechanic 1996).

Full details of the quality assessment of all studies are provided in Table 3.5.

Table 3.5 Alternatives review: included studies - quality assessment

Study Reference	**Selection Bias**	**Allocation Bias**	**Confounders**	**Blinding**	**Data collection**	**Dropouts**	**Analysis** SS (was a sample size calculation made) ITT (was analysis based on intention-to-treat)	**Intervention Integrity** E (did at least 80% of participants receive the intervention) C: (was there consistency of intervention)	**Content of care measurement?** 0 = none 1 = partially, alternative only 2 = partially, both 3 = fully, alternative only 4 = fully, both
	Rating: W = weak, M = moderate, S = strong								
1. Studies which met quality criteria for inclusion in meta-analyses									
Timko et al. (2006)	M	S	S	W	S	S	SS = No ITT = No	E: yes C: yes	4
Fenton et al. (1998)	M	S	S	W	S	S	SS = No ITT = No	E: yes C: not measured	0
Hawthorne et al. (2005)	M	S	S	W	S	S	SS = No ITT = No	E: yes C: not measured	0
Mosher et al. (1995)	W	S	S	W	W	W	SS = No ITT = No	E: No C: No (medication use)	2 (medication use)
Polak and Kirby (1976)	S	S	W	W	S	M	SS = No ITT = No	E: No C: not measured	0
Boardman et al. (1999)	M	M	S	W	S	S	SS = No ITT = No	E: No C: not measured	0

Study Reference	**Selection Bias**	**Allocation Bias**	**Confounders**	**Blinding**	**Data collection**	**Dropouts**	**Analysis** SS (was a sample size calculation made) ITT (was analysis based on intention-to-treat)	**Intervention Integrity** E (did at least 80% of participants receive the intervention) C: (was there consistency of intervention)	**Content of care measurement?** 0 = none 1 = partially, alternative only 2 = partially, both 3 = fully, alternative only 4 = fully, both
	Rating: W = weak, M = moderate, S = strong								
Hirsch et al. (1979)	S	S	W	W	S	W	SS = No ITT = No	E = yes C = not measured	0
Herz et al. (1975)	M	S	S	W	S	W	SS = No ITT = No	E: yes C: not reported	4 (but results briefly reported)
Olfson (1990)	M	M	S	W	S	S	SS = No ITT = n/a (no dropouts)	E = yes but 62.5% of experimental group also received control intervention C = yes	4
2. Studies which did not meet quality criteria for inclusion in meta-analyses									
Hawthorne et al. (1999)	W	M	S	W	S	W	SS = No ITT = No	E: yes C: not measured	0
Goveia and Tutko (1969)	W	M	W	W	S	W	SS = No ITT = No	E: yes C: not measured	0
Mosher and Menn(1978)	W	M	S	W	S	M	SS = No ITT = No	E: yes C: No (medication use)	2 (medication use)
Ciompi et al. (1993)	W	M	W	W	S	S	SS = No ITT = N/a (no dropouts)	E: yes C: Not reported	2 (medication use)

Study Reference	**Selection Bias**	**Allocation Bias**	**Confounders**	**Blinding**	**Data collection**	**Dropouts**	**Analysis** SS (was a sample size calculation made) ITT (was analysis based on intention-to-treat)	**Intervention Integrity** E (did at least 80% of participants receive the intervention) C: (was there consistency of intervention)	**Content of care measurement?** 0 = none 1 = partially, alternative only 2 = partially, both 3 = fully, alternative only 4 = fully, both
	Rating: W = weak, M = moderate, S = strong								
Bittle et al. (1986)	M	M	W	W	S	S	SS = No ITT = n/a (no dropouts)	E: yes C:not measured	0
Rappaport et al. (1987)	M	M	S	W	S	W	SS = No ITT = No	E: Yes C: Not measured	2 (medication use)
Brook (1973)	M	M	W	W	S	S	SS = No ITT = n/a (no dropouts)	E: yes C: not measured	0
Ciompi et al. (1992)	M	M	W	W	S	S	SS = No ITT = N/a (no dropouts)	E: yes C: Not reported	2 (medication use)
Readhead et al. (2002)	M	W	W	W	S	S	SS = No ITT: n/a (no dropouts)	E: yes C: not measured	0
Mendel (1966)	S	S	W	W	S	W	SS = No ITT = No	E = yes C = Not measured	0
Schneider and Ross (1996)	S	M	W	W	S	M	SS = No ITT = No	E: yes but 31% of experimental group also received control intervention C: not measured	0

Study Reference	**Selection Bias**	**Allocation Bias**	**Confounders**	**Blinding**	**Data collection**	**Dropouts**	**Analysis** SS (was a sample size calculation made) ITT (was analysis based on intention-to-treat)	**Intervention Integrity** E (did at least 80% of participants receive the intervention) C: (was there consistency of intervention)	**Content of care measurement?** 0 = none 1 = partially, alternative only 2 = partially, both 3 = fully, alternative only 4 = fully, both
	Rating: W = weak, M = moderate, S = strong								
Voineskos et al. (1972)	S	M	W	W	S	S	SS = No ITT: n/a (no dropouts)	E: yes but 46% of experimental group also received control intervention C: not measured	0
Ianzito et al. (1978)	S	M	W	W	W	S	SS = No ITT: n/a (no dropouts)	E: yes but 46% of experimental group also received control intervention C: not measured	0
Gordon et al. (2005)	S	M	W	W	M	S	SS = No ITT: n/a – service level outcomes only	E: yes C: not measured	0
Stevenson et al. (2002)	S	W	W	W	S	S	SS = No ITT: n/a (no dropouts)	E: yes C: yes	2: initial assessment and vebatim quotes in care plans
Berger et al. (2006)	W	W	W	W	S	W	SS = No ITT = No	E: yes C:No (individualised care plan)	1: % patients receiving an individualised care plan
Dodds and Bowles (2001)	S	W	W	W	S	S	SS = No ITT: n/a (no dropouts)	E: yes C: not measured	0
Lafferty and Davidson (2006)	S	W	W	W	M	S	SS = No ITT: n/a – service level outcomes only	E: yes C: not measured	0

3.4.4 Inclusion in meta-analyses

9 studies were selected for inclusion in the meta-analysis. 7 randomised controlled trials were included; of 20 non-randomised trials, 2 met the quality criteria outlined in Section 3.3.4 and were also included.

Of the 9 studies eligible for inclusion in the analysis, only 3 presented usable data. Four studies provided inadequate reporting of variance. Two included only skewed data. All three studies with usable data for meta-analysis concerned community-based alternatives. The outcomes from each study, usable and unusable for the analysis, are presented in Table 3.6 below.

Table 3.6 Alternatives review: meta-analysis - usable data from eligible studies (n=9)

Study	Usable outcomes	Unusable outcomes
Boardman et al. (1999) Haycox et al. (1999) provide costs and service use data	**Medium term** 1. LQL 12 month follow up 2. Readmission in 12 month follow up 3. VSS 12 month follow up	**Short term** 2. Length of index admission (no mean or s.d.) 4. Cost of index admission (no s.d.) **Medium term** 1. GAF, HSRD, PSE, HoNOS, CAN, SBS 12 month follow up (no n for individual arms) 2. Bed use 12 month follow up (no mean or s.d.) 4. Costs over 12 month follow up (no s.d.)
Fenton et al. (1998) Fenton et al. (2002) provide costs data	**Short term** 1. PANSS score, Discharged to the community **Medium Term** 1. employed at 6-month follow up 2. Days in hospital during 6 month follow up, readmitted during 6 month follow up, No. of readmissions during 6 month follow up	**Short term** 2. length of index admission (data skewed) 3.unpublished measure 4. Costs of index admission (data skewed) **Medium term** 1. PANSS score at 6 month follow up: no n for each arm 4. Costs at 6 month follow up (data skewed) Homeless at follow-up, arrested during study period, number of social contacts (not outcomes in this review)
Hawthorne et al. (2005)	**Short term** 1. PANSS, 3. POC **Medium Term** 1.PANSS, SF-36V(MCS) 2 month follow up	**Short term** 1. SF-36V(MCS) (data skewed) 2. Length of index admission (data skewed) 4. Costs of index episode (data skewed) **Medium term** 2. Readmissions over 2 month follow up (no n for individual arms: data given for number of participants on each arm admitted to alternative and hospital, but possibility that this includes double counting) Dug and alcohol use – ASI 2 month follow up (not an outcome included in this review) Homelessness at 2 month follow up (not an outcome in this review)
Mosher et al. (1995)	None	**Short term** 1.1 measure of clinical improvement (Mosher et al. 1971) (no s.d.) Bola and Mosher (2003) provide 2 year outcome data for a combined cohort of participants in the two Soteria studies identified in this review, but no separate data from each study
Polak and Kirby (1976)	None	**Short term** 1. Goal Attainment System (no s.d.), SDS, (no s.d.) Communiy Adjustment Scale (unspecified measure) 3. TES (no s.d.) **Medium term** 1. Goal Attainment System, SDS(no s.d.), Community Adjustment Scale (unspecified measure) 3. TES (no s.d.) 4 month follow up

Study	Usable outcomes	Unusable outcomes
Timko et al. (2006)		**Short term** 2. length of index admission (skewed data) **Medium term** 1. ASI psychiatric subscale at 1 year follow up (skewed data) 2. Number of inpatient days at 1 year follow up (no mean or s.d. for overall figure) 4. Costs over 1 year follow up (skewed data) Outpatient service use over I year follow up (not a review outcome) Drug and alcohol use - ASI total score (not a review outcome)
Herz et al. (1975) Herz et al. (1977) provide 2 year follow up data	None	**Short term** 2. Length of index admission (no.s.d.) **Medium term** 1. GAS, PSS,MSER at 8 week and 2 year follow up (no s.d.), employed at 6 month and 2 year follow up (data only given for "patients who ordinarily would have been expected to work": no n provided) 2. Inpatient bed use at 3 month and 2 year follow up (no s.d.), Number of patients readmitted at 8 week and 2 year follow up (Unclear graph only: no n for each arm) Study has 3 arms: 3rd arm (day hospital + alternative residential excluded from this review)
Hirsch et al. (1979)	None	**Short term** 2. Length of index admission (no s.d.) **Medium term** 1. PBAS (not published measure), PSE (no s.d.) at 3 month follow up, number readmitted within 1 year follow up (more than 40% lost to follow up) 4. Costs at 3 month follow up (no data provided)
Olfson et al. (1990)	None	**Medium term** 1. BPRS, GAS 3 month follow-up (no s.d.) 2. Inpatient bed days within 3 month follow up (skewed data)

3.4.5 Results of meta-analyses

Analyses of data for ten outcomes were conducted. Six outcomes concerned improvement in clinical or social functioning, two concerned service use and two satisfaction. No meta-analyses of cost data were possible. All analyses included either one or two studies, with participant numbers of between 80 and 288.

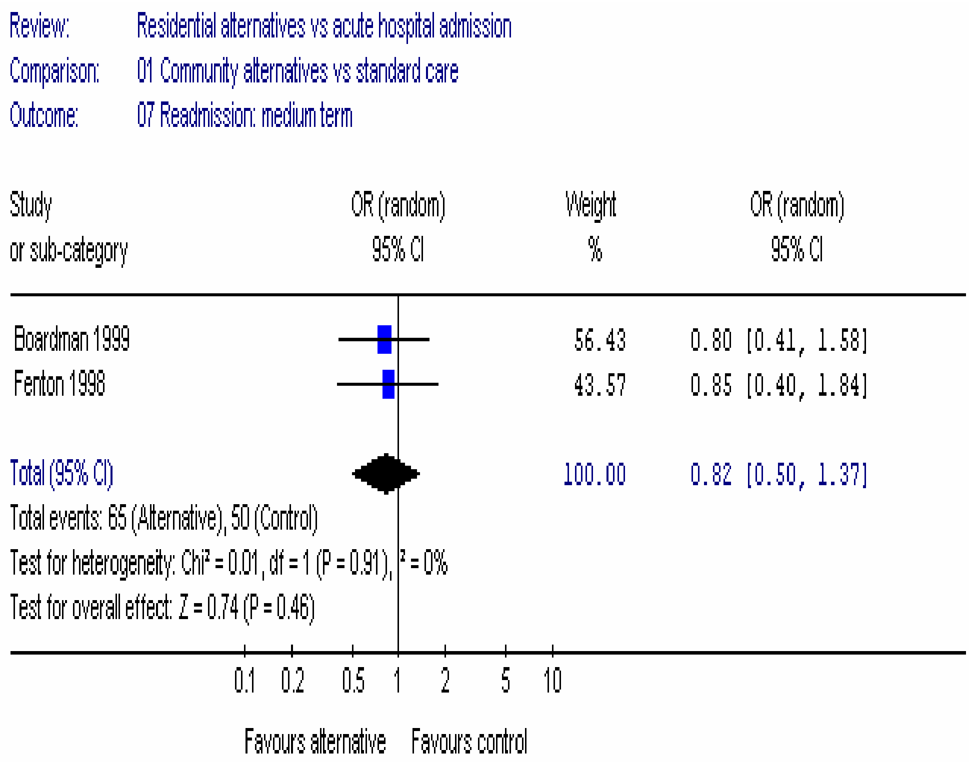
Outcomes relating to improvement which were analysed were levels of symptom severity and functioning, quality of life, independent living and employment. No significant differences were found between alternative and standard services, although all outcomes bar one (living independently – short term) relating to improvement showed non-significantly better outcomes from the alternative service. Two analyses of service use outcomes – readmission (medium term) and inpatient bed use (medium term) - also showed no significant difference between types of service. Satisfaction with services was found to be significantly greater at alternatives than standard services both in the short term (p=0.04) and medium term (p=0.02). Table 3.7 provides full results of the meta-analyses undertaken. The forest plot for one outcome (readmission medium term: of the analyses in this review, this involved the largest number of participants) is presented as an example in Figure 3.2. All the meta-analyses summarised in Table 3.7 are provided in Appendix 2, as are skewed data for outcomes from eligible studies where meta-analysis was not possible.

Table 3.7: Alternatives review: summary of meta-analyses

Comparison: Community-based alternatives vs. standard inpatient services				
Outcome	**Participants (studies)**	**Statistical method**	**Effect estimate [confidence intervals]**	**Test for overall effect**
Clinical improvement (symptoms) short term	218 (2)	WMD (95% CI)	*(negative score favours alternative)* -0.20 [-5.63, 5.24]	not significant Z = 0.07 p = 0.94
Clinical improvement (symptoms) medium term*	80 (1)	WMD (95%CI)	*(negative score favours alternative)* -3.70 [-11.08, 3.68]	not significant Z = 0.98 p = 0.33
Improvement (emotional functioning) medium term	80 (1)	WMD (95% CI)	*(>1 favours alternative)* 5.30 [-1.08, 11.68]	not significant Z = 1.63 p = 0.10
Quality of life medium term	145 (1)	OR (95%CI)	*(>1 favours alternative)* 1.03 [0.51, 2.05]	not significant Z = 0.07 p = 0.94
Employment: medium term	112 (1)	OR (95% CI)	*(>1 favours alternative)* 1.53 [0.59,3.97]	not significant Z = 0.87 p = 0.38
Living independently (short term)	119 (1)	OR (95% CI)	*(>1 favours alternative)* 0.28 [0.06, 1.35]	not significant Z = 1.59 p = 0.11
Inpatient bed-days (medium term)	119 (1)	WMD (95% CI)	*(positive score favours alternative)* -5.00 [-21.53, 11.53]	not significant Z = 0.59 p = 0.55
Readmission (medium term)	288 (2)	OR (95% CI)	*(<1 favours alternative)* 0.82 [0.50, 1.37]	not significant Z = 0.75 p = 0.46
satisfaction (short term)	99 (1)	WMD (95% CI)	*(positive score favours alternative)* 9.20 [0.25, 18.15]	favours alternative Z = 2.01 p = 0.04
satisfaction (medium term)	145 (1)	OR (95% CI)	*(>1 favours alternative)* 2.47 [1.12, 5.43]	favours alternative Z = 2.25 p = 0.02

* excluding skewed data. WMD = weighted mean difference OR = odds ratio

Figure 3.2 Community-based alternatives vs standard services: Readmission – Medium Term



3.4.6 Results from data not usable in meta-analyses

Data from eligible studies of community-based services which were unusable for meta-analyses broadly concur with results of analyses. Three eligible studies of community-based services which provided no data for analyses (Timko et al. 2006, Mosher et al. 1995, Polak and Kirby 1976) found no significant difference from standard services in clinical improvement and one which evaluated service-user satisfaction (Polak and Kirby 1976) favoured the alternative. One study (Fenton et al. 1998) reported a significantly longer duration of index admission at the alternative service; service use data from three other studies of community-based services (Boardman et al. 1999, Hawthorne et al. 2005, Timko et al. 2006) found no significant differences. No cost data were usable in meta-analyses but two studies found cost of index admission significantly less at the alternative service (Fenton et al. 1998, Hawthorne et al. 2005); two found no significant overall public cost differences over the study follow-up period (Boardman et al. 1999, Fenton et al. 1998)

Of the three moderate quality studies of time-limited alternatives, only one found that the brief-stay ward was able to discharge patients within the planned admission period and significantly more quickly than standard wards (Herz et al. 1975). One study (Olfson et al. 1990) found a majority of patients required transfer to a standard ward at the end of the brief-stay period; one study (Hirsch et al. 1979) found the eight-day planned admission period was not rigorously adhered to and that mean length of stay was not significantly shorter than standard care. None found any significant differences in clinical outcomes or readmission rates.

No studies of services using a distinctive therapeutic model were eligible for meta-analyses, all five being of low quality. All assessed multiple outcomes and reported some favourable to alternative services regarding patient satisfaction or levels of untoward incidents on wards.

3.4.7 Sensitivity Analyses

Randomisation: One outcome involved data from a randomised study (Fenton et al. 1998) and a non-randomised study (Boardman et al. 1999): readmission (medium term). Results excluding the non-randomised study (OR = 0.85, C.I. = 0.40 – 1.84) were very similar to results from combining both studies (OR = 0.82, C.I. = 0.50 – 1.37). Statistical heterogeneity between the two studies in this analysis was very low (I^2 value less than 1%).

Skewed data: One outcome involved studies with skewed (Timko et al. 2006) and non-skewed data (Hawthorne et al. 2005): clinical improvement medium term. Results do not differ significantly if skewed data are included (SMD = -0.14, C.I. = -0.39 to +0.10) or excluded (SMD = -0.22, C.I. = -0.66 to +0.22). Statistical heterogeneity between the two studies is very low (I^2 less than 1%), confirming consistency in results from the two studies.

No studies included in this review met the study criteria for high quality and no analyses identified notable heterogeneity between studies, so no further sensitivity analyses were conducted.

3.5 Discussion

In summary, the findings of this review were:

i) No studies of inpatient socio-demographic or diagnosis-specific services were identified in this review.

ii) No studies concerning services with a distinct therapeutic model were of sufficient quality to be included in meta-analyses.

iii) No studies of time-limited services provided data which could be included in meta-analyses. In two of three studies of moderate quality, a majority of patients could not be discharged home within the planned admission period.

iv) Only three, small or medium sized studies of community-based alternatives provided analysable data for a number of outcomes. Evidence from this review is therefore too limited to be conclusive.

v) The review found preliminary evidence that satisfaction with services was greater at community alternatives compared to standard inpatient wards in both short and medium term. No significant effect was identified in meta-analyses regarding clinical improvement, readmission or costs.

vi) Only three of the studies included in this review sought to measure the content of care provided at alternative and comparison services. In one of these, results regarding content of care were not fully reported. Information about the care provided in alternatives and how this may differ from standard care is therefore limited.

3.5.1 Strengths and weaknesses of the review

Two sources have informed the method and structure of this review:

i) the checklist proposed by the Quorom Group (Moher et al. 1999), a group convened to address standards of reporting of meta-analyses of clinical randomised controlled trials.

ii) the Cochrane Collaboration Handbook for Reviewers (Cochrane Collaboration 2006), a guide to writing a systematic review provided by Cochrane, a major source of systematic reviews in healthcare.

The breadth of the focus and the inclusion criteria in this review creates strengths and weaknesses.

Strengths

Two advantages of the broad approach adopted are: a) studies of all innovative acute residential services of interest are likely to be included in the review; and b) all available evidence is considered.

The lack of consistent terminology to describe residential alternatives posed difficulties in identifying all relevant studies. Sensitivity was therefore prioritised over specificity and broad inclusive search terms employed for the electronic database search. The choice of databases to search electronic and sources for searching grey literature were informed by advice from the Cochrane Schizophrenia Group. A broad, international selection of experts was also consulted to try to identify studies missed in the electronic and hand searches. Three things suggest that a comprehensive search was achieved:

i) The small number of additional studies identified through handsearching of reference list or experts
ii) The fact that no studies were identified which appeared to describe residential alternatives but did not meet the review's inclusion criteria
iii) The initial 95% rate of inter-rater agreement in study selection and greater inclusiveness of the main reviewer

The decision to include non-randomised comparison studies in the review ensured all relevant evidence was considered. Two advantages of non-randomised, quasi-experimental studies have been proposed (Gilbody and Whitty 2002):
i) They may provide some evidence about a service or intervention when a randomised controlled trial is not feasible.
ii) They may have strong real-world applicability by evaluating outcomes for cohorts of all service users in a functioning mental health service.
Given the limited available evidence about the effectiveness of alternatives suggested by previous reviews (Alwan et al. 2008, Grawe et al. 2005), the benefits of this inclusive approach outweighed the drawbacks of including poorer quality evidence.

Weaknesses

A weakness of this review regarding search strategy was the failure to record the number of duplicate studies identified from different databases in the electronic search. This means the total number of different studies initially scanned by title for inclusion/exclusion from the review cannot be identified.

The broad inclusion criteria and search strategy used in the review has two drawbacks: a) poor quality evidence may limit or distort the review’s findings; and b) The heterogeneity of services evaluated by studies may undermine the validity of synthesising or meta-analysing results from several studies. The attempts made to address the issues of quality and heterogeneity are discussed below:

a) Quality

Quality assessment of studies included in the review ensured only higher quality studies were included in meta-analysis. It also informed consideration of the strength of available evidence. The use of a published quality assessment tool allowed study quality to be assessed systematically. The Thomas instrument (2003) was used because it was evaluated in a systematic review (Deeks et al. 2003) as one of six “best” quality assessment tools for quantitative studies and as easy to use, able to deal with randomised and non-randomised studies and suitable for use in a systematic review.

Notwithstanding the guidance provided by the author on how to use the tool, some subjective judgements were required by raters (whether participants are adequately representative of the target population for the intervention, what constitute important confounders for this review, what constitutes adequate reliability or validity). Criteria for making these judgements were therefore operationalised, in order to reduce unreliability in the quality rating. The analysis and integrity sections of the tool, where comments rather than a single rating are required, were also distilled to noting the presence or absence of four methodological features in each study: this provided concision and allowed clear comparison of studies.

The multi-dimensional assessment of study quality provided by the Thomas assessment tool then required distillation to yes/no decisions about whether to include studies in meta-analyses. This process was again operationalised in order to make the criteria for including studies in analyses as consistent and systematic as possible. All randomised controlled trials, as the most robust study type, were included in meta-analyses. Non-randomised controlled trials were not automatically excluded but were required to demonstrate no significant difference between study groups for important confounders (severity of illness), in order to allay concern at the most obvious drawback of non-randomised studies, that like may not be compared with like. Non-randomised controlled trials were also required to be rated moderate or above for all quality ratings except blinding. (Blinding assessors successfully in health services research is problematic, as even if blinding is attempted, study participants may disclose which service they have used during assessment.)

These operationalised criteria were designed to exclude studies with a clearly high likelihood of bias. They allowed studies with less than optimal ratings of quality to be included however. They also left some issues affecting study quality unaddressed (e.g. procedures for allocating of participants to study arms in randomised trials, analysis based on intention-to-treat). Sensitivity analyses including only studies of higher quality were planned to address the risk of distorting meta-analyses through inclusion of studies of sub-optimal quality. However, this was not possible due to the lack of higher quality studies identified in the review: none rated strong on all the quality criteria of Thomas (2003) or on the two dimensions of quality identified by Schulz et al. (1995) as most associated with estimates of treatment effects – allocation concealment and analysis based on intention-to-treat. There are therefore caveats about the quality of the data in this review and the confidence which can be put in its results. These are considered in discussion of the review’s findings in Section 3.5.2.

b) Heterogeneity

Different types of alternatives (e.g. community or time-limited services) were analysed separately to ensure some similarity in services evaluated by studies combined in meta-analyses. Short, medium and long term outcomes were also analysed separately to help interpret results. Using a random effects model in meta-analyses accounted for the likelihood of heterogeneity in studies in the calculation of confidence intervals. However, the small number of studies eligible to be included in analyses precluded sub-group analyses or further investigation of possible reasons for heterogeneity or of potential bias through inspection of funnel plots. The small number of studies in analyses limits confidence in how far results can be generalised to all service models within an alternative service type. For instance, if satisfaction is greater at a veterans' crisis house than at standard acute wards, it is not certain that this will also be true for Soteria-style hostels or family sponsor homes.

A degree of heterogeneity in participants and services in both experimental services and standard care is inevitable when reviewing studies from different countries over a long time period. The benefits of addressing a broad question of general interest in the review outweighed its drawbacks. The small number of studies suitable for inclusion in the review and potential heterogeneity of services evaluated prompts further caution in interpreting results.

3.5.2 Interpreting the results

The paucity of evidence concerning the acceptability or effectiveness of alternatives compared to standard care is revealed by this review. No studies of services for specific demographic groups or people with specific diagnoses were identified, despite such services forming part of the current UK acute care system (Johnson et al. 2009) and despite literature providing descriptions of some such services as promising service models, e.g. Mother and Baby Units (Royal College of Psychiatrists 2001) or psychosis, affective disorder and personality disorder wards (Bonsack et al. 2001).

Studies of services with a distinctive therapeutic model were all of low quality, failing to account for differences between groups in analysis. The before and after comparison provided by most studies of services with a distinctive therapeutic model and their lack of stated primary outcomes also increase the risk of reporting and publication biases. The feasibility of brief-stay acute wards is brought into question by the finding that of the three moderate quality studies, in only one (Herz et al. 1975) was the alternative service able to discharge a majority of patients within the planned admission period. The applicability of this finding to contemporary mental health service settings may be limited however, as the moderate quality studies of time-limited services identified in this review all predate the advent of modern community resources such as home treatment teams. The most recent study of time-limited services included in the review (Schneider and Ross 1996), found that 69% of those admitted to a three-day admission ward could be discharged to the community within this period, but the comparability of patients with those admitted to general acute wards was unclear.

Of the five types of alternative identified for inclusion in the review, only studies of community-based alternatives provided any data which could be used in meta-analyses. The evidence presented here about community-based alternatives should also be accepted with some caution for two reasons:

i) Quality and quantity of the studies: results from analyses are based on a small number of studies (although data from higher quality studies of community-based services not usable in meta-analyses support the review’s findings). There is therefore a risk that positive results concerning greater satisfaction at alternatives might represent Type 1 errors or that neutral findings regarding costs and patient outcomes might represent Type 2 errors and fail to reflect real differences between alternatives and standard services in effectiveness or cost-effectiveness. Studies were all of sub-optimal quality. In particular, the study providing data on short-term satisfaction for the meta-analysis (Hawthorne et al. 2005) included some data collected by service staff, introducing an increased risk of social desirability bias (acknowledged by the authors). At the lower end of confidence intervals in analyses of service user satisfaction, the effect sizes found are too small to be clinically important.

ii) Applicability of the evidence: Studies included in the review and the meta-analyses exhibit considerable variation, both in terms of service provided and study population. All three studies in the meta-analyses used different inclusion criteria for participants and each included some criteria (such as veterans only, consenting to participate, no admissions in the previous twelve months) beyond those normally required for real-life acute admission. Two of the three studies included in meta-analyses specifically excluded detained patients (Fenton et al.

1998; Hawthorne et al. 2005), while the other four of moderate quality did not report the number of detained participants. Findings from this review may only be applicable to a sub-group of people requiring acute admission, excluding some of those who are most severely unwell or least cooperative. None of the three studies included in meta-analyses provided detailed information about the content of care provided at alternative and standard services. Therefore uncertainty remains about the population most effectively managed by community alternatives and the extent and nature of difference from standard services in what is provided.

The limited, preliminary evidence yielded by this review is favourable to alternatives however. Community-based alternatives may be a beneficial alternative to standard acute wards for some people requiring acute admission: there is evidence that they are more acceptable to service users and no indication that they are less clinically effective. Other types of alternative – socio-demographic specific, diagnosis-specific, time-limited or with a distinctive therapeutic model – are not contra-indicated by any research and remain to be thoroughly evaluated.

3.5.3 Implications for research

Despite their presence in the UK acute care system (Johnson et al. 2009), all five types of alternative service have an incomplete or absent empirical evidence base. They remain to be thoroughly evaluated. This goes some way towards explaining why some service models first described as promising several decades ago, such as crisis placements in family homes and brief stay admission wards, have yet to be widely adopted despite the wish among service planners, clinicians and service users to develop alternatives to standard acute wards. The conclusion that there is a need and an opportunity for more research is an inescapable one in this area. The studies included in this review evaluate young services or recently established service innovations. Evaluation of more established, enduring alternatives would also be desirable to investigate whether outcomes, perhaps especially satisfaction, are sustainable and not merely a function of service novelty.

Research needs to be of good quality. A minority of relevant studies identified for inclusion in this review were suitable to be included in meta-analyses. Three important requirements can be identified.

i) Studies should state primary outcome(s) in advance. This would reduce reporting bias, where only positive outcomes are published, and Type 1 errors where investigation of multiple outcomes yields some chance positive findings. A majority of studies in this review found some positive outcomes for alternatives. These were often reported prominently, providing an impression that alternatives were effective and acceptable compared to standard care. Meta-analyses in this review, however, revealed a lack of clear evidence for most types of alternative and most outcomes.

ii) When presenting data, numbers in each arm and (for continuous data) means and standard deviation must be provided in studies for data to be usable in meta-analyses.

ii) Key aspects of study quality are desirable which were absent from studies included in this review, such as arranging adequate allocation concealment and conducting analysis based on intention-to-treat - two dimensions of quality identified (Schulz et al. 1995) as most associated with estimates of treatment effects

There are particular challenges to conducting randomised controlled trials in acute mental health settings, where the need for immediate intervention makes both the logistics of randomisation and the process of informing participants and obtaining consent problematic. In such circumstances, quasi-experimental studies may be more feasible and have strong real-world applicability by evaluating outcomes for cohorts that include all service users (Gilbody and Whitty 2002). Two non-randomised, natural experiment studies met the quality criteria for this review and a sensitivity analysis revealed no significant difference in effect size with the inclusion or exclusion of a non-randomised study. This review indicates that a well-designed non-randomised study, which accounts for important confounders, may have a useful place in acute mental health services research. The development of clear protocols for ethically acceptable recruitment in mental health crises would also be very helpful, addressing issues such as how to conduct urgent randomisation out of hours when researchers are not available and how to deal with the often transient loss of capacity experienced by many people at the time of a crisis.

The lack of content of care measurement in most studies in this review is a missed opportunity. Had all studies had measured content of care consistently, there would be a wealth of information about what alternatives do, how this differs from standard care and elements of care which may be associated with positive outcomes. As it is, the nature of what is provided at alternative services remains opaque.

3.5.4 Implications for practice

Current research evidence provides clinicians and commissioners with only very limited guidance about effective models of acute inpatient mental health care. Several service models identified in this review – Soteria houses, adult family placements, time-limited wards - have been developed in more than one country or time period without ever becoming a well-established part of a national acute service system. This suggests some doubt about their sustainability and/or usefulness, although also a persisting perception of a need to seek alternatives to standard acute care. Residential services which can only cater for a proportion of people requiring acute admission may be perceived by service planners and commissioners as a luxury and be vulnerable to losing funding. The failure of alternative service models to endure may also reflect a reliance of innovative services on charismatic leaders and local champions, without whom they may not thrive. The community beds embedded in a Community Mental Health Resource Centre evaluated by Boardman and colleagues (1999) allay some of these concerns. They were able to admit a reasonably high proportion (65%) of people assessed as requiring acute admission during the study period, can accept detained patients and are still running currently, a decade later. Drayton Park, a women’s crisis house in North London, has also been established for more than a decade and evaluated in qualitative studies as providing a valuable role in local acute care (Killaspy et al. 2000; Johnson et al. 2004). This suggests that in a contemporary UK context, community crisis beds can constitute an important and sustainable part of local acute inpatient provision.

The evidence of the dissatisfaction of many service users with standard psychiatric wards (reported in Section 1.1) suggests a need for alternatives. The potential for emergency residential accommodation outside hospital to improve service user choice and thus the acceptability of services, while also relieving bed pressure on acute wards was identified ten years ago (Sainsbury Centre for Mental Health 1998). Even if alternative service models can only divert a sub-group of people requiring acute admission, the increased scope this might bring for focusing appropriate facilities and expertise in inpatient services for a higher risk, predominantly detained client group is potentially useful. Higher quality studies from this review found no evidence against alternative models of care and, consistent with previous qualitative research (Faulkner et al. 2002; Johnson et al. 2004), provides an indication that crisis beds in non-hospital settings may increase satisfaction with acute residential services for users. Certainly, this review provides no discouragement to service managers and commissioners to consider innovation in the provision of acute inpatient care.

3.6 Conclusion

The Alternatives Study (Johnson et al. 2009) found that residential alternatives constitute nearly 10% of adult acute mental health beds in England. There is no information about the prevalence of alternatives elsewhere. Literature suggests, however, that a range of alternatives have been developed and established across Europe and North America over the last 40 years. It is unclear how inpatient alternatives compare with standard acute wards. Evidence suggests community-based alternatives may be viable for many people needing acute admission and be more acceptable to them than standard care. There is no evidence that residential alternatives are unsafe or ineffective: they remain a potentially promising means to address short-comings of standard acute wards and to increase service user choice. Further research about their effectiveness and acceptability is required.

This review also suggests there is very little information available about what care is provided at alternatives and how this differs from standard inpatient care. There is little indication of whether or how care provided at community alternatives may relate to their acceptability to patients. Whether alternatives address the concerns of service users, reported in Chapter 1, for more things to do and greater availability of staff in inpatient services, is also unclear. This review suggests a need for comparison of the care provided at alternatives and standard services and how the content of care may relate to service outcomes.

Chapter 4
Assessing the content of care in mental health services: a review of measures

Given the prevalence and promise of alternatives and the uncertainty about what care is provided at different types of residential acute service (Chapters 2 and 3), there is a need to describe, quantify and compare the care provided in alternatives and standard services. This chapter reviews measures of content of care which have been used in mental health services. The context and rationale for assessing content of care, outlined in Section 1.5, is more fully considered. Existing measures of content of care for use in mental health service settings are identified and their methodological features and how they have been used to examine associations with outcomes are described. Their strengths and weaknesses and fitness for use in acute inpatient settings are discussed.

4.1 Reasons to measure the content of care in mental health services

It was proposed in the mid 1990s that mental health services research should include more emphasis on describing and measuring what services provide. Mechanic, for example, argued that detailed investigation of the content of service interventions is vital to help understand differences in outcomes between services (Mechanic 1996). Burns and Priebe drew attention to the variation in practice amongst even purportedly similar mental health services (Burns and Priebe 1996). They also identified a lack of complete or consistent approaches to describing mental health services. Johnson and Salvador-Carulla (1998) reviewed methods of describing and classifying mental health services. They noted the lack of consensus about how to describe mental health services and the paucity of established valid and reliable instruments with which to measure service content.

In recent years however, much mental health research has focused on evaluating the effectiveness of emerging service models, such as Assertive Community Treatment and Crisis Resolution Teams. This has perhaps deflected attention from refining methods of measuring service content. A review of measures of the content of care in mental health services has potential to clarify what is known and gaps in current knowledge about how content of care can best be measured in mental health service settings.

4.1.1 Content of care measurement in context

In this section, content of care is put into context within mental health services. Two broad organising frameworks are considered, followed by three frameworks which distinguish elements of the process of service delivery, including content of care.

First, Donabedian identified three types of information for evaluating the quality of health services: structure, process and outcome (Donabedian 1992). Donabedian defines **Structure** as physical and organisational properties of the settings in which care is provided; **Process** as what is done for the patients; **Outcome** as what is accomplished for the patients (Donabedian 1992).

Second, Thornicroft and Tansella propose a framework, “The Mental Health Matrix”, which adds the geographical dimension of **patient**, **local** or **country** levels to the temporal dimension of **input**, **process** and **outcome** (Thornicroft & Tansella 1999). This creates a 3 x 3 framework with nine cells to help formulate mental health service aims and practice. Process research at the patient level is required to investigate what happens in contacts between mental health staff and patients and the content of service interventions.

Three organising structures have been proposed to identify important elements of process measurement in health care by Donabedian (1980), Burns and Priebe (1996) and Johnson and Salvador-Carulla (1998).

i) Donabedian divides process interventions into two domains; technical and interpersonal (Donabedian 1980). Technical care refers to the application of the science and technology of medicine and other health sciences, i.e. what is done; interpersonal care to the social and psychological interaction between client and patient, i.e. how things are done. This distinction has been used for describing drug regimens or surgical procedures rather than for mental health interventions however. Donabedian concedes that in the application of psychotherapeutic techniques, the technical and interpersonal elements in management could be virtually inseparable (Donabedian 1980).

ii) Burns and Priebe propose a minimum data set for describing mental health services at service and area level (Burns and Priebe 1996) involving information about:

a) Context: where the service is

b) Target: what a service is and who it is for

c) System: identifying different elements of a mental health system and how they inter-relate

This data set, however, does not provide specific information about particular treatments/interventions provided by mental health services, or how much of each is provided to service users.

iii) Johnson and Salvador-Carulla identify four main ways in which mental health services have been described and classified (Johnson & Salvador-Carulla 1998), all involving description of different process factors:

a) Taxonomies of service types
b) Classification of service styles
c) Measures of the content of mental health services (focusing on the amount, nature and range of interventions delivered to patients)
d) Studies of mental health systems (involving measuring the continuity of care delivered to patients within a service system)

Johnson and Salvador-Carulla's framework usefully identifies service content as an important, distinct element of the description of mental health services, different from description of the type or stated model of a service or its style.

4.1.2 Purposes of content of care measurement

The previous section considered the measurement of content of care within broader evaluative frameworks. This section will consider why content of care measurement is important. Five reasons to measure process factors, including the content of care provided, in mental health services can be identified.

1) **To describe service content:** Measurement of process variables can provide descriptive information about what a service provides for its users. This can identify differences in provision of care to different groups of patients in a service or system. It can identify changes in the care provided by a service over time or differences in content between services.

2) **To assess model fidelity/programme implementation:** Rossi et al. (1999) identify the necessity of process measures, including content of care measures, to allow assessment of programme implementation. Measures of process variables can be compared with predetermined standards or targets, to ascertain how far a service is meeting specific service criteria or agreed objectives. This, however, requires consensus or established guidelines about the theoretical model or operational criteria to which a service is seeking to work (Rossi et al. 1999). Establishing these for mental health programmes or services is not always straightforward, given the *“atheoretical”* way many mental health programmes evolve and are implemented (Brekke 1987).

3) **To understand variation in service outcomes**: While not providing certainty, it can help generate hypotheses about why apparently similar services may exhibit wide differences in outcomes (Johnson & Salvador-Carulla 1998). It is a starting point for moving beyond “black box evaluation” (Rossi et al. 1999) where programme outcomes are evaluated without insight into what might be influencing these outcomes.

4) **To understand variation in patient outcomes:** Pawson and Tilley (1997), considering the evaluation of social programmes, stress that: i) similar process interventions may produce different outcomes for different sorts of service user; and ii) that different process factors within a complex intervention may be the important mediators affecting outcome for different groups of service users. Individual patient level data about care received can help illuminate this by

investigating whether variation in outcomes for groups of patients within a service may be due to differences in responsiveness to interventions or to differences in interventions received.

5) **To assess service quality:** If a process variable is clearly known to produce good outcomes, there is a sound basis for using it as a measure of effectiveness (Donabedian 1980). This link between process and outcomes is often unclear however (Tugwell 1979) (Brugha & Lindsay 1996). The causal relationship between process of care and subsequent health status in patients is very likely to be moderated by factors other than health care. Donabedian stresses the need to adjust for potentially confounding input variables when comparing process with outcome measurement, to be sure of comparing like with like (Donabedian 1992).

Content of care is an essential element of service process which may affect the effectiveness of services. Complex mental health interventions and programmes are likely not to be highly standardised (Mechanic 1996). Descriptions of service type may mask wide variation in service content, as for example with case management (Brugha & Glover 1998). Mechanic therefore argues that measures of service use or continuity of care are often not sufficient to understand variations in outcomes between services and why they may occur (Mechanic 1996): detailed investigation of the content of services – “the black box of service interventions” - is vital. Johnson and Salvador-Carulla (1998) also argue for the high face validity of measuring the content of mental health services, i.e. the nature and range of interventions delivered to patients, as arguably more important than the setting and organisation of the services delivering them.

4.1.3 The need for quantitative measures of content of care

Specific quantitative measures of content of care are required to achieve the five goals identified above. Neither qualitative measures of content of care nor quantitative measures of other process variables can provide similar information. Three qualitative methods of inquiry in mental health research - in-depth interviews, focus groups and participant observation – have been identified (Whitley and Crawford 2005), all of which could provide rich information about what happens at mental health services and how care is experienced. However, qualitative methods are ill-suited to comparing differences, which may be small but still significant, in the intensity or nature of care provided to representative groups of patients at services. Quantitative data is required to provide an empirical basis for identifying active ingredients of care which may positively affect service outcomes.

Quantitative outcome measures may also be used to draw inferences about the care provided at services. Most pertinently, measures of need such as the Camberwell Assessment of Need (CAN) (Slade et al. 1999), can be used to measure whether a service user’s needs in different areas are met during a period of care, from either the service user’s, carer’s or staff’s perspective. CAN is primarily an outcome measure. It is limited as a process measure of content of care for the following reasons:

i) It measures the effectiveness of care, not its provision. If someone receives considerable help with psychotic symptoms which are not alleviated, for instance, this would be recorded in CAN as an unmet need, offering no record that care has been provided.

ii) It measures whether needs are met but not how. For example, it is unclear whether someone with a met need for psychotic symptoms has received pharmacological or psychological treatment, or of what sort.

iii) It provides little scope for measuring how much care has been received. Inpatient care is always recorded as high level care for example.

Unlike CAN, the MRC Needs for Care Assessment (Brewin et al. 1987) does relate assessment of needs to types of care received. A patient’s level of functioning is assessed in 21 categories describing symptoms and behaviour problems and personal and social skills. For each of these categories, the provision, effectiveness and appropriateness of specified types of care are rated, based on information by a staff member who knows the patient well. Need is then rated for each of the 21 areas of functioning as none, met or unmet. Assessment of needs as met is based on the provision of effective or partially effective care.

Between two and nine types of care are rated for each area of functioning, as described in Box 4.1:

Box 4.1: MRC Needs for Care Assessment – ratings of interventions

Rating	Item of Care:
0	Currently provided: effective or potentially effective
1	Currently provided and worth continuing but has proved insufficient after 3 months trial
2	Not appropriate
3	Offered during the past year but refusal, premature termination or non-attendance by patient
4	Given adequate trial in the past two years and proved ineffective
5	Desirable but currently inappropriate due to incapacitating symptoms or other priorities for interventions
6	Not given adequate or recent trial

This categorisation of the types of care provided reflects the focus of the MRC Needs for Care Assessment on assessing the outcomes of care rather than the content of what is provided. It does not provide information about the intensity of care provision. It does not clearly describe when care was provided: for care which proved ineffective, no distinction is made between an intervention provided just before the assessment or one provided two years previously. Describing care through the filter of needs increases the risk of obtaining a partial record of care provided. It is unclear, for example, how help with drug or alcohol problems would be recorded using the MRC Needs for Care Assessment, which includes no level of functioning category for substance use (Brewin et al. 1987). Patient needs and content of care are conceptually distinct: each requires specific measurement tools. The limited information about care provided which can be obtained from CAN or the MRC Needs for Care Assessment illustrates this.

4.1.4 Approaches to measuring content of care

Two organising frameworks, drawn from social research literature, can be applied to describe ways of measuring the content of care in mental health services: source of information and data collection method.

a) Source of information Four sources of data for process measurement of social programmes or health services have been identified (Rossi et al. 1999): 1) direct observation by the researcher; 2) information from service records; 3) data from service providers; 4) data from service users.

b) Method of data collection Bryman identifies two ways of conceptualising how to record activity (Bryman 2004): recording in terms of time or in terms of incidents.

Time recording involves recording whatever is happening over a given period of time to specified person(s) or in a specified area. Bryman (2004) identifies three types of time recording: a) time sampling, or momentary time recording: a snapshot of activity at given instants of time; b) short period recording: a recording of the main activity or activities in a given time period (e.g. ten minutes); c) Long period, or continuous, time recording: all activity is recorded as it occurs over a longer period of time.

Incident Recording involves pre-selecting particular event(s) of interest and recording if and when these happen over a given period of time. A further distinction can be made between contemporaneous and retrospective incident recording. Here, the term **event recording** is used to describe methods of recording incidents at or very near the time they happen. **Questionnaires** (completed by staff, patients, or researchers based on interviews, observation or reference to case records) are used to record information about events of interest gathered retrospectively.

4.1.5 The focus of content of care measures

The previous sections have discussed why and how to measure the content of care in mental health services. This section considers what content of care measurement should consist of.

Hermann and colleagues reviewed process measures used for quality assessment in mental health care (Hermann et al. 2000). They found an absence of measures evaluating the content of non-pharmacological mental health care, but did identify six aspects of treatment process which can be measured: modalities, intensity, duration, patient preference, interpersonal competency and cultural competency.

Brekke (1987) acknowledges that process measurement involves selecting variables for measurement which are considered to be important, from a wide range of different possibilities. He advocates that where possible, the choice of measurement variables to investigate the content of care at a service should be guided by explicit elements of a model of care of known effectiveness. He does distinguish purely descriptive questions about what a service is like from questions about whether a service meets predetermined standards (such as targets or competency ratings). He identifies the nature, frequency, duration, scope, style and setting of care provided, i.e. *"how much of what, to whom, when and in what manner"* (Brekke 1987 p286), as fundamental descriptive variables for measurement of the content of mental health programmes. Where interventions or services lack a clear theoretical model or known associations with outcomes, there is no clear basis for selecting specific interventions or aspects of care for measurement. In such circumstances, measurement of fundamental descriptive variables may be most sensible.

4.2 Aims

This review seeks to identify and assess existing measures of the content of care in mental health services. In Phase 1 of the review, measures and the measurement methods they employ will be described. In Phase 2, the empirical associations between process variables and outcomes which have been found by studies using the measures will be summarised. How far existing measures are able to meet the goals of content of care measurement will be discussed. What is known about how best to measure the content of care in mental health services and the ability of any existing measures to assess the care provided in acute inpatient services will be considered. Directions for future research will be identified.

4.3 Method

Phase 1

4.3.1 Inclusion criteria

The review is limited specifically to measures of the content of care in mental health services. Measures are included which yield quantitative data about the intensity and nature of care at any type of specialist residential or community adult mental health service. This is consistent with Johnson and Salvador-Carulla's description of the content of mental health services as one of four distinct ways of describing mental health services, concerned with the nature and range of services delivered to patients (Johnson & Salvador-Carulla 1998).

This definition therefore excludes measures of related process factors like continuity of care or style of service. It excludes measures such as psychotherapy or pharmacotherapy rating scales which measure the quality or fidelity to a model of a specific intervention or treatment but do not yield information about any other sort of intervention. It also excludes broader measures of the quality or standard of a service if they do not provide information about what the content of service provision is.

This definition includes measures of the content of care received by individual patients, where this data can be aggregated to provide information about the care provided overall at a service. It also includes measures which do not provide information about care received by individual patients, but about the overall care given to patients at a service.

4.3.2 Search strategy

Measures were identified in the following ways:

1) An electronic search of databases: Four Medical and Nursing electronic databases (Medline, Embase, PsycInfo, Cinahl) were searched using a subject heading of mental health services or equivalent combined with terms related to either: 1) the content of mental health services (content of care, process of care, process measure); or 2) methods of process measurement (time recording, time sampling, time budget, event recording, incident recording). The search process used for one database (Medline) is shown in Table 4.1. Search terms were adapted as required for searching other databases.

Table 4.1: Measures review electronic search

Search limits: yrs: 1966-2006; fields: title and abstract	
	Search term
1	**Mental Health Services** [MeSH] explode
2	**“Content of care”**
3	**“Process of care”**
4	**“Process measure”**
5	2 or 3 or 4
6	1 and 5
7	**“Time record$”**
8	**“Event record$”**
9	**“Incident record$”**
10	**“Time sampl$”**
11	**“Time budget$”**
12	7 or 8 or 9 or 10 or 11
13	1 and 12
14	6 or 13

2) Reference lists from studies reporting the measures identified from the electronic search were hand-searched.

3) Consultation with experts: Six experts involved in previous studies of content of care were contacted. They were asked if they were aware of any current studies or methodological approaches to measurement of the content of care in mental health services in addition to what can be found through a search of the relevant literature.

4) Reference works were used to help identify methodological aspects of process measurement to consider (Rossi et al. 1999), (Bryman 2004), (Pawson & Tilley 1997), (Thornicroft & Tansella 1999), (Freeman & Tyrer 1992).

4.3.3 Data abstraction

The following characteristics of measures identified for inclusion in the review were collected:

i) Data collection method

ii) Information source

iii) Level of information provided (care provided to individual patients or overall care provided at a service)

iv) Service settings the measure has been designed for and used in

v) Established psychometric properties of the measure

4.3.4 Phase 2: Identifying use of measures to explore process/outcome associations

A second literature search was conducted to identify studies which had used measures included in the review to investigate associations between a defined content of care variable and patient outcomes. Patient outcomes were defined as subsequent inpatient admissions, clinical or social functioning or patient satisfaction. Studies presenting the measures found in the Phase 1 search were read in order to identify whether the measure had been used to investigate associations between content of care variables and outcomes. Articles citing the Phase 1 studies were additionally identified through an electronic database. (No single database provided citations for all studies: Web of Science, PsycInfo and GoogleScholar were used.) These articles were also read to find any investigation of content of care/outcome associations using measures included in the review.

The following information was collected about identified studies investigating associations between content of care and outcome: study reference; content of care variable measured; outcome variable measured; study setting; whether an association between content of care and outcomes was identified.

4.4 Results

Phase 1

25 measures were identified which have been used to obtain quantitative information about the intensity and nature of care at adult mental health services. The measures identified through literature searching are shown in Table 4.2, categorised in terms of two methodological dimensions identified from reference works: information source (Rossi et al. 1999) and method of data collection (Bryman 2004).

Table 4.2 Measures of content of care in mental health settings

	Event Recording	Time Recording	Questionnaires
Staff	6 measures	3 measures	2 measures
Service users			
Observation by researchers		8 measures	
Records	n/a	n/a	2 measures
Mixed			4 measures

Titles, references and characteristics of the individual measures identified in the literature review are provided in Sections 4.4.1 – 4.4.3, where the measures are grouped according to data collection method.

4.4.1 Event Recording Measures

6 event recording measures were identified (measures recording predefined events of interest at or near the time they occur). Their characteristics are displayed in Table 4.3.

Table 4.3: Event Recording measures

Measure	Information source	Level of data provided	Service settings used in	Established psychometric properties
Daily Contact Log (Brekke 1987)	Staff	Patient/Service	Assertive Community Teams (ACTs) (USA)	Inter-rater reliability established
Mannheim Service Recording Sheet (Salize et al. 1999)	Staff	Patient/Service	Community and Inpatient Services (Spain and Germany)	
Event Record (Burns et al. 2000)	Staff	Patient/Service	ACTs and Case Management (UK)	
Event Report (Hansson et al. 2001)	Staff	Patient/Service	"Integrated care" community service (Sweden)	
Untitled (structured record) (Patmore & Weaver 1989)	Staff	Patient/Service	Community Mental Health Teams (CMHTs) (UK)	
Service Activity Log (Fisher et al. 1988)	Staff	Patient/Service	Case Managers (community) (USA)	

Information source All included event recording measures relied on staff-report. Most measures ask individual staff to record only their own contacts with clients. The structured record described by Patmore and Weaver (1989) requires one key member of staff to record all interventions received by a client from any member of staff at a service during the recording period. Salize and colleagues do not report how many respondents were required to complete the Mannheim Service Recording Sheet for each patient included in a comparison of service use in two European countries (Salize et al. 1999).

Data collection method The Event Report (Hansson et al. 2001) required staff to use a pocket computer to complete daily records; other event recording measures have used paper forms.

Breadth of information Event records may record only face-to-face staff-patient contacts (e.g. the Manheim Service Recording Sheet (Salize et al. 1999), or different types of staff intervention. The Event Record (Burns et al. 2000), for example, identifies five types of intervention: face-to-face, telephone or failed contact with a patient, contact with a carer and contact with another professional. This measure also sets a minimum duration for some types of contact to be recorded: all face-to-face contacts (actual and failed) are recorded, but other interventions only if they are of 15 minutes duration or more.

Depth of information All measures identified the recipient of interventions as well as the provider. The measures could all therefore provide information about what care is provided to individual patients. Event recording requires respondents to categorise interventions as one of between 5 and 11 defined types of care. Three rationales for the choice of categories have been identified:

i) Consistency with an existing established measure: The Mannheim Service Recording Sheet (Salize et al. 1999) for example derives categories of care from an established instrument for describing mental health services, the International Classification of Mental Health Services (DeJong et al. 1991).

ii) To represent important elements of a model of care: The Event Report (Hansson et al. 2001), for example, draws categories of care from a model of care for people with schizophrenia - Integrated Care (Falloon & Fadden 1995).

iii) To reflect the types of care provided in practice at a service: The Event Record (Burns et al. 2000) is informed by a Delphi Process consultation with eight Intensive Case Managers (Fiander and Burns 2000). This generated ten categories of care to describe case management work practices. Such a rigorous process provides some confidence that the categories can describe what is provided in intensive case management services adequately and accurately.

Use in service settings The Mannheim Service Recording Sheet (Salize et al. 1999) provides information about patients' use of the whole local mental health system, including inpatient care, not just a single community service. Other event recording measures have been exclusively used in community rather than residential/inpatient mental health settings.

Psychometric properties Only one measure identified has been tested for reliability. Brekke (1987) tested the inter-rater reliability of the Daily Contact Log using case note vignettes and in actual workplace settings. Over 80% agreement between clinicians was found in ratings of vignettes of staff-patient contacts into different categories of care. (Kappa values of 0.59 and 0.68 were obtained in two tests.) A small study involving a researcher shadowing a clinician, both completing a Daily Contact Log, found 83% agreement between them in categories of care recorded and 90% agreement in the number of staff-patient contacts recorded.

4.4.2 Time Recording Measures

11 time recording measures were identified (measures recording all activity in predefined recording periods). They are described in Table 4.4.

Table 4.4 Time Recording measures

Measure	Type of time recording	Information source	Level of data provided	Service settings used in	Established psychometric properties
Direct Observation Schedule Shepherd and Richardson (1979)	Momentary	Researcher observation	Service	Mental Health Day Centres (UK)	Inter-rater reliability established
Untitled (staff activity measure) (Tyson et al. 1995)	Momentary	Researcher observation	Service	Adult acute inpatient wards (Australia)	Inter-rater reliability established
Dementia Care Mapping (Kitwood 1997)	Short Period	Researcher observation	Patient /Service	Residential dementia care (UK)	Inter-rater reliability unacceptably low (Thornton et al. 2004)
Untitled (Patient Observation) (Higgins et al. 1999)	Short Period	Researcher observation	Patient/ Service	Adult acute inpatient wards (UK)	
Untitled (Staff Observation) (Higgins et al. 1999)	Short Period	Researcher observation	Service	Adult acute inpatient wards (UK)	
Quality of Interactions Schedule (QUIS) (Dean & Proudfoot 1993)	Short Period	Researcher observation	Service	Elderly Mentally Ill inpatient units (UK) and adult acute wards (Nigeria) (Olusina et al. 2003)	Inter-rater reliability established (in EMI setting)
Staff-Patient Interaction Chronograph (Paul 1987)	Short Period	Researcher observation	Service	Inpatient mental health services (USA)	Inter-rater reliability established.
Time Budget (Wright et al. 1987)	Short Period	Staff	Service	Intensive Community Support Programmes (USA)	
Continuous Time Sampling (Bowie & Mountain 1993)	Continuous	Researcher observation	Patient /Service	Elderly mentally ill inpatient wards (UK)	Inter-rater reliability established
Time Budget (Wing and Brown 1970)	Continuous	Staff	Patient /Service	Long-term Adult inpatient wards (UK)	Construct validity tested, but reliability not tested.
Untitled (staff diary) (Patmore and Weaver 1989)	Continuous	Staff	Service	Community Mental Health Teams (CMHTs) (UK)	

Information source Three identified measures seek information from service staff. Two (Patmore and Weaver 1989, Wright et al. 1987) are designed for completion by staff; a third (Wing & Brown 1970) for completion by a researcher based on interview with staff. All other identified measures are completed by a researcher based on direct observation.

Data collection method Short periods in identified time recording measures vary from five to fifteen minutes. Continuous time recording, over whole days or shifts, and momentary time recording measures were also identified.

The main focus varied between all patient activity, all staff activity or staff-patient interactions, but all measures have been used to record the intensity of staff-patient interaction at services during recording periods (see Table 4.5). Measures of staff-patient interaction, such as the Quality of Interactions Schedule (Dean & Proudfoot 1993), seek to record all interactions in a defined area (e.g. one communal room in a residential service) over a series of short recording periods. All patient activity measures identified include specific recording of staff-patient contacts. All identified measures of staff activity distinguish different types of staff activity. For example, the staff observation measure of Higgins et al. (1999) distinguishes four types of staff activity: direct patient contact, indirect patient care, administrative work (e.g., record keeping) and personal activity. No basis for choice of categories of staff activity has been identified for any measure apart from face validity.

Table 4.5 Focus of Time Recording measures

All Patient Activity	Dementia Care Mapping (Kitwood 1997) Patient Observation (Higgins et al. 1999) Continuous Time Sampling (Bowie & Mountain 1993) Time Budget (Wing & Brown 1970)
All Staff Activity	Staff Observation (Higgins et al. 1999) Untitled (Tyson et al. 1995) Untitled (staff diary) (Patmore and Weaver 1989) Time Budget (Wright et al. 1987)
Staff-Patient Interaction	Quality of Interactions Schedule (QUIS) (Dean & Proudfoot 1993) Staff - Patient Interaction Chronograph (Paul 1987) Direct Observation Schedule (Shepherd & Richardson 1979)

Breadth and depth of information All measures, other than the Time Budget Wing and Brown (1970), provide specific information about the intensity of staff-patient contact - either the number of contacts in a given recording period or the amount or proportion of time a patient is in contact with or receiving care from staff.

A number of observational measures record information about the quality of staff contacts with patients: for example rating them as accepting, tolerating or rejecting (Shepherd & Richardson 1979). The Staff-Patient Interaction Chronograph (Paul 1987) provides more detail about the style and quality of interventions without providing information about their content. Only the staff-completed time recording measures categorise the types of care provided in similar detail to the Event Recording measures.

Use in service settings Observational measures have as been used in residential or day care settings. Staff-completed measures have been used in community settings, such a UK Community Mental Health Teams (Patmore and Weaver 1989) or U.S. intensive support teams (Wright et al. 1987).

Psychometric properties Inter-rater reliability has been established for time recording measures based on direct observation by researchers, most comprehensively for momentary or short-period measures focusing on measuring frequency of staff-patient interactions. Inter-rater reliability tests on four measures included in the review indicate that different observers can reliably identify what constitutes a staff-patient contact and rate whether that contact is positive, negative or neutral in nature (Shepherd and Woodward 1979, Tyson et al. 1995, Paul 1987, Dean and Proudfoot 1993).

Wing and Brown (1970) report testing the construct validity of their measure. Time spent doing nothing, not engaged with staff or others, as measured by the Time Budget, did correlate with four other measures of poverty of the social environment.

4.4.3 Questionnaire Measures

The characteristics of 8 identified questionnaire measures (measures retrospectively recording information about predefined events) are summarised in Table 4.6.

Table 4.6 Questionnaire measures

Measure	Information source	Level of data provided	Service settings used in	Established psychometric properties
Client Service Receipt Inventory (Beecham & Knapp 1992)	Staff	Patient/Service	Community settings (various) (UK)	
Untitled (staff activity questionnaire) (Kovess & Lafleche 1988)	Staff	Service	Community mental health teams (Canada)	
Quality Care Intervention Checklist (Glick et al. 1991)	Mixed: Staff, patients, carers	Patient/Service	Community mental health services (USA, Japan, Italy)	
Process of Care Review Form (Popkin et al. 1998)	Records	Patient/Service	Community Mental Health Centres (USA)	
Untitled (service receipt form) (Young et al. 1998)	Records	Patient/Service	Community Mental Health Centres (USA)	
Dartmouth Assertive Community Treatment Scale (DACTS) (Teague et al. 1998)	Mixed (unspecified)	Service	Assertive Community Teams (USA and UK)	Predictive and construct validity investigated. Mixed results from inter-rater reliability testing.
International Classification of mental Health Care (ICMHC) (DeJong et al. 1991)	Mixed (unspecified)	Service	Various European services	Inter-rater reliability established (DeJong 2000)
European Service Mapping Schedule (E.S.M.S). (Johnson et al. 2000)	Mixed (unspecified)	Service	European services/local service systems	

Information source All identified measures, except the staff-completed measure described by Kovess and Lafleche (1988), are designed to be completed by a researcher.

Three measures - the Dartmouth Assertive Community Treatment Scale (DACTS) (Teague et al. 1998), the International Classification of Mental Health Care

(I.C.M.H.C.) (DeJong et al. 1991) and the European Service Mapping Schedule (E.S.M.S.) (Johnson et al. 2000) do not specify what sources of information should be used to complete the measure. Other measures are completed following interviews with staff (Beecham and Knapp 1992, Kovess and Lafleche 1988) or reference to case notes (Popkin et al. 1998, Young et al. 1998). The Quality Care Intervention Checklist (Glick et al. 1991) is completed by a researcher following separate interviews with patient, carer and doctor and a further interview with all three.

Data collection method

Length of retrospective recording period: Measures based on interview with staff ask about the services provided over a period of time varying from one month (Beecham and Knapp 1992) to eighteen months (Glick et al. 1991). Measures using record abstraction seek information, originally recorded at or near the time of the intervention, up to one year retrospectively (Popkin et al. 1998). The DACTS (Teague et al. 1998) involves recording a rating of the frequency of staff-patient contact without specifying how this should be obtained.

Breadth and depth of information: Two identified measures (Popkin et al. 1991) (Beecham and Knapp 1992) provide information about the number of interventions provided to specific clients from a service. Other measures provide less detailed data or ratings of the intensity of patient contact provided (e.g. the DACTS (Teague et al. 1998) provides an overall rating of the frequency of contacts provided to clients at a service). The I.C.M.H.C. (DeJong et al. 1991) does not provide information about the intensity of care provided by a service, just the level of specialisation available for different types of intervention.

Of the identified questionnaire measures, the I.C.M.H.C. (DeJong et al. 1991) categorises care into ten types, the most detailed information about the nature of care provided. The DACTS (Teague et al. 1998) is designed as a measure of fidelity to the Assertive Community Treatment model, but, like the other questionnaire measures, could be used to provide information about the intensity and nature of care provided at any type of service. The European Service Mapping Schedule (Johnson et al. 2000) measures the amount and types of care provided throughout a local catchment area or service system, but does not provide detail about the sorts of care provided at individual services.

The choice of the particular information sought by identified measures was based on one of the following:
i) a reflection of current practice (e.g. the I.C.M.H.C. (DeJong et al. 1991) is based on consultation with a variety of experts about types of care provided in mental health services)
ii) to enable measurement of fidelity of services to a model of care (e.g. Assertive Community Treatment (ACT) for the DACTS (Teague et al. 1998) or published guidelines for evidence based treatment of schizophrenia (Young et al. 1998)
iii) to enable an evaluation of the cost of service provision to individual patients (the C.S.R.I. (Beecham & Knapp 1992)).

Psychometric properties: The I.C.M.H.C. has been demonstrated to have good inter-rater reliability (DeJong 2000). The DACTS shows some evidence of construct validity, replicating findings of previous measures of fidelity of services to the ACT model, and indicates potential predictive validity (Teague et al. 1998). The authors found however that inter-rater reliability for the measure is less securely demonstrated, varying dependent on the types of data used to complete it. No other identified questionnaire measures have established psychometric properties reported. It therefore remains to be established whether reliable information about the intensity and nature of care received by individual patients can be obtained using retrospective questionnaires.

4.4.4 Investigations of association between content of care and outcomes

7 measures included in this review were identified as having been used to investigate the association between content of care variables relating to intensity, setting or nature of care and patient outcomes. These investigations are summarised in Table 4.7.

Table 4.7: Studies investigating association between content of care and outcome

Content of Care Domain	Content of care measure	Association found with outcome?	Outcome variable	Setting
Intensity of care (number of staff contacts received per patient per month/year)	Daily Contact Log Brekke et al. (1987)	No Dietzen and Bond (1993)	Inpatient admissions and bed days (1 year follow up)	7 ACT services (USA)
		Yes Brekke et al. (1999)	Inpatient bed use and social functioning (1 year follow up)	Community Support Program for adults with schizophrenia (USA)
		No Brekke et al. (1999)	Symptoms (I year follow up)	
		Yes Brekke and Long (1997)	Inpatient bed use, employment, independent living status (1 year follow up)	3 community services for adults with schizophrenia (USA)
	DACTS* McGrew et al. (1994) *forerunner of DACTS: IFACT	Yes McGrew et al.(1994)	Inpatient bed use (1 year follow up)	18 ACT services (USA)
	DACTS Teague et al. (1998)	No Morse et al. (2006)	Symptoms, housing status, substance use, client satisfaction (2 year follow up)	2 ACT and 1 standard community team working with homeless dual disorder clients (USA)
	Event Record Burns et al. (2000)	No Burns et al. (2000)	Inpatient admissions (2 year follow up)	4 Intensive and standard case management services – clients with psychotic illness (UK)
	Mannheim Service Recording Sheet Salize et al. (1999)	Yes Salize et al. (1999)	Reduction in unmet needs (skills and abilities) 1 year follow up	Community services in 2 regions - adults with schizophrenia (Spain and Germany)
		No Salize et al. (1999)	Reduction in unmet needs (symptoms) 1 year follow up	
Intensity of care (time spent by patients doing nothing)	Time Budget Wing and Brown (1970)	Yes Wing and Brown (1970)	Rating of clinical improvement (4 year follow up)	3 Long-stay psychiatric hospitals (UK)
Intensity of care (duration of staff-patient contacts)	DACTS* McGrew et al. (1994)	No McGrew et al. (1994)	Inpatient bed use (1 year follow up)	18 ACT services (USA)

Content of Care Domain	**Content of care measure**	**Association found with outcome?**	**Outcome variable**	**Setting**
Setting of care (community vs office-based contacts)	DACTS* McGrew et al. (1994)	No McGrew et al. (1994)	Inpatient bed use (1-year follow up)	18 ACT services (USA)
Specificity of care (proportion of interventions with vocational focus)	Daily Contact Log Brekke et al. (1987)	Yes Brekke and Long (1997)	Employment status (3 year follow up))	3 community services for adults with schizophrenia
		No Brekke and Long (1997)	Independent Living Status (3 year follow-up)	
Specificity of care (no. of interventions providing referral or advocacy)	Service Activity Log Fisher et al. (1988)	No Fisher et al. (1988)	Reduction in severity and number of rated problems (6 months+ follow up)	1 region's Case Management services (USA)
Quality of care (proportion of staff-patient contacts rated positive by observer)	Staff Activity Measure Tyson et al. (1995)	Yes Bowers et al. (2006)	Reduction in rates of conflict and containment (1 year follow up)	2 acute inpatient services pre and post introduction of "City Nurses" (UK)
ACT-fidelity (composite measure)	DACTS * McGrew et al. (1994)	Yes McGrew et al. (1994)	Inpatient bed use (1 year follow up)	18 ACT services (USA)
	DACTS Teague et al. (1998)	Yes Resnick et al. (2003)	Employment status (1 year follow up)	7 ACT and standard community services for veterans with severe mental illness (USA)
		No Bond and Salyers (2004)	Inpatient bed use (1 year follow up)	10 ACT teams (USA)
		Yes Morse et al. (2006)	Housing status and client satisfaction (2 year follow up)	2 ACT and standard community teams – homeless dual disorder clients (USA)
		No Morse et al. (2006)	Symptoms and substance use (2 year follow up)	
	Daily Contact Log Brekke et al. (1987) (main source from which ACT fidelity rating derived)	Yes Mc Hugo et al. (1999)	Inpatient admissions and substance use (3 year follow up)	7 ACT and standard case manangement services (USA)
		No Mc Hugo et al. (1999)	Symptoms, social functioning and satisfaction with services (3 year follow up)	

Table 4.7 shows only seven of the 25 measures included in this review have been used to explore relationships between care provided and service outcomes. Consistent with the findings of the literature review in Chapter 3, this suggests process measurement is far from routine in evaluation of service outcomes.

Investigation of links between process and outcomes has been most common in studies of community-based services. Of 13 studies described in Table 4.7, 11 involved community-based services, 9 were of American services and 9 involved ACT or ACT-like services. The intensity of contact between staff and patients is the domain of content of care used most frequently to explore associations with outcomes. There has been greater focus on the possible effect of how much is done for patients than what is done.

The evidence from this review for links between content of care domains and outcomes is unclear. Table 4.7 shows that; a) the number of studies investigating specific associations is small; and b) for several content of care domains, demonstration of a positive association between service process and outcome has not been consistent.

4.5 Discussion

Progress in developing measures of content of care has been far from linear. There is variation in existing measures regarding what is measured (direct care only or direct and indirect care) and how it is measured. The methodological framework presented in Table 4.1 shows that only a minority of possible methods of measuring content of care have been used in measures described in this review. This review finds that many measures lack a clear theoretical or empirical basis and/or have not been tested for psychometric properties. Many measures have been developed and used for a particular study, but not applied or further developed in future studies or different settings.

Where the association between content of care variables and outcomes has been investigated, findings have varied. Conflicting evidence exists, for example, for the most widely examined questions: whether intensity of care (Dietzen and Bond 1993), (Brekke and Long 1997), Brekke et al. 1999), (McGrew et al. 1994), (Burns et al. 2000) or ACT fidelity (Bond and Salyers 2004), (McGrew et al. 1994), (McHugo et al. 1999) in community-based services affect inpatient bed use.

The lack of repeated, consistent demonstration of association between any content of care variable and patient outcomes in part reflects the inherent difficulties of this type of investigation, where numerous confounding factors other than received care impact on health status (Brugha and Lindsay 1996). It is not implausible, for example, that severity of illness could be associated with increased amount of treatment and poorer health outcomes for patients at a service. It is possible however, that the uncertain reliability of content of care measures used has obfuscated associations with outcomes, or that appropriate content of care domains have not been measured. This review found that the majority of studies of process and outcome associations concerned the link between quantity of direct care and outcomes. Studies which assess the nature of interventions - what staff actually do when they see patients - to investigate links between the content of care and outcomes, remain rare.

The need for effective content of care measurement in mental health services research has been highlighted repeatedly (Burns and Priebe 1996), (Brugha and Lindsay 1996), (Mechanic 1996). Criteria for effective content of care measurement, encompassing psychometric robustness, comprehensiveness, clinical credibility and feasibility, have been proposed (Tugwell 1979), (Donabedian 1980). However, current measures of content of care in mental health services only partially meet these criteria. The following are four challenges to more effective content of care measurement:

4.5.1 Psychometric Robustness

Evidence of inter-rater reliability has been provided most clearly and consistently for researcher-completed direct observation measures, which, however, provide more limited information about the nature of care provided than most other measures in this review. Whether a greater depth of information, or information from sources other than researcher observation, can be obtained with adequate reliability, remains unclear. The work of Brekke (1987) suggests that staff-report event recording measures can provide reliable information about the nature and intensity of staff-patient contact at community services, but the reliability of his Daily Contact Log has yet to be similarly demonstrated in inpatient settings or for other staff-report measures. Momentary time recording using staff report would appear to remove one source of unreliability present in event recording as respondents only have to record the type of care they are providing at the instant of recording, not the predominant type(s) of care provided in a potentially lengthy, complex meeting with a patient. This potentially useful measurement method has been used in HIV case management settings (Abramowitz et al. 1998) but not in mental health settings. Current evidence does not allow comparison of the inter-rater reliability of different measurement methods in similar settings.

A brief questionnaire measure, the I.C.M.H.C. (DeJong 2000) has been demonstrated to provide reliable information about the range of care provided at mental health services. This measure however, provides limited information as it does not assess the amount of care or how much of each type is provided to patients at a service.

There are also obstacles, whatever methodological approach is used, to creating a valid measure which accurately assesses significant elements of content of care. Case note extraction measures may rely on incomplete or inaccurate source material, as found in a study comparing information obtained from patient interviews and case notes (Young et al. 1998). Other retrospective questionnaires may be compromised by respondents' recall bias. All contemporaneous measures, meanwhile, may generate reactivity (Morley and Snaith 1992). Also known as the Hawthorne effect, this describes the phenomenon where the process of measurement may change what is being measured. Burns and colleagues for example, identify this possibility with the use of event records in the UK700 Study (Burns et al. 2000). They speculate as to whether the presence of researchers may have increased the diligence of the staff team during the recording period, or whether staff in some services may have been more identified with the study than others, leading to differences in data recording practices. All staff-completed measures may also be vulnerable to social desirability bias, i.e. deliberate distortion to present a service in a good light.

The extent or comparative impact of these factors on the validity of different methods or measures is difficult to assess. A multi-methods and multi-measure approach to assessing content of care may therefore be helpful: consistent findings from different measures could afford each a degree of convergent validity. This review suggests such an approach is rare, however: in practice, a measure is often developed for a specific study or service setting and used in isolation. The demonstration of clear links between service content and expected outcomes would also increase confidence that valid process variables are being accurately measured, but has also been rare.

4.5.2 Data completeness

The accuracy of information obtained from measures of the content of mental health services depends not just on the measures' psychometric properties but also the response rate they can achieve.

Methods requiring the involvement of researchers alone are likely to obtain more complete data than staff or patient completed measures. In two studies using staff-report measures which document response rate, Patmore and Weaver (1989) report a 66% response rate for an event recording measure; Abramowitz et al. (1998) report an 85% response rate for a momentary time recording measure. However these measures were used in different service settings over different recording periods. Future comparison of the completeness of data obtained using different methods of measurement in similar service settings might indicate benefits of particular methods.

The quality of information obtained from a measure can be compromised by a poor response rate or lack of clarity about response rate. The UK 700 Study (Burns et al. 2000) for example, used staff-completed event records to calculate the mean number of "care events" received by patients in intensive and standard case management services. However, the authors do not record the proportion of staff completing daily event records, nor whether completion rates by intensive and standard case management staff were similar. The accuracy of their figures for the mean number of care events received by patients therefore cannot be estimated.

Measures of the total number of staff contacts received by a patient over a period of time are particularly vulnerable to distortion from poor response rates. The UK 700 Study Event Record (Burns et al. 2000) was also used to calculate the proportion of staff time spent on different activities. The accuracy of these data might still be affected by a low response rate. Unlike the result for the number of contacts per patient however, these results would not be automatically deflated by a low response rate, nor would comparisons between services be so clearly invalidated by significant differences in services' response rates.

The practicality of measures may vary in different service settings, affecting the completeness of data which can be obtained. For example, contemporaneous staff-report measures have been used almost exclusively in community mental health settings (with the exception of Wing and Brown (1970)). There are potential additional difficulties for staff in inpatient services to complete content of care measures: more frequent interactions with patients, lack of time and of space to complete forms privately. It remains to be investigated whether or not staff in inpatient mental health settings can complete content of care measures with adequate response rates and reliability.

Observational measures, by contrast, have been used exclusively in residential/inpatient settings. Observation by researchers in community settings is perfectly possible: Brekke (1987) employed this method in a small-scale exploration of the reliability and validity of the Daily Contact Log, for example. However, the time and cost are much greater than for geographically contained residential settings.

Time recording and event recording measures have not been used in the same range of service settings as questionnaires such as the I.C.M.H.C. (DeJong et al. 1991) or the C.S.R.I. (Beecham and Knapp 1992). Of the most widely used measures providing contemporaneous information, the Event Record (Burns et al. 2000) has been demonstrated to be usable in various community settings, but without response rate or psychometric properties established. The Quality of Interactions Scale (Dean and Proudfoot 1993) has been used in dementia care residential homes and acute adult psychiatric wards, although only tested for psychometric properties in the former.

Measures which are sufficiently practical and reliable to provide detailed information about the amount or types of care provided at a range of mental health settings remain to be established.

4.5.3 Depth of information

A reasonable depth of information about the nature of care and types of intervention provided at services is necessary to understand what services actually do and begin to investigate what works for whom.

Only staff-report methods have been used to provide detail about types of care provided in interventions. However, even event recording measures, which seek immediate information from staff about specific care events, use extremely broad categories of care. Brekke and Test (1992) for example, clarify that the "psychotherapy" category in the Daily Contact Log should include all interventions aimed at monitoring a patient's progress and/or solving problems impeding progress. Other examples of broad categories of care in event records include "Specific Mental Health Intervention": Event Record (Burns et al. 2000)**;** **"**Support": Event Report (Hansson et al. 2001); "Follow up"**:** Service Activity Log (Fisher et al. 1988) "1:1": Daily Contact Log **(**Brekke 1987). It is not easy to infer what specific interventions these categories might include, nor what they tell us about services.

This review found that studies of content of care in inpatient mental health services have assessed the intensity and quality of care, but no measure designed for and used in inpatient settings describes the types of intervention provided. The paucity of our understanding of what happens in inpatient mental health services was discussed in Section 1.2; however, there is no measure of inpatient service content with sufficient depth to help address this issue. If feasible and reliable measures could be developed to provide a greater specificity and depth of information about care provided at services than is currently possible, this would aid attempts to describe and distinguish services.

4.5.4 Accounting for different perspectives

The scarcity of measures of content of care in mental health services using patient-report methods is notable. This seems hard to justify: consumers' perspectives on what care is provided at services are clearly important.

Young and colleagues (1998) asked patients a limited amount of information about care received to cross check with information from medical records abstractions. Wing and Brown (1970) report asking patients about their activity as well as nursing staff if this proved necessary. It is not clear how often this was necessary. Glick and colleagues address the issue of health care participants' different perspectives head on (Glick et al. 1991). Their study specifically sought information about care provided to 24 patients through interviews with physician, patient and family carer. However, it was not reported whether there were consistent differences in response from the three types of respondent. Moreover, they report conducting a subsequent interview with physician, patient and carer together, to "*reconcile discrepancies*" (Glick et al. 1991 p.56) in their accounts. It is not reported or evident how this could be achieved.

Measures of patients' needs (Slade et al. 1998) or the style of service (Rossberg & Friis 2004) have identified significant differences between the views of staff and patients. The issues of whether there are significant differences in view between patients and staff about the content of care provided at services and how best to measure this remain to be addressed.

4.6 Conclusion

The starting point of this review was that there is no consensus about how best to measure the content of care in mental health services (Burns & Priebe 1996), (Johnson & Salvador-Carulla 1998). This review suggests that there is no compelling evidence to recommend any one method or measure to investigate the nature of care provided at services and the intensity of care in community services. Measures of the intensity of staff-patient contact based on direct observation by researchers have most frequently demonstrated reliability in inpatient settings and promise fewest difficulties with response rate.

This review identifies a number of areas where further research could add to limited current knowledge about how to measure the content of care in mental health services:

- The development of measures which provide greater depth of information about the nature of care provided at services, especially inpatient services.
- More testing of the psychometric properties of measures across a range of service settings.
- More investigation of the feasibility of measures in different service settings, including routine reporting of completion rates in use of process measures in studies.
- The development of measures which include patients' perspective on the content of care at services.

Given the absence of gold standard measures or established ideal methods to measure the content of care in mental health services, and given the desirability of including consumers' and providers' perspectives, a multi-methods approach should be adopted in studies of the content of care in mental health services. Consistent findings from different information sources and data collection methods would increase confidence in results. A focus on the nature of interventions provided by services, not just their number or the type of service within which they are provided, can aid description and distinction of mental health services and the goal of understanding service outcomes.

Chapter 5

The Development of Measures of Content of Care for Inpatient Settings

5.1 Introduction

The decision to develop new measures of content of care was guided by the aim of this thesis to assess the intensity and nature of care provided at acute inpatient services. The term intensity of care is used throughout this chapter to mean the amount of staff-patient contact at services, incorporating both frequency and duration of contacts. Nature of care is used to mean the different types of interventions provided at services. Four new measures of content of care in inpatient mental health services which use different methods were developed by the author of this thesis. New instruments and this multi-methods approach were appropriate because:

i) Existing measures of content of care feasible for use in inpatient settings provide insufficient depth of information about the nature of care at services. The information provided about intensity of care by established observation-based measures is not fully representative of service provision (Section 5.2.1).
ii) Different methods are most feasible and appropriate for measuring different elements of content of care.
iii) The perspectives of key stakeholders, i.e. service users and service providers, can each be included.
iv) Comparable data from more than one measure can aid interpretation of results: convergence in results from different measures would increase confidence in findings and the measures' validity.

The development of draft measures is described in Section 5.2. Piloting and revision of measures is described in Section 5.3. Psychometric testing of measures is described in Section 5.4. The strengths and weaknesses of the measures are discussed in Section 5.5

5.2 Development of draft measures

The procedures used to develop measures reflect the stages of questionnaire design advocated by Oppenheim (1992): the study's aims were identified; the relevant literature reviewed; hypotheses to be investigated were decided; research instruments were designed, piloted, revised and psychometrically tested; the study's sample decided. Relevant literature regarding measures of content of care has been reviewed in Chapter 4. The choice of research questions for the quantitative study within this thesis and the aims and hypotheses will be presented fully in Chapter 6. The samples for the study are reported in Chapter 7. This chapter will describe the design, piloting and revision of measures.

The development of four measures is described:
CaSPAR: The Camden Staff-Patient Activity Record
CaRICE: The Camden Record of Inpatient Care Events
CCCQ-P: The Camden Content of Care Questionnaire (Patient version)
CCCQ-S: The Camden Content of Care Questionnaire (Staff version)

Final versions of the measures are provided in Appendices 3 to 5.

5.2.1 CaSPAR

Proposed purpose CaSPAR was developed as a measure of the proportion of service users in contact with staff at a service at any one time. When aggregated, it is an indicator of the overall intensity of staff-patient contact at a service. CaSPAR provides service level information only. It is not a measure of the amount of contact with staff received by specific patients.

Candidate methods Previous studies (Table 4.4) indicate that time recording, based on researcher observation, has been most clearly and frequently shown to be a reliable way of measuring the amount of care provided in mental health services. Whether based on direct observation or information from staff, time recording measures can be more feasibly completed by researchers than event recording measures, as they do not require researchers' presence with staff over lengthy continuous time periods. The potential for researcher-completion is an advantage of time recording measures, as it will maximise completion rate.

Momentary time recording was preferred to short-period time recording for three reasons. First, it has been used by researchers in observation measures with good inter-rater reliability (Tyson et al. 1995). Second, it is likely to minimise bias and unreliability in measures using staff-report, as it requires simple information about what is happening at one point of time, compared to more complex appraisal of the predominant activity during a period of time, as required for short period time-recording. Finally, momentary time recording, being the shortest time period of time recording, also makes the least demand on researchers' time, allowing the maximum amount of data to be collected using the resources available.

Information source

Patient-report was rejected as an information source for this measure: obtaining participation and adequate completion rates from patients for a measure involving sustained recording over a long period of time could not be guaranteed; nor is there evidence for inter-rater reliability of patient-report time recording measures. Observation by researchers and staff-report were considered as information sources.

Previous studies have demonstrated that researchers can reliably identify when staff are in contact with patients using direct observation (Tyson et al. 1995). This has not been demonstrated for staff-report measures in inpatient settings, although there is some evidence of good staff-observer agreement in identifying staff-patient contacts from a community-based study (Brekke 1987). Consideration of the reliability of methods of measuring the amount of contact between staff and patients in inpatient settings would therefore dictate that direct observation by researchers is used.

There are however, practical and ethical constraints to where direct observation of patients can be conducted. Previous studies in acute inpatient settings have either limited observation to a small number of consenting, not necessarily representative, patients (Higgins et al. 1999) or to all patients but only in accessible, communal areas (Olusina et al. 2003). However, it is likely that a significant proportion of staff-patient contacts in inpatient services may occur in non-communal areas, such as interview rooms, patients' rooms, outside the unit. There is no reason to assume that the proportion of total staff-patient contacts which are observable is the same in different services. It is desirable for validity that a measure of the amount of staff-patient contact in inpatient services should measure contacts with all patients in all parts of the unit; this cannot be achieved by observation alone. Service staff, however, routinely identify where patients and staff are in the unit and are well placed to provide this information for given moments in time. Ward procedures dictate that the shift coordinator and other staff are informed of staff and patients who have left the building, and are aware of staff involved in close observations with patients, potentially in areas inaccessible to researchers, and other activity such as ward rounds taking place in private.

To supplement data on staff-patient contact in communal areas gathered from researcher-observation, it was decided that an appropriate member of staff, such as the designated shift-coordinator on a ward, could be approached by a researcher and asked to identify how many patients are currently with staff, either out of the unit or in areas other than the communal areas observed by the researcher. CaSPAR uses researcher observation where possible, supplemented by staff-report where necessary to gather complete information about the contact between staff and patients at the service at recording points. This is a compromise between maximising the face validity and confidence in the reliability of the measure.

Depth of information

It was not feasible to obtain information beyond the amount of staff-patient contact at a service from CaSPAR because of its mixed sources of information. Observation measures can provide reliable information about the quality of interactions (Dean & Proudfoot 1993) (Shepherd and Richardson 1979) (Tyson et al. 1995) e.g. whether they are accepting, neutral or rejecting; staff-report measures have not been used in this way. Staff report measures can provide information about the purpose/types of interactions (Brekke 1987) but observation measures have not been used in this way. It is unlikely that researchers observing staff-patient contacts in communal areas could infer the purpose of each contact, or that staff could reliably rate the quality of their own and colleagues’ interactions.

The only information sought in CaSPAR other than whether patients are with staff or not, is whether they are in or out of the unit. This serves two purposes: 1) it can assist the member of staff providing information in focusing on where other staff are and whether they are with patients; 2) it also provides a measure of the extent to which patients retain contact with the community outside the inpatient service during an admission.

The sampling frame

Two approaches were considered for deciding the number of recording periods: a power calculation or a pragmatic approach. A power calculation of the number of recording periods necessary to identify significant differences between types of service was not feasible because the number of patients resident at each service and the proportion of patients likely to be in contact with staff at any one time was not known. A pragmatic approach was therefore adopted. The number and timings of recording periods were chosen to be sufficient to reflect adequately variations in and levels of staff-patient contact throughout the day.

The times of recording periods were decided in advance of the selection of services for The Alternatives Study, so were unaffected by knowledge of what happens when at particular services. The following considerations influenced the choice of recording times:

- Recording periods were included at evenings and weekends, so CaSPAR measures staff-patient contact overall at each service.
- It is likely, however, that most variation in amounts of staff-patient contact during weekdays occurs between 9am and 5pm, as more staff are likely to be at work at these times (e.g. doctors, occupational therapists etc as well as nursing or residential care staff). Therefore 75% of contacts each day were set between 9am and 5pm.
- Recordings were not made at night (after 8.15pm or before 8.15am) partly for pragmatic reasons given available data collection resources, but also as levels of staff-patient contact are likely to show least variation at night time because of reduced staffing levels and patients being asleep.

- Recording times were set at differing times each day so that routine daily events at services are not always excluded or included in the measure. One recording is made on a day between Monday and Friday for each 30 minute point between 9.15am and 4.15pm.

Maximum limits to the number of recordings to be made each day (2) and each week (10) were set so that:

- Recordings are made over a number of weeks (at least three), rather than one or two weeks which may or may not be typical of the service generally provided.
- The risk of reactivity – staff's behaviour being affected by knowledge that their actions are being recorded – is reduced.

Twenty eight recording points were set to allow the same number of recordings each day, recordings at a range of times each day and throughout the week and a substantial amount of data to be collected. Recording time points are detailed in Appendix 3.

Scoring

Each observation provides a figure for the proportion of all patients resident in contact with staff at the moment of recording. These data from each recording can be aggregated to provide one score for each service for the mean proportion of patients in contact with staff. Patients marked as not known in CaSPAR recordings are excluded from calculations of proportions of patients in contact with staff.

Summary

CaSPAR uses momentary time sampling to provide information about how many patients at a service are in direct contact with staff at given moments of time. A researcher directly observes and records staff-patient contacts in accessible, communal areas. The researcher then immediately asks a member of staff to identify:

a) Whether any staff are in contact with patients in parts of the unit inaccessible to the researcher (e.g. a patient's room).

b) How many patients are out of the unit, either with staff (e.g. on escorted leave) or without staff.

The total number of patients resident at the service on the day of observation is also recorded. Twenty eight observations are made in total: four per day at given moments listed in CaSPAR.

5.2.2 CaRICE

Proposed purpose

CaRICE, a second service level measure, was developed to provide information from a second data source about the intensity of staff-patient contact at services and additional information about the nature of care provided during staff-patient contacts.

Method

Event recording and time recording were considered because both these methods can provide reliable information about the nature of care provided at a service. Event recording was chosen because:

i) Event recording can provide a record of every staff-patient contact at a service during a significant period of time. A fully representative picture of the nature of care provided in staff-patient contacts can therefore be obtained from a shorter recording period than would be possible using momentary or short-period time recording.

ii) Use of event recording for CaRICE would provide a second method of measuring the intensity of care provided at services, thus allowing triangulation with CaSPAR data.

Information Source

Observation by researchers was rejected as an information source for CaRICE because of the difficulties for researchers in inferring what sort of care is being provided in staff-patient interactions (discussed in Section 5.1). Audio-recording staff contacts with patients is one means of allowing researchers to witness staff at work and assess the content of interactions with patients without needing to observe staff at the time of the interaction. This method has been used to rate the model fidelity of therapist practice, for example for rating the Cognitive Therapy

Scale (Vallis et al. 1986). Audio-recording staff contacts with patients was not feasible in inpatient settings however. Because contacts with patients are often not pre-planned, the issues of obtaining client consent and the amount of recording required to obtain representative information about staff activity were prohibitive.

Patient report was rejected because of the difficulties of obtaining consent from all or a representative group of patients and of ensuring adequate completion rates for a measure involving sustained recording over a long period of time. Staff-report was therefore chosen as the information source for CaRICE. It was decided to involve all clinical staff at the service in completing the measure, rather than just particular professional groups or representatives of each professional group. This has two benefits. First, it strengthens the validity of CaRICE, by ensuring that all types of care provided by any individual staff at a service are measured. Second, it is likely to maximise completion rates. Collecting data from as many staff as possible at each service will mean the recording period can be as short as possible, minimising the burden on individual staff.

The sampling frame

The time frame chosen for completing CaRICE was influenced by the need to obtain representative data about services while maximising data completeness and response rate. The review in Chapter 4 found that event records had not previously been used to measure the content of care in inpatient settings, so the likely response rate from staff was unclear. Completing an event recording measure may be more arduous for staff in acute inpatient services compared to community services however: they may potentially have more contacts per day with patients and more distracting, unplanned situations to attend to. It was therefore imperative to minimise the demand on staff time to complete the measure, to try to maximise the completion rate.

One working week (five days, Monday to Friday) is the minimum time required to reflect all the care routinely provided at services where some staff may work or interventions be provided only on particular days. The brevity of the completion period is not ideal: the week in which data are collected may not reflect the care provided at the service typically for two reasons. First, key staff may be absent during the recording period. Second, there may be a heightened “Hawthorne Effect” due to the short recording period, whereby staff’s normal behaviour alters due to their awareness of the research taking place. These risks to validity were counterbalanced by the need to try to maximise completion rates for the data. A five-day recording period at each service, Monday to Friday, was therefore chosen.

It was decided to record staff contacts with patients of five minutes or more in duration only, rather than a shorter duration or all momentary contacts with patients. This was primarily in order to keep the measure feasible: in residential settings, the number of momentary contacts with patients (saying hello, unlocking a door etc) might be very large indeed, making the recording process overly burdensome. A minimum duration of five minutes for a contact will improve feasibility and also focus the measure on care-giving events rather than casual contacts with patients, providing a potentially more valid measure of care provided to patients.

Depth of information

It was decided to collect only service level information about care provided with CaRICE, not data about what care is received by individual patients. This was because:

a) A longer time frame for data collection would be required to provide meaningful data about what care had been received by individual patients, which might adversely affect the measure’s feasibility.

b) As planned, CaRICE can be used to provide complete data about all care provided to all patients at a service during the recording period. If staff were asked to provide anonymised data about care provided to individual patients, this would increase the burden of completing the form and potentially jeopardise response rate. If patients’ individual consent were sought for staff to provide information about the care provided, it is unlikely that data about all patients at a service could be collected.

It was decided to seek as much detail as possible about different types of care provided at services consistent with retaining face validity, reliability and feasibility. As discussed by Brekke (1987), for services where the active ingredients of care influencing effectiveness are not known and for services which are not theory or model-guided, no basis exists for prioritising the measurement of specific elements of care provided: fundamental descriptive elements, such as the intensity and nature of care, are appropriate variables to measure. This is the case at UK acute inpatient mental health services, where knowledge of what is provided to patients on acute wards and how it is experienced by them is limited (Quirk & Lelliot 2001) and there is no established theoretical model or known link between process and outcomes for inpatient services to guide decisions about which elements of the content of care to measure. Categories of care chosen for this measure are derived from three sources:

a) The categories used in an existing event record developed for the UK700 Study, based on a systematic attempt to identify and describe types of care provided in UK community mental health services, using a Delphi Process with case managers (Fiander & Burns 2000). No similar process has been carried out for inpatient settings.
b) Descriptions of types of care provided at an inpatient ward and a crisis house in North London from a qualitative study (Johnson et al. 2001). Service users and staff were interviewed and case notes reviewed by researchers to identify types of care provided.
c) Suggestions from the advisory group for this study (researchers, user-researchers and clinicians with relevant skills and experience).

The origin of each of the categories of care used in this measure is shown in Table 5.1.

Table 5.1 Source of categories of care for CaRICE

UK700 Study Categories	**Qualitative Study additional types of care (Johnson et al. 2001)**	**Advisory Group Suggestions**	**Categories of Care piloted for this study** *(abbreviated)*
Housing			Housing
Finance			Finances
Occupation	Structuring current daily activity		Current Activity
	Help arranging new occupational activity (e.g. a college course)		Future Activity
Daily living skills			Activities of Daily Living
Criminal justice		Other forms of legal help also important (e.g. help with immigration status, appealing against Section)	Criminal Justice
	Help with legal access to children		Other legal
Carers and significant others			Carers' support
Mental Health intervention and assessment	Creating a safe environment/ Monitoring safety		Safety
	Monitoring mental state		Assessment
	Talking about current difficulties or triggers for illness	Useful to try to separate help with symptoms and help with interpersonal problems	Relationships or past events
	Coping with symptoms		Symptom coping
		Providing information/explanation about diagnosis or problems not covered	Illness education
Medication	Medication changed		Medication review or change
	Medication compliance encouraged	Practical help dispensing medication is distinct from efforts to help medication concordance	Medication practical help
			Medication concordance
Physical health			Physical health
	Detox provided/ working on drinking problems		Drugs/alcohol
Case conferences			Care planning meetings
	Referrals/encouragement for engagement with other services		Discharge/aftercare planning
		Piloting is an opportunity to check whether categories are comprehensive by including an "other" category and asking respondents to describe activity.	Other
Engagement		Problems foreseen using this category: many contacts involve engagement; few exclusively	(excluded)

Recording process:

Inpatient staff completing CaRICE were asked to record the time they spent at work and the duration and category of each contact of five minutes or more with a patient. In order to provide a consistent and valid measure of the proportion of staff time spent with patients, these were operationalised as follows:

i) Recording total time spent at the service

For staff whose whole working day is spent at an inpatient service where CaRICE is being used, calculation of their time spent at work is straightforward: their total time spent at the service each day is calculated. Criteria for recording the time spent at work were established for staff in the following two categories:

a) Sessional/casual staff: For staff, such as doctors on call for more than one ward or an advocate who provides help to patients as required, their time spent physically at the service where CaRICE was being used was recorded rather than the duration of their working day.

b) Services where not all patients are receiving acute/crisis care: for services containing a mixture of crisis beds and other beds (e.g. respite), only care provided to crisis patients was recorded in CaRICE and staff's total work time calculated pro rata: e.g. if 25% of the beds at a service were crisis beds, a worker on an eight hour shift would be considered to have 25% of eight hours (2 hours) at work at the crisis beds. This could lead to misrepresentation of staff's total time spent at work, as it is possible that staff might be expected to work in a different way and devote different amounts of time to crisis and non-crisis patients. It will not affect CaRICE data for the actual minutes of contact provided to crisis patients however.

ii) Recording contacts with patients

Three options were considered for how to direct staff to record contacts using CaRICE:

i) To record the duration of each 1:1 contact with a patient and the one most appropriate category of care to describe the main type of care provided in the contact.

ii) To record the duration of each 1:1 contact with a patient; use any number of appropriate categories of care to describe the care provided; record what time out of the total contact time was spent on each type of care.
iii) To use any number of appropriate categories of care to describe the care provided in a contact, but only to record the total duration of each 1:1 contact.

Unlike option 1, option 3 allows respondents to describe different types of care provided to a patient during a single contact. Option 3 is more straightforward to complete than option two and does not require respondents to allot time to different interventions within one contact, which may not always be clear if the interventions are concurrent. These benefits outweighed disadvantages with option 3 of not providing information about the duration of different interventions within a single patient contact and vulnerability to variation in respondents' propensity to use multiple categories to describe care events. It was considered of primary importance to create a measure which allows staff to describe all the types of care they provide and which is sufficiently user-friendly to produce an adequate completion rate. Option 3 was therefore chosen.

The potential problem of differences between respondents in recording style leading to similar actions being recorded differently was explored by a test of the inter-rater reliability of the measure, described in Section 5.4.1: the better the inter-rater reliability, the less significant this problem becomes.

Staff were not asked to distinguish between 1:1 contacts with patients and contacts with groups of patients when completing CaRICE. This has implications from how data about the intensity of staff-patient contact is calculated using CaRICE. For example, a group run by one staff member for one hour with six patients would be recorded and interpreted as one hour of staff-patient contact, not six hours. It is proposed that this reflects a reality that participants may not receive continuous direct contact with staff throughout the duration of a group. This also minimised the recording burden for staff completing CaRICE, thus maximising feasibility of the measure.

Scoring

CaRICE provides data about the number and duration of staff-patient contacts and the number of interventions of each different type in each contact (e.g. one contact with a patient recorded with three types of care counts as three interventions.). The exact duration of each intervention is not provided (and indeed might not even be clear to the person completing the form, if two interventions were provided concurrently). An estimated figure for the duration of each intervention can be calculated by dividing the total duration of a contact by the number of interventions within it.

Two types of service information can therefore be derived from aggregating data from all CaRICE forms completed at each service each day:

a) **the proportion of staff time** spent in direct contact with patients: the total time spent by staff in contact with patients can be summed from all of a day's completed forms and divided by the summed total of time spent at work. This provides one item of data for the proportion of staff time spent in contact with patients at a service on one day. Daily data can be aggregated to provide one mean figure for the proportion of staff time spent in contact with patients at each service. Data for the proportion of staff time spent providing each type of care to patients can be calculated in a similar way.

b) **the minutes of direct contact** provided at each service per patient per day: the total time spent by staff in contact with patients can be summed from all of a day's completed forms and divided by the number of patients resident that day. This provides one item of data for the minutes' contact per patient each day. Daily data can be aggregated to provide one mean figure for the minutes' contact per patient per day at each service. Data for the minutes' contact per patient per day spent providing each type of care can be calculated in a similar way.

The minutes of direct contact per patient per day is potentially the more valid measure of intensity of care, as it directly measures how much care patients receive. The relationship between the proportion of staff time spent with patients and how much care patients receive, by contrast, will be mediated by staff-patient ratios at each service. Of the two variables however, the minutes of staff-patient contact is more sensitive to response rate from the measure, as it involves summing the care events provided at services, rather than calculating contact as a proportion of time spent at work. Low response rates from the measure would lead to underestimates of the intensity of care provided to patients at services, while significant differences in response rates between services might distort comparisons of different services.

It is therefore helpful that CaRICE offers two possible variables to measure the intensity and nature of care at services. If response rates are high and uniform, the minutes of contact per patient per day can be used; if not, the proportion of staff time spent in contact can be used.

The two CaRICE variables described above can provide similar data regarding each of the 21 types of care within the measure. CaRICE could therefore be used specifically to look at the intensity of provision of care concerning, for instance, housing or medication compliance. The development of subscales for CaRICE to provide information about broader types of care is reported in Section 5.2.4.

Summary

All clinical staff at a service will be asked to complete CaRICE each day over a five-day period. Staff will record:

a) The date of the recording and the number of hours worked
b) The types of care provided at each contact of duration of five minutes or more, chosen from a list of 21 categories of care
c) The duration of each 1:1 contact of 5 minutes or more between the member of staff and a service user.

The patient involved in each contact is not identified in the form. The measure is designed to be completed by staff during the course of their working day, after each client contact. A researcher will be present at the service and give recording forms to staff. The researcher will record the number of forms given out and how many are returned.

CaRICE provides a measure of:

- The proportion of time at work spent by staff in 1:1 contact with patients and the proportion of time spent at work providing each of 21 types of care
- The mean minutes of direct care provided per patient per day and the mean minutes of each of each of 21 types of care per patient per day.

5.2.3 CCCQ-S and CCCQ-P

Purpose of measures

CCCQ-P and CCCQ-S are measures of the intensity and nature of care provided to an individual patient during an inpatient admission, from staff and patient perspectives. Individual patient data are required to investigate how the care provided to patients is associated with outcomes and affected by patient characteristics. This patient-specific information could help explore:

- Whether patient variables predict the intensity and nature of care provided to patients
- Whether identified differences in the intensity and nature of care provided at different services remain significant, controlling for patient variables.
- Which process variables appear to mediate outcomes for particular groups of patients.

Method

A momentary or short-period time recording measure would not provide information about all the care provided to a patient during an admission. Event recording could do so in theory. However, one respondent could not provide contemporaneous information about all the care provided to a patient by all members of staff at an inpatient service. Given the multiple care providers to individual patients in inpatient settings, event records would need to be completed by many staff over a patient’s whole admission in order to provide information about any one patient. This may not be feasible, as discussed in Section 5.2.

A retrospective questionnaire method was therefore chosen for this measure. This also provides a third method to measure the intensity of care, and a second method to measure the nature of care provided at services, to enable a comparison of results from different methods.

Information Source

Researcher-observation is not a practical way to obtain information about all the care provided to a patient during an admission. Case note abstraction was considered: a patient's case notes do provide a record of the care provided. However, previous studies have found shortcomings with the accuracy of information provided in case notes (Young et al. 1998). The Advisory Group for this study also expressed doubts about the completeness of information about care provided to be found in inpatient clinical case notes. This method was therefore also rejected.

Staff completion and patient completion were considered as sources of information for CCCQ. The need to include a patient perspective in measuring content of care was discussed in Chapter 4. It is not possible to guarantee that a complete cohort or a representative group of patients would agree to complete questionnaires, however. Moreover, staff and service users may have differing perspectives on what care has been provided during an admission. It was therefore decided to develop two versions of the Content of Care Questionnaire: a staff-completed version and a patient-completed version.

Staff version: One member of staff was asked to provide retrospective information about all the care provided to a patient. Inpatient and residential services in England do routinely identify one member of staff (a Primary Nurse or key worker) to take an overview of and coordinate a patient's care during their admission. This is also the most feasible way of collecting information (rather than asking all staff involved in providing care to a patient during an admission to complete a questionnaire and collating data subsequently). The possibility was acknowledged, however, that one key worker may not be aware of every intervention provided to a patient during an admission. The staff member completing the form was therefore directed to consult colleagues or case notes in completing the questionnaire if necessary. The reliability of retrospective staff-report information gathered from CCCQ-S will be investigated in Section 5.4.2.

Patient version: Individual patients, identified by staff as close to discharge from an inpatient stay, were asked to give retrospective information about the types of care each had received during an admission.

Sampling frame

Two options were considered for the questionnaire's time frame: a patient's whole admission or a fixed, consistent time period (e.g. the care provided in the last week or fortnight). The latter option was rejected for two reasons:

i) The types of care provided to a patient may well change during the course of an admission, so the care provided in any single week is not guaranteed to be typical of the care provided altogether during an admission.

ii) Any fixed time period will represent a different proportion of the whole admission for different patients: asking about care received during a whole admission may have greater validity than asking about care received during varying proportions of an admission. CCCQ potentially allows investigation of the extent to which length of stay mediates any association between type of service stayed at and the intensity or nature of care provided.

Depth of information

Three considerations were taken into account when deciding the depth of information to be sought from CCCQ: maximising feasibility and response rate; maximising the measure's reliability; providing information which can be compared with that gained from other measures. It was therefore decided to use the same categories of care as CaRICE (Section 5.2).

It was not feasible to ask respondents to provide accurate responses retrospectively about total number of staff contacts or the duration of contacts during admission. The more complex the information sought, the greater the potential problems of recall bias: it was therefore initially decided in CCCQ to ask only for simple information about whether any care from each category had been provided to the patient during the admission because it was considered likely that a respondent would be aware of this and able to provide a basic yes/no answer. An eight-point scale concerning the frequency of provision of each category of care was added following piloting (Section 5.3).

CCCQ staff and patient versions do not directly measure the duration of staff-patient interactions, but do measure variables – the frequency with which care of different types is provided – which relate, like CaSPAR and CaRICE data, to the intensity of care at services.

Scoring CCCQ-P

Two measures of the overall care provided to a patient during an admission can be derived from CCCQ:

1) CCCQ range of care score: Summing data from each category about whether any care was provided, giving a score between 0 (no care received) and 21 (care received in all categories).

2) CCCQ total care score: Aggregating data from each category about how frequently care was provided, giving a score between 0 -147.

Data can also be obtained about whether care was provided and how frequently for each individual category of care. Data for all patients can be aggregated to provide a mean score for each service.

The range of care score from 0-21 reflects the range of care provided but not its intensity. The total care score from 0-147 does provide a measure of the intensity of care provided to a patient: when aggregated to provide a mean service score, this is more comparable with CaSPAR and CaRICE data about the intensity of care provided at each service. This total care score from CCCQ data was therefore used in this thesis.

Missing values: 4 missing values may be prorated for CCCQ-P total score. One missing value may be prorated for each subscale score (see Section 5.2.4 below). Data can be missing from CCCQ-P in two ways: either no information was provided about one category of care, or information was provided about whether care was provided but not how frequently. If no information was provided, items will be prorated as 0 (piloting indicated that categories were not uncommonly left blank if no care had been provided). Where care was recorded as provided but a frequency score was missing, a score will be prorated as the mean score of other CCCQ-P categories.

Summary

CCCQ provides information about the intensity and nature of care received by individual service users over the whole of an admission to an acute inpatient service. Questionnaires can be self-completed by respondents or completed as a structured interview. Questionnaires should be completed as near as possible to the point of a service user's discharge.

CCCQ-P and CCCQ-S provide, from patient and staff perspectives respectively, measures of:

- Whether the service user has received each of twenty one types of care during his/her admission (range 0-21: high score means greater breadth of care received)
- A total care score for the frequency and range of care provided during an admission (range 0 -147: high score means more frequent receipt of care).

Both measures provide information about individual patients which can be aggregated to provide a service-level measure of the intensity and nature of care provided to patients at each service.

5.2.4 CaRICE and CCCQ subscales

Subscales were created for CaRICE and CCCQ for use in comparing alternative and standard services in the investigation reported in Chapters 6 - 10 in this thesis. Comparing the provision of all 21 individual CaRICE/CCCQ categories of care at alternatives and standard services would provide 21 results. Data reduction through creation of subscales has three benefits: a) it allows hypothesis-driven investigation; b) it reduces risk of Type 1 errors through multiple testing while; c) not, given the sample size, introducing risk of Type 2 errors through statistical correction for multiple testing (Section 10.1.3). Statistical methods of deriving subscales such as factor analysis or testing the internal consistency of predefined subscales were precluded by:

i) the small number of services in the study providing too few CaRICE data for statistical derivation of subscales

ii) the lack of clearly reliable patient level data from CCCQ (discussed in the psychometric investigations in Section 5.4)

A pragmatic data reduction strategy was therefore adopted and four subscales were developed following consultation with clinicians from different disciplines and service users. Thirty five clinicians and service users were asked to place the 21 categories of care from CaRICE into four subscales; medical, psychological and social interventions and general care organisation. These subscales mirror a commonly used way of distinguishing different types of mental health care interventions proposed by Engel (1977); the biopsychosocial model. A fourth category – general care organisation – was added acknowledging that some categories of care (e.g. assessment) were common to more than one of the other subscales.

Consultation with clinicians and service users was used in two ways. The nomenclature of subscales was refined: medical interventions was replaced by physical and pharmacological interventions following lack of agreement about whether observations and restraint constituted medical interventions. Categories of care were placed in subscales according to the majority view wherever possible. A clear majority opinion (more than 50% of respondents using the same subscale) was obtained for all but two of the categories (observations and assessment being the exceptions: for these, the most frequent response was used), indicating the subscales have a degree of face validity. The lack of a clear empirical basis for the subscales or clearly established reliability from clinicians or service users in how to place categories of care in subscales however, means that their use in comparing the amount of types of care in alternative and standard services must be exploratory. The categories of care from CaRICE and CCCQ-P comprising each subscale are identified in Table 5.5 (page 156).

5.3 Piloting of measures

5.3.1 Rationale for piloting

CaSPAR: Given limited resources, piloting CaSPAR before use was not prioritised. As a primarily researcher-completed measure, its clarity and acceptability to clinicians was less crucial than for CaRICE. It seeks less complex information than the other measures of content of care developed for this study (intensity of care only, not also the nature of care provided) and mainly uses a method (momentary time recording based on researcher-observation) which has been shown to be reliable and feasible in inpatient settings in previous studies (Tyson et al. 1995, Paul 1987).

CaRICE and CCCQ: Piloting CaRICE and the staff and patient versions of CCCQ was prioritised in order to check that the measures were clear to respondents, that the categories of care were necessary and sufficient to describe the care provided in inpatient services and to check practical considerations of how best to administer the measures. CaRICE and CCCQ were therefore piloted both at an acute inpatient ward and a Crisis House. This was designed to ensure a range of diverse staff and service users were included in the pilot and their feedback about the clarity and acceptability of measures obtained.

CaRICE was piloted at an acute ward at the Royal Free Hospital in Camden and Drayton Park Crisis House in Islington. All staff involved in patient care at each service were asked to complete CaRICE each day at work over a five-day recording period. CCCQ-S and CCCQ-P were piloted with twenty patients and staff at acute wards at the Highgate Mental Health Centre in Camden and ten patients and staff at a Crisis House in Islington. Thirty patients were a sufficient number to identify any problems with clarity and acceptability. More patients were included at the acute wards to ensure several men were included in the pilot (as the Islington Crisis House accepted female residents only). Participating patients at the Crisis House were offered £15 each in acknowledgement of their time and effort. At the acute wards, service users were not offered money (to avoid disrupting another ongoing study in which participants were not being offered money).

At both pilots, respondents were asked not only to complete the measure, but also provide brief feedback about the clarity and acceptability of the measures and suggest any possible improvements. Written feedback forms were attached to staff-completed measures. Patients were asked for their feedback orally by the researcher after completion of the measure. Four outcomes from the pilots were considered to assess the need for changes to the measures:

a) the measures' response rates
b) respondents' feedback
c) the frequency with which care categories were used
d) the discreteness with which care categories were used

5.3.2 Results of piloting

The outcomes from piloting of measures are described in this section and resulting changes to measures highlighted.

Response Rate

Response rates were considered as an indication of feasibility. The response rates from the three measures piloted are shown in Table 5.2 below. (The figure for CCCQ-P represents the number of service users who completed the measure out of all those approached. All who initially agreed to participate did complete the measure.)

Table 5.2 Pilot Study Response Rates

	CaRICE	**CCCQ-P**	**CCCQ-S**
Acute Ward	**35%** (25/72)	**63%** (20/32)	**67%** (12/18)
Crisis House	**98%** (42/43)	**83%** (10/12)	**80%** (8/10)

It was not obvious why the CaRICE response rate was so much lower than in the Crisis House. Feedback from participating staff was that the forms were clear and not onerous to complete. The very high completion rate at the Crisis House also suggests that the measure can be easily completed by staff. Staff were aware of the pilot study: a researcher had publicised the pilot study to staff in advance and was either present in person or made phone contact at every shift handover during the recording period. The active support of the Crisis House service manager in the CaRICE pilot may have positively affected the staff response rate there.

In the pilot of CaRICE, staff were not identified by name on recording forms and handed in completed forms to a central collection point. It was therefore not possible to identify whether individual staff had completed CaRICE or not. The higher response rate from acute ward staff for CCCQ-S, where the individual respondent was identified by name, suggests that asking staff to complete forms anonymously may have negatively affected the response rate at the acute ward.

The mixed response rate from the CaRICE pilot does not suggest that increasing the recording period from the five days used in this pilot is feasible.

Change 1: A record will be kept of which staff at each service have been given CaRICE forms. Staff will be asked to put their name on completed forms. Non-responders will be reminded to complete the measure by a researcher. Researchers will be present at each service during the recording period as much as possible.

The response rates for the CCCQ (staff and service user versions) were acceptable. The results suggest that paying participating service users may have a positive effect on the response rate of CCCQ-P.

Feedback from Respondents

Feedback from patient respondents about acceptability and feasibility led to two changes in CCCQ at an early stage of the pilot.

i) Asking only whether someone has had any care of a particular category, yes or no, is very blunt: respondents wished to qualify their answer, for example saying that care was only available some of the time.

Change 2: An additional question was asked for each category of care in subsequent questionnaires, asking how frequently care had been received during the admission, using an eight-point scale.

ii) The category of care: “help to resolve or manage difficulties with current relationships or negative past events” needed to be divided in two: respondents may have received help with one but not the other.

Change 3: Two categories of care were used in subsequent questionnaires and the pilot of CaRICE: Help with Past Events and Help with Current Relationships.

All patients in the pilot expressed a preference for CCCQ-P to be completed as a structured interview (i.e. filled in by the researcher), rather than self-completed.

Staff who completed CCCQ-S did feel able to provide information about all the care the patient had received during his/her admission. One staff member who completed CaRICE expressed a view that it was not possible to describe clinical interventions accurately using the categories of care provided. Apart from this however, respondents of all measures reported that they were clear and acceptable.

The frequency with which categories of care were used

The extent to which categories of care in CaRICE and CCCQ-P were used by some but not all respondents and could therefore be considered discriminating and necessary was considered by looking at the frequencies of responses for each category in the measures. Table 5.3 summarises the frequency with which the categories of care in CaRICE and CCCQ were used by all respondents in the pilot study to describe the care provided at services.

Table 5.3 Categories of care: frequency of use

	Category	% interventions from CaRICE	% of patients receiving care CCCQ-S	% of patients receiving care CCCQ-P
1	housing	4.5	27	40
2	finances	.6	37	40
3	current activity	6.2	83	85
4	future activity	0	43	40
5	adl skills	2.8	23	30
6	criminal justice	0	3	0
7	other legal	1.7	23	20
8	family/carers	2.2	33	75
9	safety	23.2	63	85
10	assessment	7.0	73	95
11	relationships	2.5	47	90
12	past events	1.4	53	90
13	symptom coping	7.3	60	90
14	illness education	9.2	53	90
15	meds review/change	1.1	60	85
16	meds practical	11.8	67	70
17	meds concordance	2.2	53	60
18	physical health	2.0	53	50
19	drug/alcohol	.3	23	45
20	care planning meeting	4.2	63	80
21	discharge planning	.6	73	80
22	other	9.2	7	15

The CaRICE data suggest that a number of types of care were provided in a very small proportion of interventions by staff; CCCQ data however suggests that these categories of care were provided to a significant proportion of service users (at least 20%) in all cases except “help with criminal justice problems”. No category of care is reported by both staff and service users as having been provided to more than 85% of service users. “Help with criminal justice problems” is the only category of care which appears to be too rarely used to be able to identify potential differences in care provided at services. Other categories of care appear to discriminate service users who have and have not received interventions of this type.

Change 4: Remove “help with criminal justice problems” as a separate category of care and merge with “help with other legal issues”.

Inspection of CCCQ responses regarding frequency of provision of care showed that a few categories of care show ceiling effects for understandable reasons (e.g. medication reviews or care planning meetings are not provided more than three times a week to any patient). For many categories of care however, ratings of frequency of care spanned the range of possible responses. Few categories exhibited bi-modal distribution of responses. This suggested that the questions about the frequency of care which were added to the CCCQ were discriminating. It was decided to retain them.

The discreteness of categories of care used in the measures

Two investigations explored whether the categories of care in CaRICE and CCCQ could be used by respondents to distinguish different types of care, or whether any categories could be merged:

i) CCCQ respondents’ descriptions of the interventions provided for each category of care marked as provided to a patient were scrutinised: where these descriptions suggested more than one questionnaire category was being used to describe the same intervention, pilot questionnaire data was examined using crosstabs in SPSS to identify categories where receipt of care was most highly correlated for individual patients. This investigation suggested the following categories were least discrete:

a) “Help with future activity” and “help with discharge planning”
b) “Relationships”, “past events”, “illness education”, “symptom coping” all appear to overlap significantly. The feedback from pilot respondents suggests the distinctions between the different sorts of help do have face validity; however the categories need to be more distinctly defined if possible.
c) The category “assessment” was used in conjunction with a number of other categories in describing particular interventions. Although this raises questions about whether the category “assessment” is sufficiently clearly defined, it is understandable that patients receiving a number of types of help are very likely to have been assessed too. Moreover, the significant proportion of patients in the questionnaire pilot reporting not receiving assessment suggests the category does have potential value in distinguishing differences in care provided to patients at different services. As an attempt to improve definition of “assessment”, respondents were directed to mark interventions using the category “assessment” if this is a major part of the intervention, rather than a concurrent, secondary part of another intervention.

ii) An initial, small exploration of the inter-rater reliability with which clinicians can describe contacts with patients using the categories of care in CaRICE and CCCQ-S was made. Ten mental health clinicians used the categories of care from the measures to code vignettes drawn from anonymised case note reports of contacts between inpatient mental health staff and patients. Care categories where responses varied most were “assessment”, “help with relationships”, “help with past events”, “safety”, “discharge planning” and “help with future activity”. This confirmed the need for changes to the measures suggested by the previous investigation described above.

Some clinicians completing the small-scale reliability test commented that the category of “help with safety” covered a wide range of possible interventions, including physical restraint, 1:1 observations or simply checking someone’s wellbeing. It was suggested that clear definition of this category would make the measure easier to complete and more reliable.

Change 5: The category of care "help with discharge planning" to be defined more specifically to refer to help planning care post discharge from mental health services.

Change 6: The category of care "help with future activity" to be defined more specifically in terms of help with work and leisure activities only.

Change 7: The categories of care "assessment", "help with past events", "relationships", "symptom coping", "illness education" and "drug and alcohol problems" all to be prefixed with "help mainly focused on". This is designed to guide respondents to pick the most appropriate of these categories where possible to describe one intervention, rather than a combination of them.

Change 8: The category of "safety" was divided into two categories: "physical restraint" and "continuous or regular observation".

The changes made to CaRICE and CCCQ as a result of piloting the measures are summarised in Table 5.4.

Table 5.4 Summary of changes to measures post-pilot

CaRICE only	CCCQ only	CaRICE and CCCQ
1. To be completed by named respondents, not anonymously	**2.** Additional question about the frequency with which each type of care is provided	**3.** Category: "help with past events or current relationships" divided into two categories **4.** Categories: "help with criminal justice problems", "other legal help" merged into "help with legal matters" **5.** Category: "help with discharge planning" defined more clearly **6.** Category: "help with future activity" defined more clearly **7.** Categories referring to types of psychological help all qualified with "help focused mainly on" each type of help. **8.** Category of help to ensure safety split into two more clearly defined categories referring to restraint and observation

The revised categories of care used in CaRICE and CCCQ are shown in Table 5.5. A brief name for each category is also provided, which will be used in this thesis from here on. Amended versions of the measures were circulated to the

Alternatives Study Steering Group (n = 25). No further amendments were suggested, supporting the face validity of the measures to a group including clinicians, user-researchers and researchers.

Table 5.5 Revised categories of care for CaRICE and CCCQ

Purpose of contact	Description	Brief Name	Subscale
1	Help with housing problems, (e.g. help finding somewhere to live or making current housing more suitable)	Housing	Social interventions
2	Help with financial problems (such as claiming benefits or managing debts)	Finances	Social interventions
3	Help with legal matters (e.g. providing a letter or report for the court or his/her solicitor, providing information about his/her legal rights or help to access legal services)	Legal	Social interventions
4	Current activity: help to plan or engage in social, leisure, occupational or religious activities while staying at this service	Current activity	Social interventions
5	Future activity: help to plan or access work, education, social, leisure or religious activities when he/she leaves this service	Future activity	Social interventions
6	Help for him/her to practice or improve skills in every day tasks (e.g. managing shopping, cooking using a washing machine, self care)	ADL help	Social interventions
7	Contact between staff and his/her family, friends or carers to help support him/her and them	Family	Social interventions
8	Assessment of his/her difficulties (help focused mainly on asking him/her about the nature or severity of current problems or his/her life history)	Assessment	General care organisation
9	Help (other than medication) focused mainly on finding ways to cope with distressing feelings, thoughts and experiences (e.g. low mood, panic attacks, intrusive or strange thoughts or hearing voices)	Coping skills	Psychological interventions
10	Help focused mainly on resolving or managing difficulties with current relationships	Relationships	Psychological interventions
11	Help focused mainly on resolving or managing difficulties with negative or traumatic past events	Past Events	Psychological interventions
12	Help focused mainly on problems he/she is experiencing with drug or alcohol use	Drugs and alcohol	Psychological interventions
13	Help focused mainly on providing him/her with information or explanation about the nature of his/her mental health or psychological problems or his/her diagnosis	Illness education	Psychological interventions
14	Help with concerns or questions he/she has about his/her medication (e.g. providing information about treatment options or side effects, discussing advantages or disadvantages of medication or problems he/she is experiencing)	Medication concordance	Physical and pharmacological interventions

Purpose of contact	Description	Brief Name	Subscale
15	A review or change of his/her current medication *(only include medication for mental health problems or side-effects of mental health medication in this category)*	Medication review	Physical and pharmacological interventions
16	Practical help with taking medication (such as staff dispensing medication or giving an injection, or reminding him/her to take medication)	Medication practical help	Physical and pharmacological interventions
17	Help with his/her physical health (treatment, investigations, tests or help to access physical health services) *(Don't include prescribed medication for mental health problems or side-effects from mental health medication in this category)*	Physical healthcare	Physical and pharmacological interventions
18	A member of staff staying with him/her continuously or at regular intervals to make sure he/she or others are safe at times when he/she has been distressed or disturbed.	Observations	Physical and pharmacological interventions
19	A member of staff physically restraining him/her to make sure he/she or others are safe at times when he/she has been distressed or disturbed.	Restraint	Physical and pharmacological interventions
20	A meeting for him/her and everyone involved with his/her care to discuss his/her current needs and concerns	Care Planning	General care organisation
21	Help to plan or arrange care from other mental health services once he/she leaves this service	Aftercare	General care organisation

5.4 Psychometric explorations

Four explorations of the psychometric properties of the measures developed for this study were conducted: a) an inter-rater reliability test of CaRICE and CCCQ categories; b) an inter-rater reliability test of CCCQ-S; c) an exploration of concordance of CCCQ-S and CCCQ-P; and d) an exploration of the convergent validity of CaRICE. The first two of these explorations were conducted post-pilot but before data collection for the quantitative study described in Chapters 6 – 10 of this thesis. The latter two explorations were conducted retrospectively with data collected from the study.

5.4.1 Inter-rater reliability of CaRICE

Following amendments to the measures post-piloting, an investigation of whether the revised categories of care in CaRICE and CCCQ allow clinicians to code real patient contacts with good inter-rater reliability was undertaken.

Method: Contacts between staff and patients documented in patients' case notes in acute wards in a North London psychiatric hospital were anonymised and used to create 21 vignettes, each describing a staff-patient contact. The vignettes were designed to include all the types of care described in CaRICE/CCCQ categories. Mental health clinicians were then asked (using similar guidance to that provided in CaRICE) to describe the type(s) of care provided in each vignette, using the 21 categories of care from CaRICE. For each vignette, the number of respondents who used and didn't use each category of care was entered onto a computer database and analysed using Stata software (StataCorp 2007). An overall kappa value for rater concordance was calculated using the "kappa pos neg" command in Stata. Descriptive data were also provided for the proportion of respondents using the most popular categories in each vignette and the spread of responses.

Results: 21 clinicians, including psychiatrists, nurses, psychologists, social workers and unqualified staff providing patient care in acute settings, coded the vignettes. The frequency with which CaRICE categories of care were used by respondents is presented in Table 5.6 overleaf. The kappa value for overall concordance among raters was calculated as **kappa = 0.71.**

The mean number of respondents using the most popular category in each vignette was 19 (91%), with a range of 14-21 respondents. The mean number of respondents using the next most popular category in each vignette was 4 (19%), with a range of 0-11. A mean of 2.6 categories of care were used per respondent per vignette. A mean of 3.5 categories (range 1-7) were used by at least one respondent per vignette.

Table 5.6: CaRICE inter-rater reliability test responses
(Total number of respondents = 21)

		Category of Care																				
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Vignette	**1**	0	0	0	0	2	11	0	0	2	0	0	0	0	0	0	0	14	0	0	0	0
	2	0	0	0	18	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0	19	4	0	0	0	0	0	0	0	0	1	0	1	0
	4	0	0	0	0	0	0	4	1	0	0	0	0	0	0	0	0	0	0	0	3	19
	5	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	3	18	0	0
	6	0	0	0	0	0	0	0	0	0	0	0	1	0	18	4	4	0	0	0	2	0
	7	0	0	0	3	0	0	0	1	20	0	0	0	0	0	0	0	0	0	1	0	0
	8	0	0	0	6	1	0	21	0	0	2	0	0	0	0	0	0	0	0	0	0	0
	9	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0
	10	0	0	0	0	1	0	0	5	4	0	1	0	19	0	1	0	0	0	0	0	2
	11	0	0	0	0	0	21	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
	12	0	0	0	0	0	0	0	2	0	0	0	21	0	0	1	0	1	0	0	0	0
	13	0	0	0	0	0	0	0	0	0	0	0	3	0	1	16	1	0	0	0	1	0
	14	0	0	21	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	15	0	0	0	0	0	0	0	7	2	0	0	0	0	0	1	0	0	19	0	0	1
	16	0	0	0	4	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
	17	0	0	0	0	0	0	3	0	0	21	1	0	0	0	0	0	0	0	0	0	0
	18	0	0	0	0	0	0	0	2	9	1	17	0	0	0	0	0	0	0	0	0	0
	19	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	20	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	3
	21	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Discussion

There are methodological limitations to this investigation of the inter-rater reliability of CaRICE. Vignettes based on case note reports may inadequately reflect the complexity of real-life interactions between staff and patients and thus pose different or fewer challenges to raters than using CaRICE in vivo. The possibility of social desirability bias - deliberate distortion by staff when completing CaRICE (for example, to show themselves or their service in a favourable light) is not tested by this vignette exercise.

There are also difficulties in interpreting results. A kappa value of 0.71 represents good inter-rater reliability, being between 0.61 and 0.8 (Altman 1991). However kappa values are influenced not just by the level of agreement between raters but also by the degree of asymmetry in ratings (Feinstein and Chicchetti 1990). Asymmetry has two components: prevalence and bias. Prevalence refers to the number of positive yes/yes agreements compared to negative no/no agreements. Bias refers to variation in the propensity of raters to make positive and negative ratings, i.e. the number of yes/no disagreements compared to no/yes disagreements. For the same overall level of agreement between raters, high asymmetry in agreement categories (yes/yes vs. no/no) will markedly reduce kappa scores; high asymmetry in disagreement categories (yes/no vs. no/yes) will slightly raise kappa scores (Feinstein and Chicchetti 1990, Lantz and Nebenzahl 1996). In this investigation of the reliability of CaRICE, large numbers of negative ratings (i.e. categories of care not being used to describe vignettes) compared to positive ones were generated. This high level of asymmetry in agreement categories (more no/no agreements than yes/yes) means the kappa of 0.71 achieved by CaRICE may be lower than expected for the overall level of inter-rater agreement in the investigation.

Dunn (1989) proposed that interpretation of kappa may be assisted by also reporting an adjusted kappa statistic: kappa max. This represents the maximum value of kappa which could be attained for a set of data within the restraints of the marginal totals, i.e. given levels of asymmetry within the data. Kappa max can be particularly useful in investigations of the agreement between two diagnostic tools or measures, where a comparison of kappa and kappa max results can help distinguish the level of concordance between the measures from the effects of variation in raters' propensity to make positive ratings or the prevalence of the condition in the sample. Worden and colleagues (2008), for example, reported both kappa and kappa max for a comparison of the concordance between two measures of needs in an elderly mentally ill population.

Feinstein and Chicchetti (1990) have criticised kappa max however because it can produce more inflated scores than kappa for data with high asymmetry in disagreement (yes/no vs. no/yes). Hoehler (2000) and Chicchetti and Feinstein (1990) have argued that the penalising effect of prevalence and bias on the value of kappa is appropriate. This is relevant to CaRICE, where (unlike a positive/negative diagnostic tool for a single condition) the low overall proportion of positive ratings is a function of the measure, not just the sample. The overall proportion of positive ratings in CaRICE is low because CaRICE offers a large number of categories to choose from to describe care events. Variation between raters in the use of multiple categories to describe care events (bias) will affect the consistency with which CaRICE can be used. The infrequency with which CaRICE categories are used to describe care events (prevalence) necessitates high levels of inter-rater agreement to allow meaningful comparison of services' provision of individual types of care. Taking account of prevalence and bias in an assessment of the reliability of CaRICE is therefore appropriate.

A second approach to refining analysis using kappa is to calculate weighted kappa scores (Dunn 1989). This can be useful with categorical or ordinal data with three or more categories, allowing disagreements between raters which are large or perceived as more serious to be given greater emphasis than others. CaRICE categories however were designed to reflect types of care which were all distinct from each other: there is no evident basis on which to prioritise agreement in some categories above others. Moreover, weighting kappa (because it requires multiple categories) would require agreement to be assessed for each vignette, rather than for each category of care within each vignette. Defining whether raters have agreed or not is not clear cut at vignette level because respondents use different numbers of CaRICE categories to describe vignettes without prioritising primary and secondary descriptions. For these reasons, weighted kappa was not used in this investigation.

Chicchetti and Feinstein (1990) recommend that, despite its limitations, if a single summary statistic is wanted to measure concordance between raters, kappa should be used. This level of analysis is proportionate given the small scale of this investigation and the limitations of using case note vignettes. Kappa is therefore the statistic reported in this investigation.

Possibly as informative as a kappa value is descriptive report of inter-rater concordance. In this investigation, the concordant use of one primary category of care per vignette by over 90% of respondents and fairly infrequent use of secondary categories to describe vignettes are an indication of a promising degree of shared understanding of CaRICE categories by raters. While its limitations are acknowledged, this investigation suggests CaRICE exhibits good inter-rater reliability. Consistent understanding of how to use the categories to describe interventions is the major factor necessary to enable staff to use CaRICE reliably. (Clinicians use CaRICE to describe their own actions almost contemporaneously, so recall bias or knowledge of what care has been provided are of less concern.) Using case vignettes to explore inter-rater reliability follows the method used by Brekke (1987) with the Daily Contact Log – the existing event record which has been most robustly tested regarding psychometric properties. CaRICE’s inter-rater reliability in this test compares favourably with that of the Daily Contact Log (Brekke 1987), the only other event recording measure to have been similarly tested, for which kappa values of 0.59 and 0.68 have been reported. This investigation therefore provides no evidence that the greater depth of information and greater number of categories of care in CaRICE, compared to previous content of care measures, has compromised its reliability.

This test also has implications for the inter-rater reliability of CCCQ-S. However, the extent to which respondents used different categories to describe similar interventions is not the only source of unreliability of CCCQ. As CCCQ is retrospective, problems with imperfect recall and, in the staff version, knowledge of what care was provided by all clinicians at a service, will also affect inter-rater reliability. Conclusions about the inter-rater reliability of CCCQ therefore cannot be drawn from this vignette exercise.

5.4.2 Inter-rater reliability of CCCQ-S

Exploration of the inter-rater reliability of the CCCQ-S can feasibly be undertaken in an actual clinical setting, by asking two or more clinicians to complete questionnaires for the same patient at discharge. The results of such an exploration are presented here. CCCQ-P cannot be investigated in this way as each patient alone experiences his own care.

Method: For 46 patients at two services participating in the Alternatives Study for whom a CCCQ-S had been completed, a second member of staff at the service who also knew the patient (e.g. an associate nurse or other staff member involved in planning the patient's care) was asked to complete a CCCQ-S form. Cohen's kappa was calculated for the inter-rater reliability of binary responses to whether any care of each category had been received by each patient. Spearman's Rho was calculated from ordinal data regarding the agreement about the frequency with which care was received by each patient. Consistent with existing guidance (Altman 1991), Kappa and Spearman's Rho values above 0.6 or above were considered to demonstrate good reliability.

A second analysis was undertaken comparing levels of inter-rater reliability at the two services from which respondents were drawn. These had markedly different lengths of stay (Alternative 3 mean stay = 6.25 days; Standard 4 mean stay = 30.62 days). It was hypothesised that if staff recall affects the reliability of the CCCQ-S, data from Standard service 4, requiring recall of care provided over a longer period on average, would demonstrate poorer inter-rater reliability than data from Alternative 3.

Results

CCCQ-S overall demonstrated poor levels of inter-rater reliability. Only 4 out of 21 categories produced adequate kappa scores; 3 out of 21 produced adequate spearman's rho scores. The results are presented in Table 5.7.

Table 5.7 CCCQ-S inter-rater reliability test

	Was care received n = 46			How frequently? (ordinal scale 0 – 7)		
	Rater 1 yes	Rater 2 yes	Kappa	Spearman’s Rho	Rater 1 median	Rater 2 median
1 housing	14	11	.618	.681	0	0
2 finances	12	13	.506	.531	0	0
3 legal	11	4	.312	.348	0	0
4 current activity	13	26	.401	.352	2	3
5 future activity	24	22	.129	.179	0	0
6 a.d.l.	32	16	.605	.603	0	0
7 family	10	28	.207	.307	2	2
8 assessment	43	40	.148	-.047	4	4
9 coping strategies	37	42	-.137	.189	5	5
10 relationships	19	17	.362	.336	0	0
11 past events	19	19	.191	.191	0	0
12 drugs alcohol	23	19	.652	.661	0	1
13 illness education	32	28	.335	.578	1	3
14 meds concordance	35	30	.328	.386	2	3
15 meds review	30	26	.276	.127	1	1
16 meds practical	33	29	.313	.399	2	6
17 physical	16	9	.306	.415	0	0
18 observations	19	13	.247	.238	0	0
19 restraint	2	4	.292	.329	0	0
20 care planning	35	24	.244	.379	1	1
21 aftercare	37	40	-.027	-.085	1	1

Inter-rater reliability was not found to be poorer at the service with the greater length of stay (Standard 4). Comparison of reliability at the two services is provided in Appendix 6 (Table A6.1).

Summary

This investigation suggests that CCCQ-S cannot be used by staff to provide reliable data about the care provided to individual patients during an admission. The reasons for this are not wholly clear but:

i) the exploration using vignettes reported in Section 5.4.1 suggests that the categories of care are not unreliable per se; staff can reliably describe interventions using these categories

ii) the comparison between services with different lengths of stay (Appendix 6) suggests that it is not the retrospective nature of CCCQ-S which makes it unreliable.

It is therefore speculated that the CCCQ-S may demonstrate poor inter-rater reliability because one member of staff at an inpatient service (even a named nurse or key worker) does not have an accurate overall impression of all the care provided to a patient during an admission.

The results of this exploration suggest that CCCQ-S cannot be used to compare the intensity and nature of care provided at alternative and standard inpatient mental health services.

5.4.3 Concordance between CCCQ-S and CCCQ-P

Method: For 108 patients in 8 services in The Alternatives Study for whom CCCQ-S and CCCQ-P data were collected, inter-rater reliability between patient and staff responses was explored using the same method outlined in Section 5.4.2 above for the CCCQ-S inter-rater reliability test.

Results: Very low levels of agreement were found between staff and patient respondents about care received: no adequate Kappa or Spearman’s Rho values were found for any of 21 types of care. The data suggests a trend (in 17 out of 21 categories) for patients to report receiving less care than staff report being provided. The results are presented in Table 5.8.

Table 5.8 CCCQ-S and CCCQ-P concordance

	Was care received n = 108			How frequently? (ordinal scale 0 – 7)		
	Patient yes	Staff yes	Kappa	Spearman’s Rho	Patient median	Staff median
1 housing	24	35	.332	.360	0	0
2 finances	20	30	.280	.299	0	0
3 legal	14	12	.301	.315	0	0
4 current activity	48	56	-.070	.003	0	1
5 future activity	33	51	.054	.023	0	0
6 a.d.l.	19	33	.104	.102	0	0
7 family	49	73	.238	.167	0	3
8 assessment	77	80	.021	.156	3	4
9 coping strategies	46	88	-.016	.041	0	4
10 relationships	15	35	.250	.346	0	0
11 past events	34	48	-.010	-.067	0	0
12 drugs alcohol	21	43	.404	.473	0	0
13 illness education	45	59	.052	-.018	0	1
14 meds concordance	49	63	.153	.237	0	1
15 meds review	57	58	.424	.396	1	1
16 meds practical	87	63	.134	.128	7	5
17 physical	50	32	.155	.178	0	0
18 observations	34	29	.084	.060	0	0
19 restraint	7	12	.484	.510	0	0
20 care planning	58	76	.264	.242	1	3
21 aftercare	56	81	.086	-.026	1	1

Summary: Staff and patient reports of care provided differed. Three possible explanations are:

- CCCQ-P data provide a reliable measure of the care provided to individual patients, which differs from the unreliable CCCQ-S data
- Patients may not interpret the categories of care in CCCQ-P reliably and may differ from clinicians in their interpretation. Patients may therefore describe interventions received unreliably and differently from staff.
- There may be real differences between patients' experience of receiving care and staff's perception of providing it: e.g. a patient may not think he has received care, even though a member of staff believes that is what he has been providing.

Had there been high levels of concordance between the primary staff respondent and patient respondent, this might have indicated that the CCCQ-S unreliability identified in Section 5.4.2 was due to unreliable ratings by the second staff respondent who perhaps did not know the patient so well. However, this investigation provided no evidence to support this explanation. Overall, the use of CCCQ-S is not supported by this investigation.

CCCQ-P data must be interpreted with some caution as its reliability is uncertain and this exploration suggests staff and patient perspectives may be significantly different.

5.4.4 Convergent validity of CaRICE

The convergence of intensity of care data from CaSPAR and CaRICE gathered in this thesis was explored. The review of content of care measures in Chapter 4 identified no gold standard measure of intensity of staff-patient contact. CaSPAR, however, being based mainly on researcher-observation, used the data collection method most clearly demonstrated to provide reliable data and was less vulnerable to social desirability bias than staff report measures. It was therefore appropriate to use CaSPAR data as a benchmark to which CaRICE data may be compared. High levels of convergence between data from the two measures would increase confidence that both are reflecting the reality of how much patient contact is provided in services and increase confidence in CaRICE's validity.

Method:

Descriptive data from individual services were presented for the following variables:

- CaSPAR: proportion of patients in contact with staff
- CaRICE: minutes' contact from staff per patient per day
- CaRICE: proportion of staff time in contact with patients

Visual inspection of data was used to assess levels of convergence of data between measures. Additionally, analyses of variance among individual services were conducted for each variable using Duncan's Multiple Range Test, a procedure to test for homogeneity in subsets. The test was used in this study to identify, for each variable assessed, clusters of services where differences within in the range of scores were not significant. Duncan's test thus helps understand the significance of differences in rankings of services in data from the different measured variables.

Results:

Descriptive data from CaSPAR and CaRICE for mean scores for individual services are presented in Table 5.9.

Table 5.9: Comparison of CaSPAR and CaRICE data: individual services

CaSPAR: proportion of patients in contact with staff (%)		CaRICE: minutes contact per patient per day		CaRICE: proportion of staff time spent in contact with patients (%)	
4.55	Alternative 3	82.56	Standard 4	11.43	Alternative 1
7.97	Standard 1	109.59	Alternative 4	15.86	Standard 4
8.55	Alternative 1	121.74	Standard 1	21.57	Standard 1
8.77	Standard 4	131.71	Standard 3	23.34	Alternative 4
10.75	Alternative 4	133.33	Alternative 1	25.00	Alternative 2
13.52	Standard 2	139.5	Alternative 2	26.71	Standard 2
14.05	Standard 3	154.12	Standard 2	31.19	Alternative 3
21.60	Alternative 2	160.37	Alternative 3	34.04	Standard 3

Descriptive data show that rankings of services exhibit some variation across different measures. This variation is most marked for Alternative 3, which ranked bottom in CaSPAR data but high in both CaRICE measures. Duncan's multiple range tests conducted on CaSPAR, CaRICE minutes' contact per patient per day and CaRICE proportion of staff time spent in patient contact data are presented in Appendix 6 (Tables A6.2 – A6.4). Duncan's test of CaSPAR data (Table A6.2) indicated that one service, Alternative 2, could be distinguished from the seven others, among which differences in the range of scores were not significant. Duncan's tests of CaRICE data (Tables A6.3 and A6.4) did not provide such a clear clustering of services: no single service was distinguished from all the others and a number of clusters of services with non-significant differences in the range of scores could be identified in data from both CaRICE variables.

Summary:

Service data from CaSPAR and CaRICE reveals that there is considerable divergence in results from the two measures. This investigation does not explain this divergence. Possible explanations will be discussed in Section 10.1.

One factor which did not explain the level of divergence of CaRICE data from CaSPAR data was unrepresentativeness of CaRICE data due to low response rates. CaRICE response rates were high (95%) with little variation between services (range 91-100%). CaRICE data regarding the proportion of staff time spent in contact with patients was not clearly more convergent with CaSPAR data than was CaRICE data for the minutes' contact provided per patient per day. This, coupled with the high response rates for CaRICE, supports the appropriateness of using the minutes' contact per patient per day as the CaRICE variable of content of care in this thesis.

This investigation does not provide CaRICE with convincing convergent validity and provides a caveat about using it as a sole measure of content of care in inpatient services.

5.4.5 Discussion: psychometric explorations

The four explorations reported above represent an investigation of the psychometric properties of the measures developed for this study. Other tests of reliability and validity were considered but rejected on grounds of feasibility or relevance.

Reliability

a) Inter-rater reliability: inter-rater reliability testing of CaSPAR was desirable but not prioritised given limited resources. The measure records only whether staff were in contact with patients, not the nature of contacts: previous studies have demonstrated that this information can be collected reliably by researcher observation (Tyson et al. 1995) (Paul 1987) and staff report (Brekke 1987).

b) Test-retest reliability: Test-retest reliability could not be assessed for CaSPAR, which uses momentary time recording. Ideally, this would have been assessed for CaRICE and both CCCQ versions. However, the consistency with which one person uses the measures was considered less likely to be a source of major unreliability than differences in use by different people. Given limited resources, investigating inter-rater reliability was therefore prioritised.

c) Split-half reliability: this is not a relevant test as none of the measures for this study use several questions to measure the same construct. CaSPAR uses the same method of observation throughout, while CaRICE and CCCQ measure different types of care.

Validity

Investigations of concurrent or predictive validity were not prioritised because of limited resources and the difficulties in achieving conclusive results. It was argued (Section 5.2.1) that purely observation-based measures of intensity of care which have demonstrated good reliability may be inadequately representative of what is provided at services through limitations of which contacts can be observed. The relationship between intensity of care and clinical improvement, met needs or needs might be confounded by patient variables, the appropriateness of care provided (quality of assessment) or the quality of the care provided. Failure to establish association between the new measure and plausible comparison measures therefore may not necessarily reflect poor validity in the new measure. Ideally, however, some comparison of measures developed for this thesis with established measures to which some relationship is likely would have been conducted: as was proposed for the comparison of CaSPAR and CaRICE (Section 5.3.4), evidence of associations between data from the two sources would have provided some evidence of validity.

The psychometric investigations of measures in this thesis suggest: a) staff can use CaRICE categories reliably to describe contacts with clients; b) there is some divergence in data about services provided by CaSPAR and CaRICE; c) staff-report measures should be limited to seeking information from staff about what they themselves have provided: individual staff members cannot provide reliable information about all the care provided to patients during an inpatient admission; and d) patient perspectives about care at services, provided by CCCQ-P, may differ markedly from those of staff. These investigations therefore reinforce the need for a multi-methods approach to measurement of content of care in inpatient services, including both consumer and provider perspectives. There is a need to triangulate results from different data sources before drawing conclusions about the care provided at services.

5.5 Strengths and weaknesses of the measures

It was decided not to use data from CCCQ-S in this study as the measure demonstrated inadequate reliability. Three measures of content of care were therefore used to compare alternatives and standard services in this thesis: CaSPAR (mean proportion of patients in contact with staff), CaRICE (mean minutes' contact per patient per day) and CCCQ-P (mean total care score). The limitations and strengths of these three measures collectively to assess the content of care in alternative and standard services are discussed below.

Brekke (1987) provided a descriptive taxonomy of dimensions of content of care needed to describe what is provided by services, identifying as important the nature, frequency, duration, scope, setting, style and competence of care. The three measures developed for and used in this study between them measure five out of these seven proposed dimensions. None measures the style of care provided at services or the competence with which interventions are carried out. The reasons for this are discussed below.

a) Style. The style of a service can plausibly be regarded as a distinct feature of a service rather than one element of the content of the care provided (Johnson & Salvador-Carulla 1998). A validated measure of the overall style of inpatient mental health services, The Ward Atmosphere Scale (Moos 1996), has been developed and used widely. Previous studies have also used researcher observation to measure the style of individual staff-patient contacts, e.g. rating them as accepting, tolerating or rejecting (Shepherd & Richardson 1979). As discussed in Section 5.2.1, however, the practical difficulties of observing a representative sample of contacts sufficiently closely for reliable inference of the nature of each contact are great in acute inpatient settings.

b) Competence. Measures of the competence or quality of care at a service require a clear theoretical basis or known influence on outcomes to validate the measurement criteria used (Tugwell 1979). There is a lack of established criteria on which to base valid measures of the competence of care in inpatient services. Obtaining valid measures of the competence of interventions would be likely to be further compromised by:
i) Lack of access to sufficient information about the nature of contacts (through lack of patient consent to observe a representative sample of staff-patient contacts in detail, or lack of detailed information available from patient records)
ii) Poor inter-rater reliability in staff or service user responses, consequent on the lack of clear, established criteria on which to assess competence

Two main limitations and two strengths of the measures developed can be identified. They can only provide a limited amount of information about the nature of care provided at services but this is greater than that obtainable from previous measures. Limited psychometric properties have been demonstrated for the measures but they allow a multi-methods investigation of care provided at services. These limitations and strengths are discussed below.

Depth and breadth of information provided

The measures only provide limited information about the nature of care provided at a service. For example, a staff member may use CaRICE to record providing help to a patient with problematic drug or alcohol use. The record will not reveal the use of which substance was being addressed, nor what specific sort of help was being provided (e.g. motivational interviewing, harm reduction strategies, health education, exhortations to abstain, relapse prevention planning). In common with existing measures of content of care (Chapter 4), the measures do not provide information about whether particular treatment modalities are being employed. More narrowly-focused therapy fidelity measures, such as developed for cognitive therapy (Vallis et al. 1986), are required for this: these would not provide general content of care measurement.

The need for feasibility and reliability in the measures dictates that more detailed information is not sought about the nature of care provided. As discussed in Section 5.1, information about the types of intervention being provided during staff-patient contacts cannot be inferred from researcher-observation, so staff or service user report measures are required. CaRICE requires respondents to scan 21 categories of care to pick one or more which describe the content of a contact with a patient. Feedback from pilot respondents and the pilot response rate suggests this could not be increased to allow more detail about types of care provided without jeopardising the practicality of the measure. Increasing the subtlety of distinctions between types of care in the measures might also reduce inter-rater reliability.

It is acknowledged that, because of these limitations, there may be important differences between services in the nature of care provided which the measures developed for this thesis are insufficiently sensitive to identify, or differences between services in the style or quality of care provided which will not be measured.

The measures can however provide more depth of information about the intensity and nature of care provided at acute residential mental health services than any of the previous measures identified in the literature review in Chapter 4. CaRICE will provide contemporaneous rather than retrospective information about the types of care being provided, which has not been obtained in previous studies of acute inpatient services. The categories of care used in CaRICE and CCCQ-P are specifically based on description of what care is provided at UK acute residential mental health services (Johnson et al. 2001). The measures developed for this study can provide a more comprehensive comparison of the content of care provided at alternative and traditional inpatient services than possible from previous measures and allow exploration of associations between process of care and patient outcomes.

Limited established psychometric properties

Previous studies of observation-based measures (Shepherd and Richardson 1979, Tyson et al. 1995) provide some confidence that CaSPAR is likely to provide reliable data about numbers of staff-patient contacts at recording times. An inter-rater reliability test of CaRICE categories suggests staff may be able to use CaRICE reliably to describe the types of care provided during contacts with patients and that its reliability may be comparable to a previous event recording measure which provides less depth of information (Brekke 1987). Both CaSPAR and CaRICE however are limited in how far robust psychometric properties have been established and confirmed through use in several studies and settings. The patient report version of the individual patient level measure – CCCQ-P – has no demonstrated psychometric robustness (Section 5.4.3). The staff report version – CCCQ-S – has demonstrated psychometric inadequacy (Section 5.4.2).

Multi-methods approach

As suggested in Chapter 4, the multi-methods approach proposed for measuring the content of care at services in this study provides advantages. Collectively, the measures incorporate assets from various methods of measuring the content of care. The use of researcher observation in CaSPAR provides as objective and reliable as possible a measure of the intensity of staff-service user contact at services. Staff-report used in CaRICE provides detailed contemporaneous information about types of care provided. Retrospective data collection in CCCQ-P allows collection of individual patient-level information over a longer time period. No single method could achieve all of this. The use of several information sources allows different perspectives to be included in the measures. Patients' perception of care is clearly important yet few previous measures have incorporated information provided by patients.

The variables measured by CaSPAR, CaRICE and CCCQ-P are related but not the same, creating potential problems for interpretation of results. The frequency of interventions has more influence on CCCQ-P scores than CaSPAR and CaRICE, which also reflect the duration of interventions. Brief contacts between staff and patients contribute to CaSPAR but not to CaRICE scores. CCCQ-P scores are more sensitive to the range of interventions provided than are CaSPAR or CaRICE. To a large extent, these differences reflect what information can feasibly be obtained from different data collection methods and information sources. They also reflect the lack of consensus identified in Chapter 4 about how best to measure content of care or which variables most validly describe the amount of care provided at services. The range of related variables measured by CaSPAR, CaRICE and CCCQ-P creates the possibility that using all three instruments to assess the care provided in services may produce some divergence in results which is hard to interpret. However, differences between services in care provision which fail to be replicated across measures are less likely than those that do to be substantial or clinically meaningful. Consistent findings from instruments measuring more than one related variable, using different methods and reflecting different perspectives, can allow more confident inference that identified differences between services are valid and important. In the absence of a gold standard measure of content of care for use in inpatient services, triangulation of data from a multi-methods approach is required. CaSPAR, CaRICE and CCCQ-P can provide this.

Chapter 6

Quantitative Study: Research Questions

The purpose of the quantitative study within this thesis was summarised in Chapter 1. Consistent with MRC guidelines for evaluation of complex interventions (Craig et al. 2008), a process evaluation of alternatives and standard services can help describe and distinguish services, develop service models and understand measured service outcomes. The study aims and hypotheses are presented fully in this chapter. Reasons are presented for the study's focus on the intensity and nature of care at services, for the decision to combine alternatives and standard services in analyses and for the investigation of the relationship of care received to patient satisfaction.

6.1 Aims and hypotheses

The study addresses three research questions:

a) Is there more staff-patient contact at alternatives than at standard services?
b) Do the types of care provided differ between alternative and standard services?
c) Can differences between alternatives and standard services in patient satisfaction be explained by the content of care provided?

The study aims to provide a quantitative comparison of the intensity and nature of care provided at four alternatives and four standard inpatient services and explores the relationship between care received and patient satisfaction. A multi-methods quantitative comparison of the intensity of staff-patient contact at alternatives and standard services will be used to test three hypotheses:

1. The proportion of patients observed in contact with staff is greater at alternatives than standard services. To test this hypothesis, CaSPAR scores will be compared for alternatives and standard services, adjusting for

clustering by service and recording point variables, in a regression analysis. Significance will be set at $p<0.05$.

2. Staff-reported intensity of care is greater at alternatives than at standard services. To test this hypothesis, the effect size for alternatives compared to standard services will be calculated from CaRICE total scores (the minutes of staff contact provided per patient per day). A medium effect size (mean difference is greater than half a standard deviation from all data (Altman 1991)) for alternatives compared to standard services will be required to corroborate the hypothesis.
3. Patient-reported intensity of care is greater at alternatives than standard services. To test this hypothesis, CCCQ-P scores will be compared for alternatives and standard services in a regression analysis, adjusting for clustering by service and patient characteristics. Significance will be set at $p<0.05$.

The nature of care provided at alternative and standard services will be compared using two measures. The following three hypotheses will be tested:

4. The staff-reported and patient-reported intensity of social interventions is greater in alternatives than standard services. This will be tested using CaRICE and CCCQ-P data. To corroborate the hypothesis, CaRICE subscale scores for social interventions will show at least a medium effect size for alternatives compared to standard services **and** CCCQ-P social interventions subscale score will be significantly greater ($p<.05$), adjusting for clustering by service and patient characteristics, at alternatives compared to standard services.
5. The staff-reported and patient-reported intensity of psychological interventions is greater in alternatives than standard services. This will be tested using CaRICE and CCCQ-P data. To corroborate the hypothesis, CaRICE subscale scores for psychological interventions will show at least a

medium effect size for alternatives compared to standard services **and** CCCQ-P psychological interventions subscale score will be significantly greater (p<.05), adjusting for clustering by service and patient characteristics, at alternatives compared to standard services.

6. The staff-reported and patient-reported intensity of physical and pharmacological interventions is greater in standard services than in alternatives. This will be tested using CaRICE and CCCQ-P data. To corroborate the hypothesis, CaRICE subscale scores for physical and pharmacological interventions will show at least a medium effect size for standard services compared to alternatives **and** CCCQ-P physical and pharmacological interventions subscale score will be significantly greater (p<.05), adjusting for clustering by service and patient characteristics, at standard services compared to alternatives.

The study will also provide exploratory analysis of relationships between care provided at services and patient satisfaction. CSQ scores will be compared at alternative and standard services. The effect on the relationship between service type and CSQ score of adjusting for clustering by service, patient characteristics and CCCQ-P variables will then be explored using linear regression.

6.2 Rationale for the focus on intensity and nature of care

Intensity of care has been identified as an important element of inpatient service provision in theory, policy and by service users (Section 1.6). The review in Chapter 4 also identified however that previous (mainly community-based) studies have not consistently found the intensity of care provided to be associated with service outcomes. The nature of service interventions may be more important. This informed the decision in this study to additionally measure the nature of care provided at services. Measuring the provision of social, psychological and physical/pharmacological interventions at alternatives and standard services could illuminate whether the nature of care provided was broadly different at alternatives compared to standard care. The descriptive data about provision at alternatives and standard services of all 21 categories of care measured by CaRICE and CCCQ-P offers hypothesis-generating data about the nature of care provided and differences between alternatives and traditional services.

The intensity and nature of care at services are not the only process elements which may differ between alternatives and standard services or influence service effectiveness and patient outcomes. However, Johnson and Salvador Carulla (1998) advocated measurement of service content to describe and distinguish services, emphasising its high face validity and arguably greater importance than the setting or organisation of services. The literature review in Chapter 3 and the Alternatives Study UK survey (Johnson et al. 2009) found that little is known about the content of care in alternatives and how this differs from standard inpatient services. Given its resource constraints, the focus in this thesis on variables relating to service content – intensity and nature of care - is therefore appropriate.

CaSPAR, CaRICE and CCCQ-P all measure intensity of care; CaRICE and CCCQ-P both measure of the nature of care. The measures therefore allow a multi-methods investigation, proposed in Chapters 4 and 5 as desirable on account of differences in perspective and in the specific variables measured by each instrument. Corroboration was sought for three hypotheses regarding the intensity of care at services. Hypotheses about the nature of care at alternatives and standard services were framed to require corroboration from both CaRICE and CCCQ-P data (Hypotheses 4-6, Section 6.3). This triangulation of information sources and data collection methods increases confidence that positive results from the study would reflect important differences in service provision.

6.3 Combining services in analyses

The services included in the quantitative investigation comprise four alternatives and four standard services. They are described in Section 7.1. The alternative services represent three different models of community alternative and one inpatient alternative. A decision was taken to combine data from all four alternatives in comparisons with standard services. The models used in all four alternatives have similarity in aims for the nature and quality of care provision and difference from standard acute care. The Tidal Model, the model of care used in the inpatient alternative, has explicitly stated aims to provide different care from standard acute wards in the theoretical model. Barker (2001) describes the model as seeking to increase contact and collaboration between staff and patients and to provide holistic care which avoids reducing care to narrowly defined treatment of symptoms. Greater collaboration between staff and patients and user-focus has been proposed as a benefit of community-based alternatives (Faulkner et al. 2002). Operating outside a hospital environment and employing fewer staff from medical or nursing professions, community alternatives might be expected to provide less medically-focused care to patients than standard acute wards. For these reasons, the alternatives were combined for the main study hypotheses but heterogeneity among alternatives was examined and secondary sub-group analyses were considered.

6.4 Rationale for providing a model of satisfaction with services

The review of the effectiveness of alternatives reported in Chapter 3 identified patient satisfaction as the only outcome domain where the difference between (community-based) alternatives and standard services has been shown to be significant. For alternative services, which are not clearly manualised or model-driven and where the content of care has seldom been measured, there is little indication of what may be responsible for differences in acceptability compared to standard services. Using data from this study to investigate associations between content of care variables and patient satisfaction with services may highlight active elements of care which impact on the acceptability of services.

This study will therefore examine whether findings from previous research that alternatives are more acceptable to patients than standard services are replicated. The effect of adjusting for patient characteristics and content of care variables on the relationship between service type and patient satisfaction will then be explored. This will provide indications of whether the content of care provided has a potential explanatory role in a model of satisfaction with alternatives and standard inpatient services.

Data from CCCQ-P will be used for this exploratory analysis because it provides information about care received by individual patients. All four CCCQ-P variables for which data is available in this study will be used: total care score physical/pharmacological, psychological and social care subscale scores. The Client Satisfaction Questionnaire (CSQ) will provide satisfaction data. Developed by Attkisson and Zwick (1982), the CSQ is a patient-completed questionnaire providing a global measure of patient satisfaction with services. It has the benefits of being brief (8 items) and having established good psychometric properties. Given the lack of existing knowledge about associations between process variables and satisfaction with inpatient services and the exploratory nature of the investigation in this thesis, the global nature of the information provided by CSQ was considered sufficient.

Chapter 7

Quantitative study: methods

7.1 Setting

Data were collected from four alternative services and four comparison local standard acute inpatient services. Each alternative service represented one particular type of alternative identified in The Alternatives Study UK national service mapping, reported in Section 2.4. The services are listed in Table 7.1 and described below.

Table 7.1 Services in the study

Location	Alternative service	Standard service
Hackney, London	Alternative 1 non-clinical crisis house	Standard 1
North Staffordshire	Alternative 2 clinical crisis house	Standard 2
Middlesbrough	Alternative 3 crisis team beds	Standard 3
Birmingham	Alternative 4 Tidal Model ward	Standard 4

Alternative 1 (non-clinical crisis house) is a nine-bedded crisis house in a residential street in Hackney, run by a voluntary sector Housing Association. It accepts patients from African and Caribbean black minority ethnic communities, who are highly represented in the local community. The service has been running for ten years. It explicitly aims to provide a culturally sensitive alternative to hospital admission and will admit patients only from the community not transferred from acute wards. Detained patients cannot be admitted. Staff are non-clinical social care workers but a counsellor and alternative therapists such as a reflexologist also provide sessional input. Any required medical care is provided to patients by their general practitioners or through the local crisis and home treatment team. Aftercare is provided by the voluntary sector service provider, but liaison with statutory mental health services is also common. An initial two week limit is set for admissions, but two additional weeks can be agreed if considered necessary by the patient and staff.

Alternative 2 (clinical crisis house) is an eight-bedded residential unit within a Community Mental Health Resource Centre in Staffordshire. It was established ten years ago and is one of five similar units within the local mental health trust designed to avert admissions to the local psychiatric hospital where possible and facilitate early discharge by transferring patients from the acute wards. The service is situated in a rural village about fifteen miles from Stoke-on-Trent in a very mono-ethnic white British area. It is staffed similarly to a standard acute ward, i.e. mainly by nursing staff with input from psychiatrists from the local Community Mental Health Team (CMHT). A daily structured programme of activity is provided within the unit, including an extensive gardening project. The CMHT gatekeep access to the beds: only known clients are admitted directly from the community, precluding the admission of people unknown to services via the police or accident and emergency units. Detained patients can be admitted directly from the community. Admissions are typically longer than for the other alternatives in the study: the service's records indicate a typical length of stay of about one month.

Alternative 3 (Crisis Team Beds) comprises four beds run by the local Crisis and Home Treatment Team, within a larger social services rehabilitation hostel. The service is situated on a residential street in a housing estate about a mile from Central Middlesbrough. It has been running for four years. Patients' basic daily care is provided by the hostel's social care staff with additional daily input from Crisis Team clinical staff, including regular scheduled time from psychiatrists and psychologists. Typical length of stay is about one week. Home treatment support is planned and provided by Crisis Team staff for patients following an admission to the crisis beds. Detained patients are not accepted directly from the community, but patients can be admitted from hospital under Section 17 leave.

Alternative 4 (Tidal Model ward) is a single twenty-bedded inpatient ward with attached outpatient unit in inner-city Birmingham. The Tidal Model has been implemented in the ward for about a year. The Tidal Model has been described in Section 2.3: as implemented at Alternative 4, it involves an expectation that daily written care plans will be agreed with patients and agendas set by patients will guide weekly ward rounds with medical staff. The service admits male and female patients: most staff are nurses or healthcare assistants but one occupational therapist is also employed on the ward.

The four comparison standard services are all general acute admission inpatient services within the same mental health trusts as the alternatives and covering a similar catchment area except where stated. **Standard 1** comprises two 20-bed mixed-sex acute wards in the mental health unit of a general hospital in Hackney, an inner city London borough. **Standard 2** is situated in a suburban area of Stoke-on-Trent and has three acute wards, two 18-bed single sex wards for lower dependency patients and one 15-bed mixed ward for higher dependency acute admission patients. All three wards were included in the study; all patients requiring general acute admission are admitted to one of the three. Standard 2 serves the whole of North Staffordshire, a larger area than Alternative 2, which includes urban as well as rural areas. **Standard 3** comprises two 25-bed wards, one male one female, within a larger psychiatric hospital. It is situated within walking distance of the Middlesbrough Crisis Team Beds about a mile from central Middlesbrough. **Standard 4** is a 22-bed mixed sex acute psychiatric admission ward within a general hospital in Solihull, a suburban area in the south of the region covered by Birmingham and Solihull Mental Health Trust, more affluent than the inner-city area covered by Alternative 4. All four standard services employed a staff mix typical of acute wards, i.e. predominantly nursing staff and health care assistants with psychiatrists and occupational therapists also represented. All were considered one of the mainstream acute inpatient services within the local service system.

Further details of the professional background of staff at services during CaRICE recording weeks will be provided in Section 8.1. Descriptive data of staff-patient ratios at services (and an exploratory investigation of associations between staffing levels and staff-patient contact) are provided in Appendix 7 and discussed in Section 10.3.3.

7.2 Measures

Three measures of content of care were used in each participating service:

1) CaSPAR: a measure of the proportion of patients in contact with staff which uses momentary time recording informed by researcher-observation and staff report. 28 recordings at each service each yielded data for the proportion of patients with staff.

2) CaRICE: a contemporaneous record of all direct patient contacts completed by staff. CaRICE recording forms were given to all staff at participating inpatient services and returned at the end of a shift/day at work. The duration and types of care provided for each patient contact are recorded. CaRICE yields data for each day at each service of the mean minutes of contact with staff per patient per day, and the minutes per patient per day of social, psychological and physical/pharmacological interventions.

3) CCCQ-P: a patient-completed retrospective questionnaire providing a measure of the intensity of overall care and of social, psychological and physical/pharmacological care provided during an admission. Data about frequency of care on a seven-point scale for each of 21 types of care were summed to provide a total CCCQ-P score from 0-147. Subscale scores were calculated for social interventions (0-49; seven items), psychological interventions (0-35; five items) and physical/pharmacological interventions (0-42; six items).

A description of the content of care measures is provided in Chapter 5. The measures themselves are found in Appendices 3, 4 and 5.

One outcome measure, the Client Satisfaction Questionnaire (CSQ) (Attkisson and Zwick 1982) was used in each service for exploratory investigations of the relationship between care received and patient satisfaction. The CSQ is a patient-completed questionnaire providing a measure of a patient’s satisfaction with a mental health service. An eight-item measure, it yields a total score from 8-32. The measure has been widely used in mental health services research and has demonstrated good psychometric properties (Attkisson and Zwick 1982).

Participants completing CSQ and CCCQ-P were also asked to give their date of birth, gender, ethnicity (using 16 categories taken from the UK national census (Office of National Statistics 2001)) and Mental Health Act status (detained or not detained).

7.3 Sample

Services

Alternative services were identified from Phase 1 of The Alternatives Study (Johnson et al. 2009), reported in Section 2.4. Typical services from different clusters of alternative from Phase 1 were chosen for this quantitative investigation. Services were also chosen to provide a geographical spread and mixture of urban, metropolitan and rural services. Managers of participating alternatives were approached to help identify the local standard acute hospital serving the same or similar catchment area.

Participants

CaSPAR: Information was sought for all patients resident at the service at the 28 momentary recording times at each service. Where staff report data was needed to supplement researcher observation in making CaSPAR recordings, the shift coordinator or equivalent was approached. 28 recordings were made at each service (n = 224).

CaRICE: All staff at work at participating services were included in collecting CaRICE data over a five day (Monday – Friday) recording period at each service. Visiting staff from other services, such as community mental health teams, were not included in CaRICE data collection.

CCCQ-P: 40 questionnaires were sought at each service (n = 320). All service users at or close to the point of discharge from the service were eligible for inclusion as study participants, except those whom service staff advised lacked capacity to consent to or complete the questionnaire.

4) CSQ: 40 questionnaires were sought at each service (n = 320). Eligibility criteria for participants were as for CCCQ-P.

7.4 Procedures

Recruitment and consent

a) Services

Managers of candidate participant services were approached and their provisional agreement to participate in the study was obtained. National, multi-site ethical approval was then obtained, local Principal Investigators were identified with help from participating service managers and local ethical and NHS Trust R and D approval were obtained. Researchers applied for honorary contracts with the NHS Trusts in which participating services were located. Managers of participating wards and services were then contacted to facilitate starting data collection. Researchers visited each participating service to present the study to the staff team before starting to collect data. Posters and information sheets about the study were also displayed prominently in the services.

b) Participants The recruitment procedures for participants for each measure were as follows:
CaSPAR: No individual consent was obtained for collection of CaSPAR data, which involved no direct patient contact or information about identifiable staff or patients. Posters were displayed prominently at participating services informing patients about the research and asking them to let staff or researchers know if they were bothered by the research. (The purpose of the research could be explained to patients or individual observations could be curtailed if necessary.) Researchers sought to include all resident patients in CaSPAR recordings.

CaRICE: Staff’s consent for CaRICE recordings was given through completion of CaRICE forms: no signed consent procedure was used. No patient consent was required for CaRICE, which involved no participation from patients and provided no data about identifiable individual patients.

CCCQ-P and CSQ: Individual written informed consent was obtained from patients completing CCCQ-P and CSQ following provision of an information sheet.

Researchers asked staff to identify patients close to discharge and judge whether they had capacity to consent and participate in the study. Staff approached these patients initially to ask if they were prepared to talk to a researcher about the study. If so, researchers provided information, answered questions and took written consent. If participants changed their mind during completion of data or failed to complete forms, consent was considered to have been withdrawn and forms were returned to the participant if possible or destroyed. Participants who completed the measures were each offered £15 in cash in acknowledgement of their time and efforts.

Data collection

CaSPAR: Guidance regarding completion of CaSPAR, shown in Appendix 3, was followed. For each recording, a researcher walked through the communal areas of the service observing how many patients were in contact with staff. The researcher then approached a member of staff and asked:
i) whether any staff were with patients in inaccessible areas of the service (e.g. in a patient’s room)
ii) how many patients were out of the unit, either in another part of the service or away from the service
iii) whether any of these patients were with staff
Whenever possible, the researcher would seek corroboration of staff report information (e.g. going to the occupational therapy room or the garden to observe whether patients were in contact with staff). When possible, observation was used to provide data; otherwise, staff report was used. If it could not be identified whether a patient was with staff or not their status was marked as not known on the CaSPAR recording form.

CaRICE: During the data collection period, a researcher was present at all times when staff were expected to start or finish work at the service (every shift handover and constantly between 9am-5pm as a minimum). Researchers distributed and collected recording forms individually from staff at the beginning and end of their shifts. They helped explain how to use the form whenever required. Researchers kept a record of every inpatient staff member at work each day and whether they had completed a CaRICE form. When a researcher left the ward, forms were left prominently in the staff office in the service. On returning to the unit, researchers checked with staff whether any other staff had been to the unit (e.g. an on call doctor) and whether a form had been completed.

CCCQ-P and CSQ: Once written consent was obtained, CCCQ-P and CSQ were completed with the patient as a structured interview in a private room. To minimise the number of patients who left the service without being approached to participate in the study, researchers visited services regularly (typically twice a week) and additionally contacted staff by phone to identify patients close to discharge and seek to recruit them. Numbers of patients participating and declining to participate were noted by researchers.

Data management

Paper forms from all measures were brought by researchers to their workplace and data were entered in SPSS version 14.0 software (SPSS 2006). CaSPAR, CCCCQ-P and CSQ data were transferred from SPSS version 14.0 to Stata IC version 10.0 software (StataCorp. 2007) for data analysis. All paper forms were kept in locked filing cabinets and archived at the Institute of Psychiatry for 10 years. Electronic data were stored in password protected files.

7.5 Analysis

Analysis plans for testing the six hypotheses presented in Section 6.3 and the exploratory model of patient satisfaction with services are described below.

7.5.1 Hypothesis 1: The proportion of patients observed in contact with staff is greater at alternatives than standard services.

Hypothesis 1 was tested using linear regression. with CaSPAR proportion of patients in contact with staff as dependent variable and service type (alternative, standard) and recording point data (day and time of recording) as independent variables. Two models were estimated:

a) the relationship between CaSPAR score and service type (alternative or traditional), adjusting for clustering by individual service.

b) the relationship between CaSPAR score and service type (alternative or traditional), adjusting for clustering by individual service and adjusting for recording point variables (day and time of recording). Significantly ($p<.05$) higher CaSPAR scores at alternatives compared to standard services estimated by this model would provide corroboration of Hypothesis 1.

7.5.2 Hypothesis 2: Staff-reported intensity of care is greater at alternatives than at standard services.

The mean difference in minutes' contact per patient per day at alternatives and standard services was calculated from CaRICE data, together with the standard deviation. The mean difference as a proportion of the standard deviation was then calculated. A figure greater than 0.5 indicated a medium or greater effect size (Altman 1991), which was required to corroborate Hypothesis 2. Confidence intervals for the effect size were reported. A t-test comparing minutes' contact per patient per day at alternative and standard services was also conducted and 95% confidence intervals were reported. The reasons for using an estimate of effect size in hypotheses involving CaRICE data are discussed in Section 10.1.3. This analysis strategy does not provide a statistical test of the significance of differences between alternatives and standard services. Although the lack of power in CaRICE data meant that significant differences were unlikely to be found by a t-test, confidence intervals for the effect size and the mean difference in CaRICE scores were presented to inform discussion of CaRICE results by indicating the potential magnitude of differences between service types suggested by CaRICE data.

7.5.3 Hypothesis 3: Patient-reported intensity of care is greater at alternatives than standard services.

Hypothesis 3 was tested using linear regression, with CCCQ-P total score as dependent variable and service type (alternative, standard) and patients' characteristics as independent variables. Two models were estimated:
a) The relationship between CCCQ-P score and service type, adjusting for clustering by individual service.
b) The relationship between CCCQ-P score and service type, adjusting for clustering by individual service and adjusting for the following potentially confounding variables concerning patient characteristics:

- MHA status (detained during admission, not detained, MHA status unknown)
- Age
- Gender
- Ethnicity (White, Black, Asian, Other)

The four ethnicity groups used in analysis were created from the 16 categories presented to participants during data collection. Given the sample size, data reduction of ethnicity groups was required to facilitate analysis using linear regression. The four categories (White, Black, Asian, Other) derive from organising groups used in the 2001 UK census (Office of National Statistics 2001), from which the original 16 ethnicity categories came.

Significantly higher ($p<0.05$) CCCQ-P total scores at alternatives compared to standard services estimated by this model were required to corroborate of Hypothesis 3.

7.5.4 Hypotheses 4, 5 and 6:

Hypothesis 4: The staff-reported and patient reported intensity of social interventions is greater in alternatives than standard services.

Hypothesis 5: The staff-reported and patient reported intensity of psychological interventions is greater in alternatives than standard services.

Hypothesis 6: The staff-reported and patient-reported intensity of physical and pharmacological interventions is greater in standard services than in alternatives.

CaRICE data for minutes per patient per day of social, psychological and physical/pharmacological interventions at alternative and standard services were analysed similarly to CaRICE total contact data, as described for Hypothesis 2, estimating effect size by calculating the mean difference in CaRICE scores between alternatives and standard services divided by the standard deviation of all data. Confidence intervals from a t-test comparing alternatives and standard services were also presented for each CaRICE subscale.

CCCQ-P social, psychological and physical/pharmacological subscale data were analysed similarly to CCCQ-P total score data, as described in Hypothesis 3, using linear regression, adjusting for service type, clustering by service and patient characteristics.

7.4.5 Exploratory analysis of the relationship of content of care to patient satisfaction

Linear regression was used to model patient satisfaction, with CSQ score as dependent variable and service type, patient characteristics and CCCQ-P variables as independent variables. Six models of the relationship between CSQ score and service type were estimated, adjusting for:

a) clustering by service
b) patient characteristics (mental health act status, age, gender, ethnicity) and clustering by service
c) patient characteristics and CCCQ-P social interventions subscale score and clustering by service
d) patient characteristics and CCCQ-P psychological interventions subscale score and clustering by service

e) patient characteristics and CCCQ-P physical and pharmacological interventions subscale score and clustering by service
f) patient characteristics and CCCQ-P total score and clustering by service

P values were used to assess whether relationships between CSQ score and patient characteristics and CCCQ-P variables were significant ($p<0.05$). Regression coefficients and p values for service type were compared in each model to estimate the effect of patient characteristics and CCCQ-P variables on the relationship of service type to CSQ score. R^2 values were used to identify the total amount of variance in CSQ scores explained by variables included in each model.

Skewed data: Regression analyses, used to analyse CaSPAR, CCCQ-P and CSQ data, are parametric tests which assume normal distribution of data from dependent variables. Two strategies were used to address the possibility of non-normal distribution of CaSPAR, CCCQ-P or CSQ data.

i) Robust standard errors were calculated in all analyses, providing conservative estimates of confidence intervals.

ii) The distribution of data from CaSPAR, CCCQ-P and CSQ was assessed by visual inspection of histograms. If data were non-normally distributed, estimates of residual error were calculated following regression analyses and the distribution of residuals assessed. If residuals were also non-normally distributed, this was reported (in Chapter 8) and possible effects on results discussed.

All analyses involving CaSPAR and CCCQ-P data were conducted using Stata IC version 10.0 (StataCorp 2007). Analyses involving CaRICE data were conducted using SPSS version 14.0 software (SPSS 2006).

Chapter 8

Quantitative study results: comparing alternatives and standard services

8.1 Response rates and sample characteristics

Data from three measures of content of care - CaSPAR, CaRICE, CCCQ-P – and one measure of satisfaction - CSQ - were collected from four alternative and four standard services. Response rates and characteristics of the sample for each measure are provided below:

CaSPAR: 28 momentary time recordings were made at each participating service (224 recording points). The 224 CaSPAR service data comprised information from 4581 patient identifications (alternatives = 889; standard services = 3692). (Most patients will have contributed to more than one recording at a service.) CaSPAR data were generated from a mean of 20 service users at each recording (alternatives mean = 8; standard services mean = 33). The status (with staff or not) of three patients could not be ascertained during CaSPAR recordings (1 at alternatives; 2 at standard services). CaSPAR therefore provides data for 99.9% of the possible patient identifications at recording points.

CaRICE: Data were gathered from staff at each participating service over a five day recording period. A maximum of 919 forms (all staff working each day during the recording period) could have been completed during recording periods (alternatives = 263; standard services = 656). 871 completed CaRICE forms were obtained from staff (alternatives = 256; standard services = 615), a response rate of 94.7% (alternatives = 97.3%; standard services = 93.8%). The mean duration of shift during which CaRICE forms were completed was six and a half hours (391 minutes). The professional groups of respondents are described in Table 8.1.

Table 8.1: Professions of CaRICE respondents

Professional group	Alternatives	Standard services	All services
Qualified mental health professionals			
Psychiatrists	26 (10.2%)	110 (17.9%)	136 (15.6%)
Nurses	68 (26.6%)	165 (26.8%)	233 (26.8%)
Occupational Therapists	5 (2.0%)	36 (5.9%)	41 (4.7%)
Social Workers	12 (4.7%)	None	12 (1.4%)
Psychologists	1 (0.4%)	10 (1.6%)	11 (1.3%)
Unqualified staff and non-mental health professionals			
Nursing Assistants	44 (17.2%)	167 (27.2%)	211 (24.2%)
Social Care staff	85 (33.2%)	None	85 (9.8%)
Other	15 (5.9%)	127 (20.7%)	142 (16.3%)

52% of staff at standard services were qualified staff from core mental health professions, i.e. psychiatrists, nurses, occupational therapists, social workers, psychologists, compared to 44% at alternatives. The largest difference between alternatives and standard services was in the background of unqualified staff: in standard services these were mainly nursing assistants; in alternatives unqualified staff most frequently had a social care background. The majority of staff in the "other" category were students but a variety of other workers were represented including welfare rights advisors, advocates, pharmacists, faith group representatives and physiotherapists.

CCCQ-P: 447 patients were asked to participate in completing a CCCQ-P questionnaire (alternatives = 186; standard services = 261). Completed forms were obtained from 314 respondents (alternatives = 142; standard services = 172), a response rate of 70.2% (alternatives = 77%, standard services = 66%). A minimum of 40 questionnaires was completed from each service except one, Alternative 2 (clinical crisis house), where delays starting data collection and a slow throughput of patients meant only 20 questionnaires could be completed. The characteristics of patients who completed CCCQ-P questionnaires are summarised in Table 8.2.

Table 8.2: Characteristics CCCQ-P respondents (n = 314)

Patients' Characteristics		Alternatives (n=142)	Standards (n=172)	All services (n=314)
Gender	Male	63 (44%)	99 (58%)	162 (52%)
	Female	79 (56%)	73 (42%)	152 (48%)
Mean age at admission		39.9 years	38.9 years	39.4 years
MHA status at admission	Voluntary	113 (80%)	110 (64%)	223 (71%)
	Detained	27 (19%)	52 (30%)	79 (25%)
	Not known	2 (1%)	10 (6%)	12 (4%)
Ethnicity	White	79 (55%)	127 (74%)	206 (66%)
	Black	51 (36%)	25 (15%)	76 (24%)
	Asian	8 (6%)	6 (3%)	14 (4%)
	Other	4 (3%)	14 (8%)	18 (6%)

Three of the four alternative services contributed detained patients to the sample, the exception being Alternative 1 (non-clinical crisis house). The higher proportion of Black service users at alternative services can partly be explained by the fact that Alternative 1 (non-clinical crisis house) only accepted patients from black minority ethnic groups while, in a predominantly white mono-ethnic area, Alternative 2 (clinical crisis house) contributed a smaller sample than its paired standard service.

The staff-patient ratios at services during CaRICE recording weeks are reported in Appendix 7. No significant difference in staffing level between alternatives and standard services was found. (Staffing level was found to be positively associated with minutes' of staff contact provided per patient per day and negatively associated with proportion of staff time spent in direct patient contact. These results are discussed in Section 10.3.3.)

8.2 Main results

Results from CaSPAR and CCCQ-P data are presented together. Descriptive data are provided in Table 8.3 and results from regression analyses are presented in Table 8.4.

Table 8.3: CaSPAR and CCCQ-P – descriptive data

Measure	**Alternatives**	**Standard services**	**All services**
CaSPAR total score (mean proportion of patients recorded in contact with staff) 224 observation points (alternatives 112; standard 112)	11.9% (s.d. = 18.8)	11.0% (s.d. = 9.1)	11.5% (s.d. = 14.7)
CCCQ-P total score (frequency and range of interventions - patient reported) n = 314 (alternatives 142; standard 172)	27.5 (s.d. = 16.1)	31.6 (s.d. = 16.5)	29.8 (s.d. = 16.4)
CCCQ-P social interventions subscale score n = 314 (alternatives 142; standard 172)	7.8 (s.d. = 6.6)	7.8 (s.d. = 6.6)	7.8 (s.d. = 6.6)
CCCQ-P psychological interventions subscale score n = 314 (alternatives 142; standard 172)	5.4 (s.d. = 6.1)	5.2 (s.d. = 6.1)	5.3 (s.d. = 6.1)
CCCQ-P physical and pharmacological interventions subscale score n = 314 (alternatives 142; standard 172)	9.1 (s.d. = 5.7)	13.1 (s.d. = 5.8)	11.3 (s.d. = 6.1)

The significance of differences in CaSPAR and CCCQ-P scores presented in Table 8.3 is reported in Table 8.4.

Table 8.4: CaSPAR and CCCQ-P regression analyses

Dependent variable	Independent variables	R^2	Regression coefficient: service type*	95% confidence intervals	t	p
CaSPAR total score	1. Service type, adjusting for clustering by service	.001	-0.87	-9.30, 7.57	-0.24	.82
	2. Service type, adjusting for timepoint variables** and clustering by service	.028	-0.87	-9.46, 7.73	-0.24	.82
CCCQ-P total score	1. Service type, adjusting for clustering by service	.016	4.11	-2.54, 10.76	1.46	.19
	2. Service type, adjusting for patients' characteristics*** and clustering by service	.028	3.25	-1.07, 7.58	1.78	.12
CCCQ-P social interventions	1. Service type, adjusting for clustering by service	.000	0.04	-1.22, 1.30	0.07	.94
	2. Service type, adjusting for patients' characteristics*** and clustering by service	.014	-0.15	-1.22, 0.85	-0.03	.98
CCCQ-P psychological interventions	1. Service type, adjusting for clustering by service	.000	-0.18	-1.89, 1.52	-0.26	.81
	2. Service type, adjusting for patients' characteristics*** and clustering by service	.013	-0.28	-1.42, 0.85	-0.59	.58
CCCQ-P physical and pharmacological interventions	1. Service type, adjusting for clustering by service	.107	4.00	-1.14, 9.13	1.84	.11
	2. Service type, adjusting for patients' characteristics*** and clustering by service	.183	3.24	-.0.01, 6.47	2.38	.049

* negative regression coefficient = higher score at alternatives

** timepoint variables = day of the week and shift (early, late, night)

*** patients' characteristics = Mental Health Act status at admission (detained, not detained, not known), age, gender and ethnicity (black, white, asian, other)

Table 8.4 shows that there was no significant difference between alternatives and standard services in the proportion of patients in contact with staff, measured by CaSPAR. There was no significant difference between alternatives and standard services, before and after adjustment for patient characteristics, in the frequency and range of total care, social interventions or psychological interventions, measured by CCCQ-P. CCCQ-P scores for standard services were significantly higher than for alternatives, adjusting for patient variables and clustering by service, indicating that standard services provide a greater frequency and range of physical and pharmacological interventions.

CaSPAR data exhibited positive skew and (unlike CCCQ-P and CSQ data), residuals calculated following regression analysis were also not normally distributed. The conservative estimate of confidence intervals provided by calculation of robust standard error in regression analysis (through adjusting by clustering by service) is therefore appropriate. Given the clearly non-significant nature of the difference between alternatives and standard services in CaSPAR scores, the skewed distribution of data will have minimal effect on findings.

Descriptive data and estimates of effect size from CaRICE data are reported in Table 8.5.

Table 8.5 CaRICE descriptive data and estimates of effect size

CaRICE domain **(minutes of staff contact per patient per day)**	**All services** **(40 days' data)**	**Alternatives** **(20 days' data)**	**Standard services** **(20 days' data)**	**Mean difference**	**95% confidence intervals (t-test)**	**Effect size*** **(95% C.I.)**
Total score	131.0 (s.d.=36.3)	139.6 (s.d.=39.9)	122.5 (s.d.=31.0)	17.1	-5.9, 39.9	0.47 (-0.16, 1.10)
social interventions	48.7 (s.d.=26.1)	56.4 (s.d.=31.7)	41.0 (s.d.=16.3)	15.4	-0.7, 31.5	0.59 (-0.03, 1.23)
psychological interventions	22.9 (s.d.=16.4)	26.9 (s.d.=22.2)	19.0 (s.d.=4.8)	7.9	-2.4, 18.3	0.48 (-0.15, 1.11)
physical and pharmacological interventions	36.2 (s.d.=23.5)	31.1 (s.d.=26.8)	41.3 (s.d.=19.1)	-10.2	-25.1, 4.7	0.43 (-1.06, 0.20)

* Effect size = mean difference / standard deviation

Table 8.5 shows that, measured by CaRICE, alternatives provided greater intensity of total care and psychological interventions than standard services and standard services provided greater intensity of physical and pharmacological interventions, but that the effect size was small in each case (mean difference/standard deviation = 02 – 0.5). Alternatives provided greater intensity of social interventions than standard services, measured by CaRICE and for social interventions there was a medium effect size for service type (mean difference/standard deviation = 05 – 0.8).

8.2.1 Study hypotheses

Tables 8.4 and 8.5 show that no study hypotheses were corroborated. Hypotheses 1-3 were that intensity of total care was greater at alternatives than standard services, measured by CaSPAR, CaRICE and CCCQ respectively. The intensity of staff-patient contact was not significantly different at alternative and standard services, measured by CaSPAR or CCCQ-P, and the effect size for alternatives was not medium or large, measured by CaRICE. Hypotheses 4-6 were that, confirmed by CCCQ-P and CaRICE, intensity of social and psychological care were greater at alternatives than standard services and intensity of physical and pharmacological interventions was greater at standard services than at alternatives. Medium effects and significant difference between alternatives and standard services were not identified by both CaRICE and CCCQ-P for provision of social, psychological or physical and pharmacological interventions.

The wide confidence intervals for mean differences between alternative and standard services, obtained from linear regressions of CaSPAR and CCCQ-P data and t-test of CaRICE data, indicate the possibility of clinically important differences between alternative and standard services despite the non-significant findings from investigation of study hypotheses. However, the low R^2 values in CaSPAR and CCCQ-P regression analyses suggest the lack of significant differences found between alternatives and standard services for intensity of care was not due to the study being insufficiently powered, but that service type had weak ability to explain variance in data measuring levels of staff-patient contact.

For social interventions, results from CaRICE and CCCQ-P were divergent: CaRICE indicated a medium effect for greater provision of social interventions at alternatives while CCCQ-P mean score was marginally higher at standard services and adjusted CCCQ-P mean score marginally higher at alternatives, with no significant difference shown. Results from the two measures were more complementary for psychological and physical/pharmacological interventions. Both CaRICE and CCCQ-P showed a trend towards greater provision of psychological interventions at alternatives but it did not achieve significance. CCCQ-P data showed significantly greater provision of physical/pharmacological interventions at standard services; CaRICE data also showed greater provision of physical/pharmacological interventions at standard services, but the estimated effect size fell short of being medium.

As with Hypotheses 1-3, the width of confidence intervals in analyses in Hypotheses 4-6 mean that there may be clinically important differences between alternatives and standard services in the broad types of care being provided, despite the thesis' hypotheses not being confirmed. R^2 values from analyses of CCCQ-P data indicate that service type is not able to explain most of the variance in data about types of care provided but makes some contribution to predicting intensity of provision of physical and pharmacological interventions.

8.2.2 Alternatives and standard services: individual categories of care

Table 8.6 provides descriptive data from CaRICE for all 21 individual categories of care and estimates the effect size of service type for each category of care by calculating the mean difference between scores for alternative and standard services as a proportion of the standard deviation for all data.

Table 8.6: CaRICE data – individual care categories: alternatives vs. standard services

	Alternatives (20 days' data)	Standard (20 days' data)	All services (40 days' data)	Estimate of effect size (mean difference/s.d >0.5)
	Mean minutes contact per patient per day (s.d.)			
More at alternatives: medium or large effect size compared to standard services				
10. Relationships	4.4 (5.6)	1.8 (1.5)	3.1 (3.2)	0.80
1. Housing	5.0 (9.1)	1.5 (2.0)	3.2 (6.7)	0.51
4. Current activity	33.6 (34.8)	19.9 (12.9)	26.7 (26.9)	0.51
12. Drugs/alcohol	5.6 (8.8)	2.2 (2.2)	3.9 (6.6)	0.51
No difference: small or no effect size for alternatives compared to standard services				
11. Past events	4.4 (5.4)	2.4 (1.4)	3.4 (4.0)	0.49
8. Assessment	17.9 (13.0)	13.1 (5.2)	15.5 (10.1)	0.48
7. Family support	4.0 (4.8)	2.6 (2.1)	3.3 (3.7)	0.39
2. Finances	2.7 (3.4)	1.7 (2.5)	2.2 (3.0)	0.33
14. Medication compliance	2.9 (3.3)	2.7 (1.9)	2.8 (2.7)	0.07
5. Future activity	4.1 (5.1)	4.0 (4.0)	4.0 (4.5)	0.02
9. Coping strategies	9.8 (7.7)	9.7 (3.4)	9.7 (5.9)	0.01
13. Illness education	2.8 (3.1)	2.9 (1.3)	2.8 (2.3)	0.03
20. Care Planning	6.1 (9.0)	6.8 (7.3)	6.4 (8.1)	0.08
16. medication practical	7.0 (5.6)	7.5 (5.4)	7.3 (5.4)	0.09
15. medication review	2.8 (3.5)	3.2 (2.5)	3.0 (3.0)	0.14
19. Restraint	1.1 (1.7)	1.3 (2.4)	1.2 (2.1)	0.14
21. Care coordination	1.2 (1.7)	1.5 (0.9)	1.4 (1.4)	0.19
18. Observations	15.5 (21.4)	21.4 (20.0)	18.5 (20.7)	0.28
6. A.D.L.	6.3 (3.9)	8.6 (5.6)	7.4 (4.9)	0.47
More at standard services: medium or large effect size compared to alternatives				
3. Legal	0.8 (2.0)	2.7 (2.8)	1.8 (2.6)	0.75
17. Physical health	1.8 (2.3)	5.1 (4.5)	3.5 (3.9)	0.85

Table 8.7 shows descriptive data from CCCQ-P and reports a t-test comparing mean scores from alternative and standard services for each category of care.

Table 8.7: CCCQ-P data - individual care categories: alternatives vs. standard services

	All services (n=314)	Alternatives (n=142)	Standards (n=172)	t	p
	CCCQ-P score (standard deviation)				
Greater provision at alternatives					
1. Housing	0.7 (1.4)	1.0 (1.7)	0.6 (1.2)	2.4	.02
10. Relationships	0.6 (1.5)	0.8 (1.7)	0.4 (1.3)	2.3	.02
No significant difference					
11. Past events	1.1 (1.95)	1.2 (2.01)	1.0 (1.87)	1.3	.19
3. Legal	0.4 (1.1)	0.5 (1.1)	0.4 (0.9)	1.0	.31
21. Care coordination	1.2 (1.6)	1.2 (1.7)	1.2 (1.6)	0.3	.71
7. Family support	1.9 (2.2)	1.9 (2.2)	1.9 (2.2)	0.1	.92
6. A.D.L.	0.9 (1.8)	0.9 (1.9)	0.9 (1.7)	-0.1	.94
12. Drugs/alcohol	0.8 (1.8)	0.8 (1.8)	0.8 (1.8)	-0.3	.77
2. Finances	0.5 (1.1)	0.4 (1.1)	0.4 (1.1)	-0.4	.72
8. Assessment	2.9 (2.3)	2.8 (2.3)	3.0 (2.3)	-0.5	.60
14. Medication compliance	1.3 (1.8)	1.2 (1.8)	1.3 (1.8)	-0.6	.57
9. Coping strategies	1.6 (2.2)	1.6 (2.2)	1.7 (2.3)	-0.6	.57
5. Future activity	1.0 (1.7)	0.9 (1.6)	1.0 (1.9)	-0.6	.55
4. Current activity	2.4 (2.5)	2.2 (2.6)	2.6 (2.5)	1.4	.16
13. Illness education	1.3 (1.9)	1.1 (1.8)	1.4 (1.9)	-1.6	.13
Greater provision at standard services					
16. medication practical	5.4 (2.7)	5.0 (2.9)	5.7 (2.4)	-2.3	.02
20. Care Planning	1.4 (1.5)	1.1 (1.4)	1.6 (1.5)	-2.6	.01
15. medication review	1.2 (1.3)	1.0 (1.3)	1.4 (1.4)	-3.0	.003
17. Physical health	1.8 (2.0)	1.1 (1.8)	2.3 (2.1)	-5.3	<.001
18. Observations	1.4 (2.3)	0.7 (1.8)	1.9 (2.5)	-5.0	<.001
19. Restraint	0.2 (0.7)	0.1 (0.3)	0.4 (0.9)	-3.9	<.001

Tables 8.6 and 8.7 show that CaRICE and CCCQ-P both found no difference between alternatives and standard services in a majority of care categories. Medium effect sizes and significant differences between alternatives and standard services were found by CaRICE and CCCQ-P for three types of care, while no significant difference was found by both measures for ten types. For eight categories, a medium effect or significant difference was found on one measure but not the other. Both data sources indicated significantly higher scores for help with housing and relationships at alternatives compared to standard services and significantly higher scores for physical health care at standard services. Because no account has been made for multiple testing however, it is possible positive results could represent Type 1 errors: inferences about the provision of individual categories of care at alternatives and standard services can only be preliminary.

The categories of care comprising the physical and pharmacological interventions subscale exhibited a consistent trend. Standard services scored higher than alternatives for all six categories in CCCQ-P data and five out of six (all except medication compliance) in CaRICE data.

8.3 Post hoc analysis: community alternatives versus standard services

The quantitative comparison of care provided at alternative and standard services was driven by six hypotheses, none of which was corroborated. Multi-methods investigation revealed no consistent significant differences between alternatives and standard services in the intensity of staff-patient contact or of provision of different types of care.

Community alternatives were combined and compared to standard services in a secondary analysis for the following three reasons:

1) Comparisons of community and hospital services are common in mental health services research. A distinction between community-based and

inpatient alternatives was used by Braun and colleagues (1981) to dichotomise types of alternative to standard inpatient care. There is thus some face validity to considering community-based alternatives as a group to compare to standard inpatient services.

2) Unpublished research not forming part of this thesis conducted for The Alternatives Study provides some support for this approach. Qualitative interviews with service users and other stakeholders of Alternative 4, the inpatient alternative using the Tidal Model included in the quantitative investigation in this thesis, suggested that implementation of the Tidal Model at the service was incomplete and that the service was in practice very similar to standard inpatient care. This raises the possibility that differences in care provided at alternatives and standard services might be more apparent if community-based alternatives were considered separately and not aggregated with data from an inpatient Tidal Model ward.
3) Descriptive data for individual services, presented in Tables 8.11 and 8.12 also indicated that content of care in community alternatives might differ more than the Tidal Model Ward from standard services. The results of individual service data from study measures are discussed in Section 10.2.

An exploratory, post hoc comparison of the care provided at community-based alternatives and standard services was conducted, using similar analyses to those initially undertaken for all four alternatives and four standard services. For these analyses, data from Alternative 4, the inpatient Tidal Model alternative, and Standard 4, its local comparison standard service, were excluded. Thus data from six services – the three community-based alternatives and their comparison local standard services – were included.

8.3.1: community alternatives versus standard services: results

Descriptive data from CaSPAR, and CCCQ-P comparing care at community alternatives and standard services are provided in Table 8.8

Table 8.8 Secondary analysis – community alternatives: CaSPAR and CCCQ-P descriptive data

Measure	**Community alternatives**	**Standard services**	**All services**
CaSPAR total score (mean proportion of patients recorded in contact with staff) 168 observation points (alternatives 84; standard 84)	12.3% (s.d. = 20.9)	11.8% (s.d. = 9.4)	12.0% (s.d. = 16.1)
CCCQ-P total score (frequency and range of interventions - patient reported) n = 227 (alternatives 102; standard 225)	25.7 (s.d. = 16.2)	30.6 (s.d. = 15.5)	28.4 (s.d. = 15.9)
CCCQ-P social interventions subscale score n = 227 (alternatives 102; standard 225)	7.4 (s.d. = 6.6)	7.9 (s.d. = 6.4)	7.7 (s.d. = 6.5)
CCCQ-P psychological interventions subscale score n = 227 (alternatives 102; standard 125)	5.7 (s.d. = 6.3)	4.7 (s.d. = 5.8)	5.2 (s.d. = 6.0)
CCCQ-P physical and pharmacological interventions subscale score n = 227 (alternatives 102; standard 225)	7.7 (s.d. = 5.8)	12.7 (s.d. = 5.8)	10.4 (s.d. = 6.3)

The significance of differences between community alternatives and standard services presented in Table 8.8 is reported in Table 8.9.

Table 8.9: Secondary analysis – community alternatives: CaSPAR and CCCQ-P regression analyses

Dependent variable	**Independent variables**	**R^2**	**Regression coefficient: service type***	**95% confidence intervals**	**t**	**p**
CaSPAR total score	1. Service type, adjusting for clustering by service	0.00	-0.45	-12.68, 11.78	-0.09	.93
	2. Service type, adjusting for timepoint variables** and clustering by service	0.03	-0.45	-12.98, 12.03	-0.09	.93
CCCQ-P total score	1. Service type, adjusting for clustering by service	0.02	4.85	-2.70, 12.41	1.65	.16
	2. Service type, adjusting for patients' characteristics*** and clustering by service	0.06	3.31	-2.27, 8.90	1.53	.19
CCCQ-P social interventions	1. Service type, adjusting for clustering by service	0.00	0.49	-1.00, 1.97	0.84	.44
	2. Service type, adjusting for patients' characteristics*** and clustering by service	0.03	0.02	-2.09, 2.14	0.03	.98
CCCQ-P psychological interventions	1. Service type, adjusting for clustering by service	0.01	-0.97	-2.80, 0.85	-1.38	.23
	2. Service type, adjusting for patients' characteristics*** and clustering by service	0.04	-1.33	-2.48, -0.18	-2.98	.03
CCCQ-P physical and pharmacological interventions	1. Service type, adjusting for clustering by service	0.16	5.03	-1.01, 11.08	2.14	.09
	2. Service type, adjusting for patients' characteristics*** and clustering by service	0.25	4.35	0.75, 7.96	3.10	.03

* negative regression coefficient = higher score at alternatives

** timepoint variables = day of the week and shift (early, late, night)

*** patients' characteristics = Mental Health Act status at admission (detained, not detained, not known), age, gender and ethnicity (black, white, asian, other)

Table 8.9 shows there was no significant difference in intensity of total care between community alternatives and standard services, measured by CaSPAR or CCCQ-P. There was also no significant difference for social interventions, measured by CCCQ-P. Adjusting for patient variables, CCCQ-P data show community alternatives provided significantly more psychological interventions and significantly less physical and pharmacological interventions than standard services.

Descriptive data and estimates of effect size for community alternatives compared to standard services from CaRICE are reported in Table 8.10.

Table 8.10 Secondary analysis – community alternatives: CaRICE descriptive data and estimates of effect size

CaRICE domain **(minutes of staff contact per patient per day)**	**All services** **(30 days' data)**	**Community alternatives** **(15 days' data)**	**Standard services** **(15 days' data)**	**Mean difference**	**95% confidence intervals (t-test)**	**Effect size* (95% C.I.)**
Total score	142.7 (s.d.=32.3)	149.6 (s.d. = 39.1)	135.9 (s.d. = 23.1)	13.7	-10.6, 38.0	0.42 (-0.31, 1.14)
social interventions	52.2 (s.d.=28.9)	60.7 (s.d.=35.6)	43.8 (s.d.=17.7)	17.0	-4.4, 38.3	0.59 (-0.15, 1.32)
psychological interventions	24.7 (s.d.=18.2)	29.7 (s.d.=24.6)	19.8 (s.d.=5.7)	9.9	-3.9, 23.7	0.54 (-0.19, 1.27)
physical and pharmacological interventions	40.1 (s.d.=25.3)	31.9 (s.d.=30.0)	48.3 (s.d.=16.6)	-16.4	-34.8, 2.0	0.65 (-1.39, 0.08)

* Effect size = mean difference / standard deviation

Table 8.10 shows that the higher CaRICE total score for community alternatives compared to standard services represents only a small effect size. Measured by CaRICE, community alternatives provided greater intensity of social and psychological interventions and the effect size was medium. Standard services provided greater inensity of physical and pharmacological interventions than community alternatives with a medium effect size.

The results from comparison of community-based alternatives and standard services revealed similar trends to comparisons of all alternatives and standard services, but differences in the nature of care provided were more pronounced. As with comparisons of all alternatives, all measures found no significant difference in overall intensity of staff-patient contact in community alternatives and standard services. CaRICE and CCCQ-P social interventions data diverged as in the main analysis: CaRICE indicated a medium effect size for community alternatives but CCCQ-P data showed non-significantly greater scores at standard services. CaRICE and CCCQ-P data both indicated medium effect/significantly greater provision of psychological interventions at community alternatives and of physical and pharmacological interventions at standard services.

For both psychological and physical and pharmacological interventions, CaRICE estimated effect sizes were larger and p values from CCCQ-P regressions were smaller comparing community-based alternatives to standard services than for comparisons of all alternatives to standard services. The exploratory analyses in Section 8.4 provide a preliminary indication that community-based alternatives may differ more from standard services than inpatient alternatives do, providing more psychological and less physical and pharmacological care.

R^2 values in regression analyses of CCCQ-P and CaSPAR data indicate service type and measured patient characteristics cannot explain most of the variance in total or subscale scores. Service type though has some explanatory power in accounting for variance in physical and pharmacological scores; for other subscale and total scores, service type is a weak explanatory factor.

8.4 Descriptive data for individual services

It is possible that there may be substantial variation in care provided at different alternative services in this study which has not been revealed by the comparisons of groups of alternatives with standard services presented so far in this chapter. Further descriptive data will be provided in this section to allow exploration of heterogeneity among alternative services.

Descriptive data for individual services about intensity of staff-patient contact (CaSPAR, CaRICE and CCCQ-P total scores) are provided in Table 8.11.

Table 8.11: CaSPAR CaRICE and CCCQ-P total scores: individual services

CaSPAR (proportion of patients in contact with staff)		**CaRICE total score** mean minutes of contact per patient per day		**CCCQ-P** mean total score
Alternative 2 21.6%		**Alternative 3** 160.4		**Standard 4** 34.4
Standard 3 14.1%		**Standard 2** 154.1		**Alternative 4** 31.9
Standard 2 13.5%		**Alternative 2** 139.5		**Standard 2** 31.8
Alternative 4 10.8%		**Alternative 1** 133.3		**Standard 1** 30.6
Standard 4 8.8%		**Standard 3** 131.7		**Alternative 2** 29.8
Alternative 1 8.6%		**Standard 1** 121.7		**Standard 3** 29.4
Standard 1 8.0%		**Alternative 4** 109.6		**Alternative 3** 29.1
Alternative 3 4.6%		**Standard 4** 82.6		**Alternative 1** 20.6

Alternative 1 = non clinical crisis house
Alternative 2 = clinical crisis house
Alternative 3 = crisis team beds
Alternative 4 = Tidal Model ward

Descriptive data about the nature of care provided at individual services (CaRICE and CCCQ-P subscale scores) are presented in Table 8.12.

Table 8.12: CaRICE and CCCQ-P subscale scores: individual services
(CaRICE = minutes of contact per patient per day; CCCQ-P = mean subscale score)

Social interventions subscale score		**Psychological interventions subscale score**		**Physical/pharmacological interventions subscale score**	
CaRICE	**CCCQ-P**	**CaRICE**	**CCCQ-P**	**CaRICE**	**CCCQ-P**
Alternative 2 98.0	**Standard 2** 8.8	**Alternative 1** 41.7	**Alternative 3** 6.8	**Alternative 3** 64.2	**Standard 4** 14.1
Alternative 1 59.2	**Alternative 4** 8.7	**Alternative 3** 36.6	**Standard 4** 6.6	**Standard 2** 60.4	**Standard 2** 13.1
Standard 2 57.0	**Standard 1** 8.0	**Standard 3** 22.9	**Alternative 2** 5.6	**Standard 3** 54.3	**Standard 1** 13.0
Alternative 4 43.3	**Alternative 2** 7.9	**Standard 1** 20.9	**Standard 2** 5.2	**Standard 1** 30.3	**Alternative 4** 12.5
Standard 1 43.2	**Alternative 1** 7.7	**Alternative 4** 18.4	**Alternative 4** 4.8	**Alternative 4** 28.8	**Alternative 2** 12.4
Standard 4 32.6	**Standard 4** 7.6	**Standard 4** 16.4	**Standard 3** 4.7	**Standard 4** 20.4	**Standard 3** 12.0
Standard 3 31.1	**Standard 3** 6.9	**Standard 2** 15.6	**Alternative 1** 4.7	**Alternative 2** 16.9	**Alternative 3** 9.8
Alternative 3 24.9	**Alternative 3** 6.8	**Alternative 2** 10.8	**Standard 1** 4.3	**Alternative 1** 14.6	**Alternative 1** 3.3

Alternative 1 = non clinical crisis house, Alternative 2 = clinical crisis house,

Alternative 3 = crisis team beds, Alternative 4 = Tidal Model ward

Tables 8.11 and 8.12 provide scores and rankings of individual services for content of care measures' total scores and subscale scores respectively. Mean scores from for individual services for all twenty one categories of care measured by CaRICE and CCCQ-P are provided in Appendix 8. Characteristics of the nature of care at individual services suggested by Tables 8.11 and 8.12 are discussed in Section 10.2.

Chapter 9

Quantitative study results: a model of patient satisfaction with inpatient services

Results from an exploratory model of patient satisfaction with residential and inpatient mental health services are presented in Chapter 9. The contributions of service type (alternative or standard), patient characteristics and care received (CCCQ-P variables) to patient satisfaction with services are assessed and the extent to which differences in care provided contribute to variation between alternatives and standard services in patient satisfaction are explored.

9.1 Client Satisfaction Questionnaire: descriptive data

Client Satisfaction Questionnaires (CSQ) were completed with 314 patients at four alternative and four standard services (alternatives = 142; standard services = 172). Data were collected from the same sample and at the same time as CCCQ-P questionnaires. CSQ response rates and characteristics of participants are the same as those reported for CCCQ-P respondents in Section 8.1. Mean CSQ scores at alternative and standard services are shown in Table 9.1.

Table 9.1

CSQ descriptive data: alternatives and standard services

Service Type	n	Mean CSQ score (s.d.)	Range	Interquartile scores		
				25%	50%	75%
Alternatives	142	25.7 (5.7)	8-32	23	27	30
Standard services	172	23.5 (6.2)	8-32	20	25	28
All services	314	24.5 (6.1)	8-32	21	26	29

The widest possible range of total CSQ scores (from 8-32) was provided at both alternatives and standard services. Mean scores in Table 9.1 indicate that patients at alternatives and standard services were overall fairly satisfied with the service received.

Table 9.2 reports mean CSQ scores for patients with different socio-demographic characteristics within the study sample.

Table 9.2 : CSQ descriptive data : patient groups

Patient Group	**n**	**Mean CSQ score (s.d.)**
Men	162	25.1 (5.1)
Women	152	23.9 (6.9)
White	206	24.6 (6.3)
Black	76	24.7 (5.2)
Asian	14	24.0 (6.4)
Other	18	22.3 (6.1)
Voluntary	223	25.5 (5.1)
Detained	79	21.6 (7.5)
MHA status unknown	12	24.7 (5.3)

Table 9.3 describes the mean CSQ score at individual services.

Table 9.3 : CSQ descriptive data: individual services

Service	n	Mean CSQ score (s.d.)
Alternative 3	41	26.6 (5.2)
Alternative 1	41	26.6 (3.7)
Alternative 2	20	25.7 (6.5)
Standard 3	43	25.1 (5.9)
Standard 4	47	24.6 (4.9)
Alternative 4	40	23.8 (6.9)
Standard 1	42	22.0 (6.8)
Standard 2	40	22.0 (6.7)

Table 9.3 indicates that satisfaction with services was highest at the three community-based alternatives. Alternative 4, the Tidal Model inpatient alternative scored in the middle of the range of standard service scores. Mean scores indicate that typical responses at all services approximate to fairly satisfied with services.

9.2 The relationship of service type and patient characteristics to patient satisfaction

CSQ scores at alternative and standard services were compared using linear regression. Adjustment was made for clustering by service and patient characteristics. Results are presented in Table 9.4.

Table 9.4:

CSQ linear regression: patient satisfaction, service type and patient characteristics

Dependent variable = CSQ score

Model: relationship to patient satisfaction	R^2	Regression coefficient	t	p	95% confidence intervals
1. Service type, adjusting for clustering by service	.072				
Service type: standard (reference category = alternative)		-2.19	-2.12	.07	-4.63, 0.25
2. Service type, adjusting for patient characteristics (MHA status, age, gender, ethnicity) and clustering by service	.135				
Service type: standard		-1.96	-3.31	.01	-3.35. -0.56
MHA status: detained (reference category = voluntary)		-3.80	-9.07	<.01	-4.79, -2.81
MHA status: other (reference category = voluntary)		-0.24	-0.11	.91	-5.49, 5.00
Age		0.04	1.38	.21	-0.02, 0.10
Gender: women (reference category = men)		-1.85	-5.06	<.01	-2.71, -0.98
Ethnicity: Black (reference category = white)		-0.53	-1.01	.34	-1.76, 0.70
Ethnicity: Asian (reference category = white)		-0.41	-0.22	.83	-4.84, 4.02
Ethnicity: Other (reference category = white)		-2.29	-1.33	.23	-6.35, 1.78

Table 9.4 shows that patients' gender and Mental Health Act status both had a significant effect on their satisfaction with services. Ethnicity and age did not significantly affect patient satisfaction with services. Adjusting only for clustering by service, alternatives narrowly failed to show significantly greater patient satisfaction than standard services ($p = 0.07$). However, adjusting for patient socio-demographic characteristics and Mental Health Act status as well as clustering by service, CSQ scores were significantly higher at alternative than standard services ($p = 0.01$). Table 9.4 provides preliminary evidence that alternative services are more acceptable than standard services to patients.

9.3: The relationship of care received to patient satisfaction

CSQ scores for patients at alternative and standard services were compared using linear regression, adjusting for clustering by service and patients' characteristics as in Table 9.4, but additionally for patients' CCCQ-P subscale and total scores. Results of additional adjustment for CCCQ-P social interventions score, psychological interventions score, physical and pharmacological interventions score and total score are presented in Table 9.5.

Table 9.5: CSQ linear regression 2: patient satisfaction, service type, patient characteristics and CCCQ-P variables

Dependent variable = CSQ score

Model: relationship to patient satisfaction	R^2	Regression coefficient	t	p	95% C.I.
1. Service type, adjusting for patient characteristics* CCCQ-P social interventions subscale score and clustering by service	0.22				
Service type: standard (reference category = alternative)		-1.90	-2.50	.04	-3.69, -0.10
CCC(P) social interventions score		0.26	10.13	<.01	0.20, 0.31
2. Service type, adjusting for patient characteristics, CCCQ-P psychological interventions subscale score and clustering by service	0.24				
Service type: standard		-2.06	-3.61	.01	-3.41, -0.10
CCCQ-P psychological interventions score		0.31	7.68	<.01	0.21, 0.40
3. Service type, adjusting for patient characteristics, CCCQ-P physical and pharmacological interventions subscale score and clustering by service	0.18				
Service type: standard		-2.85	-3.47	.01	-4.79, -0.91
CCCQ-P physical and pharmacological interventions score		0.22	4.37	<.01	0.10, 0.34
4. Service type, adjusting for patient characteristics, CCCQ-P total score and clustering by service	0.27				
Service type: standard		-2.53	-3.10	.02	-4.46, -0.60
CCCQ-P total score		0.14	7.94	<.01	0.10, 0.18

*(MHA status, age, gender, ethnicity)

Table 9.5 shows that all types of care and total care received, as measured by CCCQ-P, are significantly associated with patient satisfaction. This relationship is positive for all types of care: patients who report receiving more care are more satisfied with services. This investigation therefore suggests that receiving more of any of the three care groupings has a positive impact on satisfaction.

R^2 values for linear regressions in Table 9.5 suggest that CCCQ-P variables have a modest but not negligible role in explaining variance in patient satisfaction. Adjusting for CCCQ-P total score in addition to adjusting for service type and patient characteristics doubles the amount of variance in CSQ scores explained from 13.5% to 27%. This indicates however, that nearly three quarters of variance in patient satisfaction is not explained by variables included in this model. Standardised regression coefficients for CCCQ-P variables indicate that variation in psychological interventions subscale scores is most closely related to variation in patient CSQ scores, with social interventions scores also more closely associated than physical and pharmacological interventions score.

P values for service type in linear regressions in Table 9.5 show that the relationship between service type and patient satisfaction remained significant after adjustment for each CCCQ-P variable. Additional adjustment for CCCQ-P variables, in addition to patient characteristics, produced modest changes in regression coefficients for service type, suggesting that the influence of care received, measured by CCCQ-P, on the relationship between service type and patient satisfaction is weak. The largest change in regression coefficient for service type is following adjustment for CCCQ-P physical and pharmacological interventions, reflecting that this was the CCCQ-P domain with greatest difference between mean scores at alternatives and standard services. The increase in regression coefficient (from 1.96, adjusting for patient characteristics, to 2.85, adjusting for patient characteristics and CCCQ-P physical and pharmacological subscale score) indicates that satisfaction with alternatives would increase compared to standard services if the level of physical and pharmacological interventions were similar: i.e. that alternatives are more acceptable than standard services despite providing less physical and pharmacological care rather than because of it.

Chapter 10

Discussion

10.1 Limitations

There were limitations to the study’s scope, measures and analysis.

10.1.1 Scope of the study

This thesis did not evaluate the effectiveness of alternatives and standard inpatient services, except for patient satisfaction with services. Thorough service evaluation requires assessment of outcome domains including mental and physical health, needs, social functioning, quality of life, costs and service use (Higginson 1994, Gilbody et al. 2002). Higginson (1994) cautions against measuring elements of structure or process at services to infer service quality, when their relationship to desired health outcomes is uncertain. The content of care investigation provided in this thesis is therefore an insufficient basis on which to make recommendations about provision of alternatives or draw conclusions about their effectiveness.

As the MRC framework for evaluating complex interventions (Craig et al. 2008) proposes, process evaluation is useful to describe services, identify variation in service provision and define service models, and understand service outcomes. The content of care in services represents only one means to describe and distinguish services however (Johnson and Salvador Carulla 1998). Two other aspects of care relevant to understanding variation in services and their outcomes were not measured in this thesis:

I. Continuity of care: achieving consistent, uninterrupted care provision to patients within a service and between services is widely recognised as an important service aim and may influence patient experience of care and outcomes (Crawford et al. 2004). Measurable elements of inpatient care relating to continuity include: a) consistency of contact provided to patients with key staff during admission (e.g. a named nurse); b) levels of staff absence or use of bank or agency staff at a service; c) patients’ pathways in and out of care, e.g. the duration of time between referral to an inpatient service and admission or between discharge and any community follow-up.

II. Service style: the social climate of a residential or ward unit can be measured. The Ward Atmosphere Scale (Moos 1996) for example, uses ten subscales to assess elements of ward environment including the degree of staff control, support and orientation towards practical or personal problems, and has demonstrated a relationship to patient satisfaction and outcomes in hospital inpatient settings (Jorgensen et al. 2009). The quality of interactions between staff and patients might also be assessed through measures such as therapeutic alliance scales, which have been used in routine mental health settings (Catty et al. 2007).

Within the process domain which was assessed in this thesis – content of care – the depth of information provided was limited (Section 5.5.2). Specific interventions or treatment modalities were not identified by the 21 categories provided to describe care in CaRICE and CCCQ-P. Differences between services in the intensity and nature of indirect care (care not provided during face-face contact with clients) provided for patients were not assessed by CaSPAR or CaRICE, which focused on the care received by patients during direct contact with staff.

10.1.2 Measures

The psychometric testing of study measures has been reported (Section 5.4). The remaining uncertainty about the quality of measures' psychometric properties limits confidence in study results and also complicates their interpretation. Individual service results from CaSPAR, CaRICE and CCCQ-P exhibit divergence between measures (Table 8.27). This divergence may represent: a) psychometric limitations of the measures; b) differences in the variables being measured; or c) differences in the perspectives of respondent groups.

Psychometric limitations

The inter-rater reliability of CaSPAR was inferred from tests of previous measures which used similar methods, but was not demonstrated for CaSPAR itself. Factors affecting the inter-rater reliability of CaRICE which were not adequately reflected in the inter-rater reliability test conducted for this study may include the greater complexity of interventions in vivo than as described in vignettes, the distractions of a busy working environment for staff completing CaRICE in vivo and social desirability bias. The spread and range of scores from CCCQ-P data provide some reassurance that patients were not oblivious to care provided or using the measure crudely to praise or criticise services with high or low scores. (Only two out of 314 respondents reported receiving no care; the highest score was 97 out of 147; data were normally distributed within this range.) However, possible factors influencing patients' completion of CCCQ-P include severity of illness (leading to recall bias) or satisfaction with the service (less satisfied patients might be expected to experience or report a smaller proportion of staff interventions as care received than more satisfied patients did).

Differences in variables measured

An important difference in the focus of the measures is that CaSPAR and CaRICE measure the frequency and duration of staff interactions with patients but not the range of care provided; CCCQ-P measures the frequency and range of direct care but not its duration. This has the potential to create substantial differences in scores between CCCQ-P and the other two measures. For example, Alternative 2 provides a daily timetable of structured activity for patients and scored highest of all services for help with current activity on both CaRICE and CCCQ-P. However, while its score for the CaRICE current activity category contributed 62% of its total CaRICE score and was more than double the next highest service, its CCCQ-P current activity score contributed only 12% to its total CCCQ-P score and less than one point higher than four other services. Other features of the measures' scoring and sampling which may contribute to divergent results are summarised in Box 10.1. The cumulative effect of these features of the measures in divergence in results for services is unclear.

Box 10.1 Contributors to divergence: features of measures

Feature of measures or data collection	Difference between measures	Possible effects
Sampling time frame	CaSPAR: recordings made between 8.15am -8.15pm only, most between 9am – 5pm CaRICE and CCCQ-P: all care included	Services providing a greater proportion of care outside of office hours score lower on CaSPAR than other measures.
Duration of interventions included	CaRICE: only contacts of 5 minutes or more included CaSPAR and CCCQ-P: any contacts included	Services with more brief staff-patient contacts score higher on CaSPAR than CaRICE
Group and individual contact	CaRICE: 1:1 and group contacts not distinguished CaSPAR: all patients in groups counted individually CCCQ-P: care provided in groups may be included by all participants	Services with more group interventions score lower on CaRICE than other measures
Scoring	CaSPAR and CaRICE: no limit on how much different types of care contribute to total score CCCQ-P: 21 categories of care each contribute score of 0-7 to total score of 0-147	Services with the greatest range of care score high on CCCQ-P. Services offering fewer interventions of longer duration score higher on CaRICE and CaSPAR than CCCQ-P
Data collection period	CaSPAR: minimum 3 weeks CaRICE: 5 days CCCQ-P: at least two months (40 respondents needed close to discharge)	CaRICE data may be less representative if data collection occurs during an atypical week. CaRICE may be most vulnerable to Hawthorne effect as change to normal activity required over a short period only.
Completion rates	CaSPAR: 99% (of patients accounted for) CaRICE: 94% (of all staff) CCCQ-P: 70% (of patients approached)	CCCQ-P sample may be less representative because not all patients could be approached and some declined to participate.
Completeness of sample	CCCQ-P: Incomplete sample collected from Alternative 2	Alternatives compare less favourably to standard services on CCCQ-P than other measures for variables where Alternative 2 scored high.

Differences in perspectives of respondent groups

The comparison of results from CCCQ-P and CCCQ-S (Table 5.8) indicated the possibility of systematic difference between staff and patient perspectives on care provided. Staff activity which may be differently appraised by observers, staff and patients – and thus recorded differently in CaSPAR, CaRICE and CCCQ-P – can be envisaged. For example, a member of staff spending half an hour watching television with and intermittently talking to a patient might consider himself to be in contact with the patient for all of that time to assess, engage or check the patient's safety. An observer conducting a momentary time recording may or may not record the staff member as in contact with the patient, depending on whether they were explicitly interacting at the moment of observation. The patient may or may not judge that he has received care. Types of intervention such as observation or assessment, which do not necessarily involve continuous interaction, are potentially particularly vulnerable to inconsistent interpretation as contact or care.

The relative contribution of psychometric limitations, differences in variables measured and differences in respondent group perspectives to divergence in results from study measures is unclear. Uncertainty therefore remains about the extent to which divergence reflects shortcomings of the study or important results.

Subscales

In addition to divergence, a second complication in the interpretation of study results concerns the subscales used in CaRICE and CCCQ-P, described in Section 5.2.4. With both measures, services and service types which scored highly on one subscale item did not necessarily do so on others. For example, measured by CaRICE and CCCQ-P, Alternative 2 ranked highest of all the services for the current activity item but lowest of all on the help with housing item within the same subscale. There is therefore a risk that similar subscale scores could mask substantial differences in service provision between different services or types of service. Findings regarding types of care provided are therefore discussed in Section 10.2 with reference to individual item scores to ensure valid interpretation.

10.1.3 Analysis

Three considerations for the robustness of the study analysis concern: a) the methods used to compare alternatives and standard services; b) the extent to which potential confounding factors were accounted for in analyses; and c) the effect of multiple testing.

Method of comparison

Linear regression was used to analyse CaSPAR (the researcher-completed measure) and CCCQ-P (the patient-report questionnaire) data. This provided a robust statistical comparison of scores from alternatives and standard services which reflected quantity of data and its variance in tests of significance and the width of confidence intervals. The lack of independence in scores derived from each service was also accounted for in analyses through adjustment for clustering by service by calculation of robust standard errors. The width of confidence intervals for results from CaSPAR and CCCQ-P does however leave open the possibility that there are clinically important differences between alternatives and standard services in provision of care, despite study hypotheses not being corroborated.

The method of comparison used with CaRICE data was less robust. An estimate of effect size was used because CaRICE data were insufficiently powered to undertake statistical tests of significance (Section 7.5.2). Estimating effect size as proposed by Cohen (1988), was preferable to just presenting descriptive data because it allowed pre-planned hypothesis testing. It was preferable to setting a threshold for clinically meaningful difference in CaRICE scores for alternatives and standard services and using this for hypothesis testing because estimating effect size takes some account of clustering by service. (Greater variation in individual service scores would increase the standard deviation of all data, increasing the mean difference between alternatives and standard services required to achieve positive effect size estimates.) Medium rather than small effect sizes were required to corroborate study hypotheses to reduce the risk of small, potentially unimportant differences between service types producing positive results.

Estimating effect size is however an unconventional and limited way of comparing alternatives and standard services. Although it provides an established way to describe the magnitude of difference in CaRICE scores between service types, it gives no information about the statistical significance of differences in mean scores or confidence intervals. The resulting uncertainty about whether differences in CaRICE scores were significant or happened by chance is a limitation of this study. The wide confidence intervals for estimated effect sizes in all analyses of CaRICE data reinforce the limitations of this approach as an analysis strategy.

Confounding factors

Differences in care at alternatives and standard services might reflect different service approaches to patient care or a response to differences in the respective patient groups. Multivariate analysis was required to assess the extent to which patient variables moderate relationships between type of service and care received. There were insufficient CaRICE data for multivariate analysis. Adjustment for patient characteristics was not feasible with CaSPAR data. Because CaSPAR recordings provide service-level rather than patient-level data, it would have been necessary to gather data about the characteristics of all patients at each recording and aggregate this to produce patient characteristic variables to include in regression analyses. The help of staff would have been required to gather this information about patients, which, unlike the location of other staff on shift, may not have been readily accessible to an available staff member. To minimise the burden on staff and maximise the feasibility of gathering CaSPAR data, information about patient characteristics were therefore not sought.

Information was provided by CCCQ-P respondents about gender, age, ethnicity and Mental Health Act status and included in analyses. The resource demands of obtaining information from patients about their diagnosis or health status and the likelihood that seeking more complex or personal information from respondents might jeopardise response rates outweighed the potential usefulness of the additional information which might be gained. Permission was not sought from patients to seek further information from staff because of similar concerns about a negative effect on response rates.

While Mental Health Act Status was assessed, diagnosis, severity of illness and needs are all examples of patient variables which could not be accounted for and which might be expected to impact on care. Overall this study could not adequately identify whether differences in care between alternatives and standard services reflect differences in presenting problems of the client groups served or inherently different treatment approaches.

Multiple testing

Six hypotheses were tested in this thesis. This multiple testing increases the risk of Type 1 error, i.e. positive results which occurred by chance rather than reflected genuine differences between alternatives and standard services. A Bonferroni correction was considered for use in analyses to address this problem of multiple testing. In a Bonferroni correction, the standard of proof for each individual comparison is made more stringent in order to maintain a desired total Type 1 error rate: it proposes setting a statistical significance level of 1/n times what it would have been if only one hypothesis were tested (Abdi 2007). In this thesis with six hypotheses, a significance level of .008 (.05/6) would therefore have been required for individual analyses. The significant differences between community alternatives and standard services found by CCCQ-P for psychological and physical and pharmacological interventions in this thesis (Table 8.9) would not have achieved significance had a Bonferroni correction been used.

In this thesis, however, a Bonferroni correction may have been unduly conservative because there is a degree of dependence in the variables measured in different hypotheses. (Total care scores will relate to subscale scores.) The raised threshold for statistical significance set by a Bonferroni correction would have increased the risk of Type 2 errors (Perneger 1998), i.e. where important differences between service types on variables measured failed to achieve statistical significance. For these reasons, a Bonferroni correction was not appropriate in this study. It also could not be applied to CaRICE data, where no tests of statistical significance were possible. Two steps were taken to reduce the risk of Type 1 errors in this investigation. First, hypotheses concerning the nature of care required corroboration from more than one data source. Second, comparisons between alternatives and standard services using the 21 individual categories of care in CaRICE and CCCQ-P were not used in hypothesis testing.

10.1.4 Limitations of the study: conclusions

Study methods reflect a tension between developing new measures to maximise the informativeness of data collected and using established measures to maximise the robustness of data collected. The reasons for developing content of care measures for inpatient services to provide greater depth of information than available from existing measures have been presented (Section 5.1). The CCCQ-S inter-rater reliability test (Section 5.4.2) provided additional evidence that in inpatient settings, a single staff respondent with access to case notes may not be able to provide reliable information about the care provided to individual patients, as would have been required with the use of well-established measures such as the MRC Needs for Care Framework (Brewin et al. 1987). However, a decision to use only established measures with some demonstrated good psychometric properties would have increased confidence in results and could have provided some information about the intensity and nature of care at alternatives and standard services. For example, using the observation measure developed by Tyson and colleagues (1995) and the International Classification of Mental Health Care (DeJong et al 1991) could have achieved this.

The reasons for focusing in this study on the intensity and nature of direct patient care (Section 1.6) and the barriers to developing measures which provide greater depth of information about content of care (Section 5.4.5) have been presented. A broader assessment of the process of care at alternatives and standard services, including attention to service style and continuity of care, would have increased the informativeness of this study about what is provided at services. Qualitative interviews with staff or patients or ethnographic research might provide rich information about elements of the process of care in inpatient settings which may be important to patient experience. Qualitative or case series approaches might also suggest how patient preferences at admission about styles of care or types of intervention impact on the experience of care provided.

The measures used in this investigation are thus a compromise between robustness and informativeness, within the constraints of available resources and the study's chosen focus. The limited state of knowledge about how best to assess content of care in inpatient services (Chapter 4) and knowledge gains from the development of new measures support the use of innovative measures in this thesis.

A major strength of the study is its multi-methods approach. No study identified from the review in Chapter 4 triangulated results from different information sources and data collection methods to the same extent as this investigation in assessing the content of care in mental health services. In particular, the inclusion of a patient perspective in assessing content of care is novel and a strength of the study. The conclusions of Schmidt and colleagues (2000) regarding patient reported outcome measures also apply to content of care measurement: while obtaining data from patients with some forms of mental illness may be difficult and patient-report measures may not always yet meet stringent psychometric criteria, the information provided by a patient's own evaluation of his condition [or experience of care] is unique and valuable. The thesis provided preliminary evidence, not previously available, about the intensity of staff-patient contact and nature of care provided at alternatives and standard acute inpatient services.

10.2 Main findings

The main results from the quantitative comparison of the care provided at four alternatives and four standard acute wards can be summarised as:

- Intensity of staff-patient contact is not significantly different at alternatives and standard services
- Significant differences between alternatives and standard services in provision of social, psychological or physical and pharmacological

interventions were not consistently identified by staff and patient report measures.

- A secondary sub-group comparison suggests community-based alternatives may provide more psychological care and less physical and pharmacological care to patients than standard services.
- Patient satisfaction is positively associated with receipt of social, psychological and physical and pharmacological interventions.
- Measured differences in care provided do not explain greater patient satisfaction at alternatives compared to standard services.

Amount of staff-patient contact

Consistent results from three measures, encompassing staff, patient and researcher-observer perspectives, provide evidence that the intensity of staff-patient contact was not significantly different at alternative and standard services. This applies to both inpatient and community alternatives.

Results suggest that services which organised pre-planned, structured activities (such as recreational and activity groups, formal observations or dispensing medication) provided greatest intensity of staff-patient contact. Services which offered less pre-planned, organised activity did not appear to provide comparable levels of staff-patient contact through other means. For example, CaSPAR data indicated a significantly greater proportion of patients in contact with staff at Alternative 2 (the clinical crisis house) than all other services (Table A6.2); Alternative 2 was the only service to provide a daily programme of recreational group activities within the main residential unit. By contrast, Alternative 3 (the Crisis Team Beds), which scored lowest on CaSPAR, was the only service which provided no therapeutic or recreational group activities.

Types of care provided

This investigation provides no clear evidence of differences in provision of broad types of care between alternatives and standard services. Significant differences were found by both CaRICE and CCCQ-P for only three of twenty one types of care: help with housing and relationships may be greater at alternatives and help with physical health care may be greater at standard services, but overall, similarities in service provision outweigh differences.

Results from this investigation did not suggest that care provided at the Tidal Model ward, Alternative 4, was distinctive or significantly different from standard services. Total and subscale scores for Alternative 4 (Tables 8.11 and 8.12) on all measures were similar to mean scores for standard services (Tables 8.3 and 8.5). It was not an outlier among services in the study on either CaRICE or CCCQ-P for any category of care (Tables A8.1 – A8.8).

The sub-group comparison of community-based alternatives and standard services suggests community alternatives may provide more psychological interventions and less physical and pharmacological care. The magnitude of differences may not be of great clinical importance however. CaRICE data indicates for example (Table 8.10) that community alternatives provide less than 10 minutes more psychological interventions per patient per day than standard services. It is doubtful that ten minutes more per day would satisfy the wishes of service users (Baker 2000, Gilburt et al. 2008) to talk and be listened to. Regarding physical and pharmacological interventions, differences between community alternatives and standard care are found more consistently for physical healthcare, restraint and observation than for items specifically relating to pharmacological treatment of mental health problems.

These items of care within the physical and pharmacological interventions domain of care, where differences between community alternatives and standard wards were consistent and greatest – physical interventions, observations and restraint - are potentially clinically important. People with enduring mental health problems have poor physical health outcomes (Harris and Barrowclough 1998) and make less use of general health services (Jeste et al. 1996, Phelan et al. 2004). Data from this thesis suggest patients admitted to alternatives may get less screening or treatment for physical health problems during their admission than patients on standard wards. Risk management in the forms of observation or restraint is also provided less at alternatives than standard services. Given potential differences in the clinical populations served, the extent of supplementary care provided by other organisations or other forms of risk management practiced, the result of these differences in care are uncertain. They do suggest a need for evaluation of alternatives and standard services to include long-term health outcomes and rare adverse events such as suicide.

Some differences in care provision among community-based alternatives were indicated (Table 8.12). The most marked distinctive features of individual community alternatives indicated by the data are:

- Alternative 1, the non-clinical crisis house, on both measures provided less physical and pharmacological care than all other services. CaRICE and CCCQ-P individual item data (Tables A8.5 and A8.6) suggest that typically at Alternatives 2 and 3 (clinical crisis house and crisis team beds), patients receive a medication review during admission and medication is dispensed daily but this is not the case at Alternative 1.
- Alternative 2, the clinical crisis house scored highest of all services on CaRICE and CCCQ-P for help with current activity, providing 50 minutes more than any other service per patient per day according to CaRICE data (Table A8.1).
- Alternative 3, the Crisis Team Beds scored lowest of all services on CaRICE and CCCQ-P for social interventions. It was the only alternative to score higher than all standard services on both measures for provision of psychological interventions: both data sources indicated more help was provided than at other alternatives with assessment and coping strategies for symptoms.

Comparison of the care provided by individual services is however highly exploratory. Data for individual services are insufficiently powered to assess the significance of differences between services. Each alternative service is a single exemplar of an alternative service model and may not reflect the typical care provided by services of this type.

Patient satisfaction with services

Patient satisfaction was found to be significantly greater at alternatives than standard services, adjusting for measured patient characteristics (Table 9.4). The magnitude of the adjusted mean difference (just under 2 points on CSQ, which has a range of 24 points) does not represent a very stark contrast in the acceptability of alternatives and standard services. Mean CSQ scores at both service types (23.5 at standard services; 25.7 at alternatives) correspond to typical responses of “fairly satisfied” to CSQ questions. CSQ scores provide a global and fairly crude measure of patient satisfaction but this study does not reinforce the poor acceptability of acute wards found by previous studies (Section 1.3).

CCCQ-P total and subscale scores were all positively associated with patient satisfaction with services (Section 9.3). These results may hide more fine-grained distinctions between types of care acceptable and unacceptable to patients. They suggest however, that no broad types of care at standard and alternative services are experienced by patients as aversive. The stronger association with satisfaction found for intensity of care, rather than any type of care, offers evidence for prioritising providing more care, rather than different care, to patients in acute inpatient and residential services. The stronger associations with satisfaction found for psychological and social care than for physical and pharmacological care suggest an increased focus on psychological and social interventions might also increase services’ acceptability to patients. In terms of the conceptual model of Bowers and colleagues (2009), this may suggest that the extent to which secondary admission tasks are considered and an admission bonus achieved are important to patients’ experience of admission, not just the extent to which the primary admission tasks are addressed.

The modest amount of variance in patient satisfaction explained by content of care variables (Table 9.5) suggests factors other than the nature and intensity of care provision are important to patients’ experience of inpatient admission. That greater satisfaction with alternatives than standard services remained significant after adjustment for all CCCQ-P variables (Table 9.5) indicates that measured differences in content of care were not a major cause of alternatives' greater acceptability to patients.

10.3 Implications for policy and practice

Four recommendations for service planners, managers and clinicians can be made from this thesis. First, community alternatives are acceptable to patients. Second, they offer valuable flexibility in acute care provision. Third, providing more staff-patient contact should be a priority for inpatient services and alternatives. Fourth, this investigation provided no evidence to support the implementation of the Tidal Model.

10.3.1 The acceptability of community alternatives

This thesis suggests the provision of community-based alternatives as part of local acute care systems can be supported. The literature review in Chapter 3 found no evidence that community-based alternatives are less clinically effective than standard inpatient services and limited evidence that patient satisfaction may be greater. This investigation also found greater patient satisfaction at alternatives than standard services and limited differences in the extent and nature of care provided at standard services and community alternatives, considered collectively. This suggests that the conclusions of the literature review may remain valid in a contemporary UK context. The community alternatives in this study all demonstrated greater patient satisfaction than standard wards, despite some variation in care provided, the setting of services and the target client group. The potential benefits for acceptability of community-based crisis services may be retained however the service is configured. Community residential crisis services represent a promising service model which may be able to increase service user choice and provide an acceptable alternative to inpatient admission for some patients. These findings offer a degree of support to government guidance (Department of Health 1999, Department of Health 2005) that community alternatives to acute admission, such as crisis houses, should be provided.

10.3.2 Flexibility in care provision at community alternatives

Community alternative services, typically being smaller than inpatient services with fewer tiers of management and often serving a more specific geographic area or population group than hospital acute wards, may offer greater flexibility than standard wards to adapt care provided to meet local needs. Decisions about service provision can be informed by: a) characteristics of the local area; b) target client groups; and c) the available support elsewhere in the acute care system.

Local factors: Two examples how service delivery in alternatives can adapt to address local factors, identified in unpublished qualitative research (which was not part of this thesis) from The Alternatives Study, are:

i) Alternative 1, the non clinical crisis house, is sited in a London borough with a large black population and is run for and by people from black and minority ethnic communities. It explicitly seeks to provide a service where care is less medically orientated than standard wards in order to provide a more acceptable, culturally tailored alternative to hospital admission. This is reflected in its low scores for physical and pharmacological interventions found in this study. It seeks to address the problematic pathways to acute care for black minority ethnic service users, who experience higher rates of detention and police involvement, and over-representation in acute inpatient settings (Morgan et al. 2005).

ii) Alternative 2, the clinical crisis house, is situated in rural Staffordshire, embedded within a Community Mental Health Resource Centre. Its large catchment area and poor local transport links create difficulties in attending daycare or outpatient appointments for many service users. There are few local social or leisure amenities which can be accessed by patients during inpatient admissions. The longer mean length of stay in Alternative 2 than all other alternatives in this study was explicitly identified by the local crisis team manager as a consequence of the more limited accessibility of other support and outpatient care available. The greater provision of structured activity, reflected in its high score for current activity on both CaRICE and CCCQ-P, might also be seen as a response to the paucity and inaccessibility of leisure facilities in the community for people during inpatient stays.

Target client group: Pawson and Tilley (1997) comment that in complex health service interventions, different process interventions may affect outcomes for different patient groups. One source of variation in care provision at alternatives and standard services will be differences in the needs and presentation of patients admitted. For example, there was a smaller proportion of detained patients in community-based alternative services than standard wards (Table 8.2). The lower levels of observations and restraint at community alternatives than standard wards (Tables A8.5 and A8.6) may be appropriate for a typically less high risk and more cooperative client group. A building like a house on a residential street, where swift observation of clients is less easy than in a hospital ward, may also be adequate if a less high risk client group is planned.

Availability of additional services: Two examples of how additional external care can serve to address gaps in care provision at community alternatives are:

i) Alternative 1, the non-clinical crisis house, has a strong working relationship with the local crisis resolution team (CRT) and can arrange for CRT staff to visit patients daily to dispense medication and to provide medication reviews. This allows medical care to be provided to patients despite the absence of trained medical or nursing staff at Alternative 1.
ii) Alternative 3, the Crisis Team Beds are managed by the local CRT who routinely follow up patients at home following discharge from the beds. This facilitates brief admissions to the beds (the briefest of all services in this study) which focus on alleviating the immediate crisis. Help with longer term needs or social systems may then be provided to patients following discharge.

Differences in the overall care received by patients admitted to different acute residential services during a crisis may be slighter than differences in the care provided specifically by the alternative services. Collaboration with other mental health services may allow community alternatives to employ unqualified staff or staff from a smaller range of professional groups without compromising patient care. It may also help reduce length of stay through planned provision of aftercare.

These examples illustrate how the national UK implementation of crisis resolution teams (CRTs) in the last decade provides an opportunity for collaboration with alternatives. Models of collaboration between alternatives and CRTs and potential benefits of such partnerships have been identified (Lloyd-Evans et al. 2008). CRTs may be well placed to provide additional rapid response, frequent, expert interventions to patients at alternatives. Formal arrangements by which integration may be achieved include home treatment teams providing regular sessional input at alternatives, gate keeping beds or managing the residential service directly (as with Alternative 3). As well as providing continuity for patients between inpatient and outpatient care, such collaboration can help clarify referral criteria and processes and provide access to formal and informal supervision and training for alternative service staff.

The variation in care provision found by this investigation among community crisis houses provides a challenge and an opportunity for service planners. Because it may be less clear than for an acute ward what a community crisis house will provide, service commissioners may need to consider more carefully what sort of service they aim to establish and which target client groups should be served. The inclusion of alternatives in a local acute care system may increase the need to establish, and the complexity of, coherent acute care pathways and the nature and extent of collaboration between services. The potential for flexibility in care provision at community alternatives may however enhance the adaptability of a local service system to meet perceived local needs.

10.3.3 The need to prioritise increasing staff-patient contact

Increasing the amount of staff-patient contact on acute wards is an explicitly stated aim of a recent UK government policy document, the Mental Health Policy Implementation Guide for Adult Acute Inpatient Care Provision (Department of Health 2005). The need for this is reinforced by the association between intensity of care received and patient satisfaction found in this thesis.

Service planners should not conclude that inpatient or community alternatives necessarily provide more staff-patient contact than standard wards. Similarities in the organisation of alternatives and standard services may help to explain the negative finding from this thesis regarding greater intensity of care at alternatives. All four alternative services in this study included a staff office separate from the communal areas accessible to patients. All four implemented formal handover meetings in the office at the start and end of shifts, routinely taking all staff away from the parts of the building used by patients for periods of each day and perhaps creating a culture where staff's default location minimises contact with patients. Three out of the four alternatives were run by statutory secondary mental health services: staff therefore had many of the same administrative demands on their time as staff in standard services (such as writing up patient notes, entering details of patient care on electronic information systems and completing assessment documentation). Table 8.1 in shows that over 60% of staff at alternatives were qualified mental health professionals or nursing assistants, similar to staff in standard services. Many staff in alternative services may have received similar training or had previous experience of working on standard wards, with consequent effects on their attitudes and working practices.

A secondary analysis of CaRICE data, reported briefly in Section 8.1 and presented in Appendix 7, suggests staffing levels at services are not the major determinant of the intensity of staff-patient contact. The number of staff per patient at services was comparatively weakly correlated with the minutes of contact provided per patient per day (Pearson's correlation coefficient = 0.40, p = 0.01). A stronger, negative correlation was found between the number of staff per patient and the proportion of time at work staff spent in direct contact with patients (Pearson's correlation coefficient = -0.59, p = <0.001). This analysis is exploratory but provides an indication that if a service increases staffing levels, individual staff may spend less time with patients, so patients experience only a marginal increase in total staff contact. This suggests to service planners that increasing the staffing levels in acute inpatient services is by itself an inefficient way to increase the amount of staff contact received by patients.

Planned activities such as groups, observations and dispensing medication were major contributors to services' total scores on content of care measures in this study. Increasing structured activity on wards should therefore be considered by clinicians and service planners as an effective means to increase staff-patient contact. Whether this style of intervention can meet all the wishes of patients for more time with staff (Baker 2000) is not addressed in this thesis. While structured activity is likely to alleviate boredom, it is less clear whether it can meet the expressed desire of patients to be listened to and understood by staff (Gilburt 2008). Strategies to increase informal contact and 1:1 therapeutic engagement between staff and patients may also be required.

The similarity in proportion of staff time spent with patients found by this and previous studies (Tyson et al. 1995, Higgins et al. 1999), the similar intensity of care at alternatives and standard wards found in this thesis and the limited impact of variation in staffing levels all suggest that influencing the intensity of care provided in acute inpatient services is challenging. Three approaches which managers or service planners could adopt are considered here: commissioning service models more fundamentally different from standard wards than those evaluated in this study; implementing initiatives specifically targeting an increase in staff-patient contact; establishing regular audit and targets for services.

Other alternative service models: Achieving major differences from standard wards in care provided at alternatives may require fundamental differences in setting, personnel or organisation. The narrative survey of alternative service models in Chapter 2 identified services with working practices more different from standard care than any of the alternatives included in this investigation. Two examples are Soteria hostels and family sponsor homes. Soteria houses, established by Loren Mosher in California in the 1970s, employed unqualified staff who typically worked shifts of up to 48 hours including sleeping at the service (Mosher et al. 1975; Mosher 1999). This was specifically designed to help them understand, engage and be alongside the residents. Family sponsor homes, developed by Paul Polak in Colorado in the 1970s, (Polak et al. 1979) dispensed with formal staff as the primary care-givers altogether. Placing patients in a family home provided opportunity for substantial inter-personal contact. While the intensity of contact provided to patients at these specific service models is uncertain (as is evaluation of their effectiveness in a contemporary health service setting), they illustrate that more radically alternative service models than those evaluated in this thesis have been established and could be considered by commissioners and service planners.

Targeted initiatives: Service initiatives which provide a mechanism by which staff-patient contact might be increased, as well as an aim to do so, may be required to change staff practice. Patient Engagement Time (PET) (CSIP 2005) is an example of an initiative designed to increase staff-patient contact on wards, with an explicit process for how this could be done. Developed in inpatient wards but equally applicable to non-hospital services, it involves regular set periods of time where distractions for staff from direct contact with patients are minimised. The staff office is vacated and shut, staff are required to refrain from paperwork or answering the phones and the service is shut to visitors. The effects of Patient Engagement Time on staff-patient contact have yet to be evaluated. The direct focus on how increased contact with patients might be achieved holds promise however.

Audit: Targets and publicised performance data have been widely used in public and private sectors as a spur to effect change and improve performance. Grol and Grimshaw (2003) reviewed approaches to implementing change in healthcare settings, noting the practical difficulties and a lack of clarity about the most effective approaches. They report that feedback on performance, combined with education or reminders, has been recommended. Audit, such as the type conducted on inpatient wards by Ford and colleagues (1998), could be instituted regularly in inpatient services by senior managers or commissioners to provide specific data about activity and staff-patient contact at services. An accreditation scheme for inpatient wards developed by the Royal College of Psychiatrists – AIMS – includes targets for provision of daily structured group activities (Royal College of Psychiatrists 2007). Focus on staff-patient contact in staff performance targets and routine audit might affect the priority with which increasing staff-patient contact is viewed by service managers and the practice of service staff.

10.3.4 Lack of support for the Tidal Model

The evidence from one inpatient ward using the Tidal Model found no indication that it impacted on the content of care provided or the acceptability of the service to patients. The specific aims of the Tidal Model, to increase contact between nursing staff and patients and provide more holistic, less narrowly medically focused care than standard wards (Barker 2001), do not appear to have been met at the service in this study. The literature review in Chapter 3 found a lack of robust evidence regarding the effectiveness of the Tidal Model. This thesis therefore provides no support for its implementation. Guidance from the devolved Welsh Assembly (Welsh Assembly Government 2005) that all psychiatric inpatient wards in Wales should implement the Tidal Model or equivalent is premature. The apparent difficulty in implementing the Tidal Model found by this study may indicate that making changes to culture and practice in hospital acute wards is difficult and increase the appeal to service planners of establishing community-based services.

10.4 Implications for research: measuring content of care

10.4.1 Changes to study design

Some of the limitations of this study identified in Section 10.1 could be addressed in future studies by the following four changes to study design: including more services in the study; increasing the sampling frame for study measures at each service; randomising participants to alternative or standard care; using a repeated measures design in exploring links between content of care and patient satisfaction.

Including more services would have three benefits: a) involving several alternative services of each type could identify whether service types did consistently differ in amount of staff-patient contact or types of care provided; b) CaRICE data could be collected from enough services to allow statistical comparison of alternatives and standard services; and c) a larger study could also generate sufficient CaRICE data to allow factor analysis of aggregated service data for individual categories of care, from which to generate internally consistent subscales, allowing clearer interpretation of subscale scores than possible in this study. Increasing the sampling time frames and sample sizes for content of care measures could increase confidence in the representativeness of data from each service. Randomising patients to alternatives or standard services would minimise the impact of confounding factors on the relationship between service type and care provided and increase clarity that differences between services in care provision (and patient satisfaction) were due to differences in their models of care rather than the differing needs or characteristics of admitted patients. A repeated measures study design, which obtained patient ratings of satisfaction with services and care received at regular intervals during admission, could help clarify the direction of causation for association found. Of the three elements proposed by Bollen (1989) as necessary to establish causation – association, isolation and direction – the exploratory model of satisfaction and content of care achieved the first, the second to a limited extent through adjusting for some confounders, but not the third. While a plausible inference from the association between CCCQ-P and CSQ scores is that care received contributed to patient satisfaction, it is also possible that patient satisfaction affected reported CCCQ-P scores. If in a repeated measures study, changes in care received preceded changes in satisfaction during an admission, this would provide stronger indication of causality than available from this thesis.

The four changes to study design proposed above all require more resources than this investigation. The problems of feasibility with conducting randomised trials in mental health services research have been noted (Gilbody and Whitty 2002) and are particularly challenging for studies involving crisis services and detained patients. A methods paper (Howard et al. 2009) from a recent UK study whose results have not yet been published reflects these difficulties: Howard and colleagues (2009) report that only 41% of eligible patients in their study agreed to be randomised to a crisis house or an acute ward and only a minority of participants could be recruited prior to admission. The literature review in Chapter 3 found studies where randomisation was compromised by unavailability of beds at alternative services (Timko et al. 2006) or pressure to use available beds (Goveia and Tutko 1969). To ensure that differences in service provision did not reflect differences in patient presentation, service level measures like CaRICE and CaSPAR would require all patients admitted to participating services to be randomly allocated. The obstacles to a large scale, genuinely randomised trial of community alternatives and standard services might therefore be prohibitive.

Extending the sample size or recording period for measures at services would increase the demands on staff time, particularly for CaRICE as a staff-completed measure. Studies using event recording measures do not routinely report response rates (Section 4.5.2). The measure described by Patmore and Weaver (1989) reports a 65% response rate over several months from staff in community mental health team settings. The City128 study (Bowers and Simpson 2007) achieved an overall 60% response rate from staff on acute inpatient wards, reporting use of formal observation and untoward incidents each shift over a two-year recording period. There is therefore reason to believe that the 94% CaRICE response rate obtained in this investigation may not be sustainable over longer recording periods. Gains in the representativeness of CaRICE data from extending recording periods might therefore be lost through reduced response rates.

Including more services in a study or using a repeated measures study design would also increase the demands on researcher time. The resource-heavy, time-consuming nature of collecting data using the measures developed for this study may tend to preclude their use in large scale studies (in particular CaRICE, where collecting one week’s data from one service required a full week of researcher time). The information generated in this thesis about CaRICE service scores and their variance could be used to calculate the number of services required for statistical comparison of alternatives and standard services using CaRICE. A sample size calculation provides a means to estimate the number of units needed to detect a certain size of effect, within given Type 1 and Type 2 error rates – typically a significance of 0.05 and minimum power of 80% respectively (Machin et al. 2007). Using each service to represent one unit and calculating effect sizes using the standard deviation from CaRICE data, a sample size calculation indicated a study including 25 services on each arm would give 80% power to detect a large effect size (mean difference = 0.8 standard deviation) at 0.05 level of significance. 59 services on each arm would be required to give 80% power to detect a medium effect size (mean difference = 0.5 standard deviation). This scale of data collection would be feasible only for very highly-resourced studies.

The four changes to study design proposed in this investigation would not address all the limitations of the measures described in Section 10.1.2. Interpretation of results would still be hampered by uncertainty about whether divergence in results from measures reflected differences in respondent groups' perspectives, the variables measured by each instrument or psychometric inadequacies of the measures. Further development of methods to measure content of care is required.

10.4.2 Development of methods of content of care measurement

Two future studies which build on the work of this thesis and could inform future development of content of care measures are proposed: a) to investigate differences in patient, staff and researcher perspectives regarding content of care; and b) to investigate extending the scope of researcher-observation measures.

a) Differences in perspective

Exploration of concordance between staff and patient respondents using CCCQ suggested patients may tend to experience less care than staff think has been provided (Section 5.4.3). The divergence between data obtained from staff, patients and observers in this study (Section 10.1.2) also indicates possible differences in perspective between respondent groups. A single, objective measurement of content of care may not be achievable: as with needs (Slade 1996) or ward atmosphere (Moos 1996), there may be different, valid perspectives. A single, multi-perspective measure of content of care is desirable to allow direct comparison of data from staff and patients. This thesis suggests such a multi-perspective measure may not be possible for assessing content of care in inpatient services however, because individual staff members cannot reliably report what care has been provided to individual patients (Section 5.4.3). Further exploration of how the perspectives of staff, patients and observers differ regarding content of care is therefore required to inform future instrument development and decisions about what account should be taken of differing perspectives in measuring content of care.

Future study 1: Direct comparison of respondent groups' perspectives is hampered by the practical difficulties of obtaining reports of care events from staff, patients and researchers, especially from real-life acute inpatient settings. Three approaches are:

I. Consent could be sought from individual patients for a researcher to shadow them for an agreed period of time. For each contact the patient had with a staff member, the patient, researcher and involved staff member could be asked to describe the nature of care provided using CaRICE categories. Higgins and colleagues (1999) did obtain ethical approval for a researcher to shadow patients on acute wards and did recruit patients to their study (although the researcher did not interact with participating patients in their study), suggesting this approach may be feasible. Differences in perspective about what constitutes a staff-patient contact would not be adequately explored by this approach, but differences regarding the nature of care provided during staff-patient interaction could be investigated.

II. Triads consisting of a consenting patient, staff member and researcher could all sit in a communal area of a ward and complete an observation measure such as the Quality of Interactions Schedule (Dean and Proudfoot 1993) for a defined time period. This would provide information about any differences in perspective about what constitutes a staff-patient contact and about which contacts are perceived as positive, negative or neutral.

III. Patients, staff and researchers could all be recruited and asked to categorise the care described in vignettes derived from case notes using CaRICE categories, in an exercise similar to the one described in Section 5.4.1. This approach might be most feasible and raise fewest ethical considerations, although the complexities of real care events may be inadequately reflected in case note vignettes, compromising the validity of the exercise.

Data collected from studies described above would allow investigation of inter and intra-group differences in perspective about what constitutes contact between staff and patients and the nature of care being provided. It would help clarify whether distinct staff, patient or researcher perspectives about care provided at services exist and how much they differ. Substantial intra-group differences would indicate a need for training respondents in use of measures before collecting data. Inter-group differences would reinforce the need for a multi-methods approach to measuring content of care, including data collected from patients and staff.

b) Extending the scope of observation measures

Three advantages of researcher-observation measures of content of care are; a) demands on the time of service staff are minimised; b) high completion rates can be guaranteed; and c) inter-rater reliability for identifying staff-patient contact has been most clearly established (Section 4.4.2). The challenges to developing feasible and reliable patient or staff-report measures of content of care reinforce the desirability, if possible, of using researcher-observation to assess the nature of care at services as well as its intensity. The validity of this approach would be greater if Future Study 1 (above) found evidence of good correlation between observer and staff and patient perspectives regarding content of care.

Evidence regarding observation measures' ability to provide reliable information about the nature of care in services is limited and mixed. Staff-patient contacts can reliably be categorised by observers as positive, neutral or negative (Shepherd and Richardson 1979, Tyson et al. 1995). In an unreplicated study, Paul (1987) found very good reliability for coding observed staff behaviour during interactions with patients into 21 categories. These mainly related to the style of care (positive or negative, verbal or non-verbal) but provided some information about the types of care being delivered, including group or individual contact and whether physical force was used. The Dementia Care Mapping tool (Kitwood 1997), which categorises patient activities including receiving care, has demonstrated unacceptably low inter-rater reliability (Thornton et al. 2004).

Future study 2 A future study could investigate the inter-rater reliability of records from two researchers observing the same events in inpatient services and categorising different types of care. This would require the following steps:

I. A measure would be developed – for example, based on an existing measure such as the Quality of Interactions Schedule (QUIS) (Dean and Proudfoot 1993). Researchers would record the number and time of each staff contact with patients in a defined area during a pre-planned recording period. The nature of each contact would also be described using categories which might be more easily inferred from observation than those developed for CaRICE, such as: group activity, formal meeting, 1:1 engagement; informal contact, restraint, other.

II. Pairs of researchers would use the measure to record staff-patient contacts in acute settings over the same time period and geographical area (e.g. 40 pairs of researchers for 1 hour per pair).

III. Data from each pair of researchers would be entered into an electronic database. The inter-rater reliability of records for a) number of contacts and b) type of care provided during contacts could then be calculated using Cohen's kappa.

Good inter-rater reliability for categorising the nature of care during contacts would indicate that researcher observation could be used to provide useful information about the nature of care in inpatient services. It has been argued in this thesis (Section 5.2.1) that not all contact between staff and patients in inpatient settings can be observed by researchers, that observable contacts (i.e. in communal areas) may be qualitatively different from unobservable ones and that the proportion of contacts which can be observed may vary across services. Triangulation of results from such new measures with data from a measure like CaSPAR, which measures all staff-patient contact in services, could address some of the concerns with the validity of purely observational measures.

10.5 Implications for research: alternatives to standard wards

Two negative findings from this thesis suggest agendas for future research. First, as neither community nor inpatient alternatives provided more staff time with patients than standard wards, how can the intensity of staff-patient contact be increased in acute inpatient services? Second, if the intensity and nature of care received by patients are weak explanatory factors regarding patient satisfaction and do not explain the greater patient satisfaction with community alternatives than standard wards, what does influence patient satisfaction with acute inpatient care and why do patients like community alternatives more than standard wards? Two proposed future studies will be described which would help to answer these questions. Future directions for broader evaluation of inpatient services will then be discussed.

10.5.1 Enhancing intensity of staff-patient contact

Results suggest specific initiatives with a mechanism for achieving change may be required to increase the intensity of staff-patient contact in inpatient services. Protected Engagement Time (PET) (CSIP 2005) is one such initiative, (described in Section 10.3.3). A UK study has been funded by the National Institute for Health Research (NIHR) Research for Patient Benefit Programme (PB-PG-0808-17014) which will evaluate PET. The author of this thesis is a collaborator on the PET study, which uses content of care measures developed for this thesis. It is described below.

Future study 3

The study will involve three main components

I. A telephone survey of 100 adult acute wards in England will ascertain how widely and in what way PET is implemented. This will be used to clarify an operational definition of PET and establish minimum criteria for adequate model fidelity.
II. A quantitative process and outcomes evaluation will compare 12 acute wards with PET and 12 without. CaSPAR will be used to assess the intensity of staff-patient contact at services. CCCQ-P will be used to assess the frequency and nature of interventions received by patients. Patient satisfaction and perceived recovery, staff burn-out and patient and staff ratings of ward atmosphere will also be assessed using established questionnaire measures with total samples of 300 patients and staff. Main hypotheses will be that on PET wards patients will be more satisfied, staff will be less burnt out, and the mean proportion of patients in contact with staff at any given time will be greater than on wards without PET.
III. Qualitative interviews will be used to explore patient and staff experiences of PET on 3 wards. These will provide depth of information about what happens during periods of protected engagement time, its impact on staff-patient relationships and staff and patients' experience of the ward. They

will explore barriers and facilitating factors to the implementation and effectiveness of PET. Twenty staff and twenty patients will be interviewed.

The PET study will use a naturalistic design: if preliminary evidence for the use of PET is positive, it may inform a future cluster randomised trial evaluating a pre-defined model of PET. The association between intensity of care received and patient satisfaction found in this thesis supports increasing the level of staff-patient contact at inpatient services. The PET study describes how one mechanism to achieve this aim can be evaluated.

10.5.2 Understanding patient satisfaction with inpatient services

Patient satisfaction with services may be explained by compositional or contextual factors (Bjorngaard et al. 2009). Compositional explanations focus on differences in patient characteristics. For example, Leese and colleagues (1998) found evidence from a study of community services in London that the number of outstanding needs patients had was negatively associated with patient satisfaction and a stronger predictor of satisfaction with services than how many of their needs had been met by service interventions. Soergaard and colleagues (2008) found that high levels of social problems were associated with lower satisfaction with services for patients in inpatient settings. Contextual explanations call attention to organisational, physical, cultural and social factors of services. Contextual factors are of particular interest to service planners and providers because they are potentially amenable to change by service managers and staff.

If, as this thesis found, the intensity and broad types of care provided at services do not explain the greater patient satisfaction with crisis houses than inpatient wards and a modest proportion of overall variance in satisfaction, a study exploring other contextual factors influencing patient satisfaction is needed. Qualitative research from the Alternatives Study (not forming part of this thesis) suggested that the quality of relationships patients formed with staff and peers and not feeling unsafe or intimidated during admission were important in patients' experience of admission to alternatives or inpatient wards. A future study, on which the author of this thesis is collaborating, is planned to compare staff-patient and patient-patient relationships at alternatives and standard wards and their association with patient satisfaction. This is described below.

Future study 4 Data will be collected from four community alternative services and four standard inpatient wards. The study will employ mixed methods. Quantitative data will be collected from a total sample of approximately 200 patients (a sample size calculation based on previous use of the study's primary outcome measure will inform the precise sample). Structured measures will be used to assess:

I. Patient satisfaction, using the Client Satisfaction Questionnaire (Attkisson and Zwick 1982)
II. Patient-staff relationships: to provide an overall measure of the quality of relationships with staff in inpatient services, patients will be asked to rate their alliance with for instance three staff they nominate as important in their care, using a measure of therapeutic alliance such as STAR (McGuire-Snieckus et al 2007), with the mean of their scores used in analysis.
III. Patient-patient relationships: perceived conflict with and support from other residents will be rated using an adaptation of a measure of interpersonal relationships such as the Inter-Personal Relationships Schedule (Tilden et al. 1990), asking specifically about relationships with other patients.
IV. Patients' exposure to negative events during admission will be assessed using a schedule derived from Alternatives Study qualitative research, e.g. asking about assault, intimidation, theft, restraint or forced treatment experienced or witnessed during admission.
V. Data about patient characteristics will be collected from patients and case notes, e.g. age, gender, ethnicity, Mental Health Act status, length of stay, diagnosis, symptom severity, previous service use.

Mean scores for alternatives and standard wards will be compared for alternatives and inpatient services. The primary hypothesis is that patient–rated therapeutic alliance is stronger at alternatives than inpatient wards. Variables will then be included in a multivariate model of satisfaction to investigate independent predictors of satisfaction (adjusting for a number of compositional variables), the relative influence of different variables and the amount of variance in patient satisfaction explained by the model.

Qualitative interviews with patients and staff at alternatives and standard services will explore experiences of staff-patient and patient-patient relationships and, in particular, structural and organisational features of services which facilitate or hinder good relationships. A sample of 24 patients and 12 staff is planned.

This proposed study seeks to confirm the finding reported in this thesis of greater patient satisfaction at community alternatives than inpatient wards and help to explain it. Factors enhancing patient satisfaction and relationships with staff and peers at inpatient and community services may be suggested, which could be further evaluated and, where confirmed, incorporated into service models and intervention studies.

10.5.3 Developing an evidence-based model of inpatient care

As proposed in MRC guidance for evaluating complex interventions (Craig et al. 2008), process evaluation can help define service models and inform the design of future service evaluations. The information provided about service content in this thesis can aid consideration of whether alternatives, community alternatives or sub-types of community alternative should be distinguished and compared with standard acute inpatient care in future evaluations. Results suggest that, for two reasons, attempts to define and evaluate a number of different models of acute inpatient care and compare their effectiveness may not be the most useful way forward to establishing effective inpatient care. First, the similarities in care provision between alternatives and standard services appear to be greater than the differences, many of which may result more from local factors and the presentation of patients than conceptual model differences. Second, the findings from development of methods of measuring content of care suggest obtaining reliable information about relatively subtle differences between service models in service provision is problematic.

A more profitable next step in evaluating inpatient care may be to establish a single model of good inpatient care, then test its association with service outcomes. The Evidence Based Practices (EBP) Programme in the USA (Mowbray et al. 2003, Mueser et al. 2003) gives a blueprint for how conceptually robust, testable models of mental health services can be developed and then evaluated. The EBP approach has been applied with a number of complex mental health interventions, including Assertive Community Treatment (McGrew et al. 1994) and supported employment (Bond et al. 1997), but not acute inpatient care.

Future study 5: The stages of developing an evidence-based model using the EBP approach (McHugo et al. 2007) could be applied to acute inpatient care as follows:

I. Develop a model of inpatient care, based where possible on empirical evidence about the critical ingredients resulting in best outcomes; where this is not available, based on expert consultation and qualitative evidence from stakeholders including service users, carers and clinicians.
II. Develop a scale to measure fidelity to the model: this is likely to assess input factors such as staffing levels and experience, organisational aspects of the service and its working with other parts of the mental health care system and the interventions provided to clients. Information sources may include service protocols, case note audit, interviews with managers, staff or patients.
III. Test whether model adherence is associated with better outcomes
IV. Develop and implement a resource kit to help services achieve high model fidelity: this may include providing written guidance and learning materials, training to staff and support and supervision to service managers focusing on implementing change.
V. Evaluate the ability of the implementation resource kit to enhance model fidelity and recheck associations between model adherence and outcomes.

The EBP research into Assertive Community Treatment (ACT) illustrates how content of care measurement has a role to play in the development and evaluation of EBP service models. An event recording measure, the Daily Contact Log, was used to describe proto-ACT services (Brekke 1987) and inform the development of a model of ACT, and has subsequently been used to test criteria relating to frequency of contact in model fidelity measurement (McHugo et al. 1999). Results from this thesis suggest that requirements for intensity of staff-patient contact should be included in a model of inpatient care and assessed using content of care measures in a fidelity scale, at least if patient satisfaction is seen as an important outcome.

The approach of this thesis towards an ultimate goal of establishing how inpatient services should achieve good outcomes involved: first, developing measures of the care provided in inpatient services; second, using these measures to explore associations between content of care variables and outcomes; leading to third, developing empirically-based service models. The EBP approach by contrast, first develops a service model and then seeks to establish its empirical support. Two disadvantages of the EPB approach are:

a) It involves considerable investment in developing a service model and fidelity measures with no guarantee that the model is effective. If subsequent evaluation of the model finds no positive associations between model fidelity and good outcomes, the process may need to be repeated and the model revised.
b) Even if a pre-defined service model demonstrates effectiveness, it may omit other variables which influence service outcomes. These might have been identified had more comprehensive process measurement been used to explore associations with outcomes before a model was developed.

An advantage of the EBP approach is that it provides a clearly defined service model, based on best available information, which guides more focused process measurement and can be rigorously evaluated. The methodological challenges and resource implications of comprehensive content of care measurement in inpatient services suggest that a more prescriptive, EBP-style approach may be a better means to approach the goal of establishing effective inpatient care. If future research found that a model of inpatient care which had been demonstrated to be associated with good outcomes in inpatient settings could also be adhered to and was associated with good outcomes in community alternatives, then the conclusion that a single model of acute inpatient care was sufficient would be warranted. If not, additional model(s) of community-based residential acute care could be developed, evaluated and compared for defined clinical populations.

Chapter 11: Conclusion

This thesis has added to knowledge about alternative services and methods to measure content of care. It informs direction for future evaluations of acute inpatient care.

11.1 Alternative services

Alternative services may experience a tension between seeking to provide a genuine alternative to inpatient care – i.e. fulfilling the same functions as and catering for a similar population to standard acute wards – and being genuinely alternative – i.e. providing a distinctly different experience of admission and care for people once they are admitted. In this thesis, no hypotheses about differences in care between alternatives and standard services were corroborated and similarities in care were more marked than differences. These results are congruent with the Alternatives Study UK national survey (Johnson et al. 2009), which found that UK alternatives typically have high levels of integration with local acute care systems, considerable overlap in populations with acute wards and lack a defined therapeutic model. Together, these findings support a view that in a current UK context, alternatives may be capable of forming part of mainstream acute care provision but may not provide fundamentally different care from standard services. The findings from this thesis suggest that the rhetoric surrounding alternatives may not always reflect the reality. For example, this study found little impact on care from the implementation of the Tidal Model in one inpatient ward and no evidence that community community-based alternatives provided more staff-patient contact than standard services, despite assumptions that they would by local managers and service planners (found in qualitative interviews for the Alternatives Study, not forming part of this thesis).

A literature review (Chapter 3) found preliminary evidence for greater patient satisfaction with community alternatives than with standard inpatient services. This was also found in a current UK context by the quantitative investigation in this thesis, for all community alternatives assessed despite some variation in service set-up and care provision. The evidence of some stability for this finding increases confidence in the acceptability to patients of community alternatives. This thesis could not adequately explain why patients are more satisfied at community alternatives than standard wards but suggests it is not because of differences in the intensity or broad nature of care provided. This indicates a need for future research to investigate the relationship to patient satisfaction and service type of process and input variables not assessed in this thesis.

Current evidence is insufficient to draw conclusions about the clinical effectiveness of alternatives (Chapter 3). The evidence synthesised and generated by this thesis does suggest however, that community alternatives are a promising service model: they are not contra-indicated by any current evidence and offer a means to increase choice and the acceptability of acute care for some patients with acute mental health problems.

11.2 Content of care measurement

The development of outcomes measurement in mental health services research in recent decades has been accompanied by repeated calls for process measurement, and assessment of service content in particular, to understand variation in implementation and in service outcomes (Donabedian 1966, Tugwell 1979, Brekke 1987, Mechanic 1996, Campbell et al. 2000, Craig et al. 2008, Burns et al. 2009). Yet process measurement remains a comparatively neglected field: the review in Chapter 4 found no consensus about content of care measurement methods or gold standard measures and limited research investigating associations between service content and outcomes. The review of measures and development of new measures for this thesis found that measuring what services do is a deceptively simple aim. Valid categorisation of service interventions, accounting for different perspectives, achieving reliable measures and limiting the resource demands of measurement methods to feasible proportions all present challenges which have not been fully met. In inpatient settings, the multiple care-givers for individual patients and acute environment, with unwell patient respondents and emergencies requiring immediate staff attention, create additional problems. The difficulty of obtaining adequately informative, valid and reliable information may explain why relatively little attention has been paid to process measurement in mental health services, despite it being frequently advocated.

Developing a single, simple, validated, multi-perspective measure of content of care, as has been achieved for measures of needs (Slade et al 1999), ward atmosphere (Moos 1996) and therapeutic alliance (McGuire-Snieckus et al 2007), was not achieved in this thesis. The means to achieve this desirable goal in the future are not wholly clear. However, three conclusions about how content of care assessment in inpatient settings should be conducted in future research can be drawn from this thesis

1) Seeking information of some depth about the care provided in inpatient settings is feasible. Response rates in this thesis from patient and staff measures were good (Section 8.1). Staff can categorise care provided into 21 types with comparable inter-rater reliability (Table 5.6) to that found for a previous measure with seven categories (Brekke 1987).
2) Staff should only be asked to provide information about care they have personally provided. CCCQ-S reliability testing (Table 5.7) indicated that a key worker such as a named nurse does not have sufficient knowledge about all the care received by a patient during an inpatient admission to provide reliable information.
3) Content of care may be subjectively experienced with differing, valid perspectives. Staff and patient reports showed poor agreement with a trend for patients to experience less care as received than staff report as provided (Table 5.8). Divergence in results from CaSPAR, CaRICE and CCCQ-P reinforce the suggestion that researchers, staff and patients may not concur about what constitutes care or its nature. A multi-methods approach to content of care measurement is therefore required which accounts for different perspectives.

Further development of content of care measurement methods is required. The measures developed for this thesis do provide means to achieve desirable triangulation of data from different perspectives about the content of care in inpatient services, greater depth of information than was available from existing measures and preliminary evidence of inter-rater reliability for a staff-report measure, CaRICE.

11.3 Future inpatient service evaluation

The quantitative investigation in this thesis was a naturalistic, non-randomised study and involved only one exemplar of each of several different sorts of alternative service, in terms of the typology developed by Johnson and colleagues (2009). These limitations preclude definite conclusions about whether any differences in service content found by the study reflect systematic differences in service models or merely variation in implementation and populations served among broadly similar services. However, the extent of differences found between alternatives and standard services was modest. This does not provide a clear rationale for future studies to prioritise distinguishing and comparing different models of inpatient acute care. Defining one model and evaluating how model fidelity relates to outcomes may be a preferable next step towards establishing effective acute inpatient care. The Evidence Based Practices (EBP) programme (Mueser et al. 2003) provides a blueprint for how to do this. Refinements to the model for different clinical populations and evaluation of any additional effects from a non-hospital service setting could follow. Such an approach would still require measurement of service content in order to assess model fidelity.

This thesis only explored how service content relates to one outcome – patient satisfaction. Its contribution to understanding what constitutes effective inpatient care is therefore limited as it does not address most of the broad range of outcome domains necessary for service evaluation (Higginson 1994, Gilbody et al. 2002). Patient satisfaction is an important outcome for evaluating inpatient care though, because of high service-user dissatisfaction with standard inpatient care (Chapter 1) and an increasing emphasis that service evaluation should not focus on narrowly-defined clinical outcomes but should assess its impact on people’s whole-person needs, accounting for their preferences (CSIP 2008). This thesis provides preliminary evidence for elements which therefore should be included in a model of inpatient care. Most clearly, it supplies empirical corroboration, not previously established, of the findings from qualitative research of the impact of availability and time with staff on patients' experience of admission. The intensity of staff-patient contact was found to be more strongly related to patient satisfaction than the broad types of intervention patients received (Table 9.5). This supports the importance of "presence" as a component of inpatient care, identified in the conceptual model by Bowers and colleagues (2009). Targets for intensity of staff contact with patients should be included in any EBP-style model of inpatient care and fidelity measurement. The stronger relationship to satisfaction for psychological and social interventions than physical and pharmacological interventions (Table 9.5) may further indicate that requirements for multi-disciplinary care and demonstration of service focus on a broadly defined range of patient needs should be included in a model of inpatient care.

Through literature reviews, instrument development and quantitative data collection, this thesis has contributed to achieving two goals for mental health services research: developing measures to assess the content of care delivered by services and establishing evidence-based models of acute inpatient and residential care. The thesis can thus contribute to the overarching goal of more effectively supporting people experiencing mental health crises.

References

Abdi,H. (2007) *“Bonferroni and Sidak corrections for multiple comparisons”* in Salkind,N. ed “Encyclopedia of measurement and statistics” Thousand Oaks, Sage

Abramowitz,S., Obten,N., & Cohen,H. (1998), "Measuring HIV/AIDS Case Management", *Social Work in Healthcare*, vol. 27, no. 3, pp. 1-28.

Altman,D. (1991) *“Practical statistics for medical research”* Chapman and Hall, London

Altman,D. & Bland,J. (1996) "Detecting skewness from summary information.", *British Medical Journal*, vol. 313, p. 1200.

Alwan,N., Johnstone,P. & Zolese,G. (2008) *"Length of hospitalisation for people with severe mental illness"* Cochrane Database of Systematic Reviews

Arce,A. & Vergare,M. (1985) "An overview of community residences as alternatives to hospitalization", *Psychiatric Clinics of North America* vol. 3 pp 423-436.

Attkisson,C. & Zwick,R. (1982) "The Client Satisfaction Questionnaire: psychiattric properties and correlations with service utilization and psychotherapy outcome", *Evaluation and Program Planning*, vol. 5, pp. 233-237.

Baker,S. (2000) *“Environmentally friendly? Patients views of conditions on psychiatric wards”* MIND publications

Barker,P. (2001) “The Tidal Model: a person-centred approach to psychiatric and mental health nursing” *Perspectives in Psychiatric Care* vol. 37(3) pp 79-87

Bebeau,C. (1971) *“Development and validation of a scale to measure treatment effectiveness in psychotherapy”* Rocky Mountain Psychological Association Convention, Denver.

Beecham,J. & Knapp,M. (1992), "Costing Psychiatric Interventions," in *Measuring Mental Health Needs*, G. Thornicroft & C. W. J. Brewin, eds., Gaskell, pp. 163-183.

Bennett,R. (2002) “The crisis home program of Dane County” In *“Alternatives to the Hospital for Acute Psychiatric Treatment”,* ed. R. Warner. Washington DC: American Psychiatric Press, pp. 227–36.

Berger,J. (2006) "Incorporation of the Tidal Model into the interdisciplinary plan of care - a program quality improvement project", *Journal of Psychiatric and Mental Health Nursing*, vol. 13, no. 4, pp. 464-467.

Bittle,R. (1986) "Acute care treatment services (ACTS): A model program for providing acute psychiatric services in a homelike environment in an institutional setting", *Behavioral Residential Treatment* , vol. 1, no. 4, pp. 275-288.

Bjørngaard, J.H., Rusted, A-B, Kjelisberg, E. (2009) The prisoner as patient – a health services satisfaction survey. BMS HSR 9: 176

Boardman,A., Hodgson,R., Lewis,M. & Allen,K. (1999) "North Staffordshire Community Beds Study: longitudinal evaluation of psychiatric in-patient units attached to community mental health centres. I: Methods, outcome and patient satisfaction", *British Journal of Psychiatry* vol. 175, pp. 70-78.

Bola,J. & Mosher,L. (2003) "Treatment of acute psychosis without neuroleptics: two-year outcomes from the Soteria project", *Journal of Nervous and Mental Disease* vol. 191, no. 4, pp. 219-229.

Bollen,K. (1989) *“Structural equations with latent variables”* Wiley and Sons, New York

Bond GR, Becker DR, Drake RE, Vogler KM. (1997) “A fidelity scale for the individual placement and support model of supported employment” *Rehabilitation Counseling Bulletin* vol.40(4) pp 265-84.

Bond,G. & Salyers,M. (2004), ""Prediction of outcome from the Dartmouth Assertive Community Treatment fidelity Scale"", *CNS-Spectrums*, vol. 9, no. 12, pp. 937-942.

Bonsack,C., Borgeat,F. and Dubois,A. (2001) “Impact of mental health services' specialization by diagnosis in Lausanne: I. Qualitative Study.” *Annales Medico-Psychologiques* vol. 159 pp 645-651

Bowers (2005) “Reasons for admission and their implications for the nature of acute inpatient psychiatric nursing” *Journal of Psychiatric and mental health nursing* vol. 12 pp 231-236

Bowers,L.; Simpson,A.; Alexander,J.; Hackney,D.; Nijman,H.; Grange,A.; Warren,J. (2005b) “The nature and purpose of acute psychiatric wards: the Tompkins acute wards study” *Journal of Mental Health* vol. 14(6) pp 625-635

Bowers,L., Brennan,G., Flood,C., Lipang,M., Oladapo,P. (2006) "Preliminary outcomes of a trial to reduce conflict and containment on acute psychiatric wards: City Nurses" *Journal of Psychiatric and Mental Health Nursing* vol.13 pp 165-172

Bowers,L. and Simpson,A. (2007) *“*Observing and engaging*” Mental Health Practice* vol.10(10) pp12-14

Bowers,L.; Chaplin,R.; Quirk,A.; Lelliott,P. (2009) “A conceptual model of the aims and functions of acute inpatient psychiatry” *Journal of Mental Health* vol. 18(4) pp 316-325

Bowie, P. & Mountain, G. (1993), "Using Direct Observation to Record the Behaviour of Long-Stay Patients with Dementia", *International Journal of Geriatric Psychiatry*, vol. 8, pp. 857-864.

Braun,P., Kochansky,G., Shapiro,R., Greenberg,S., Gudeman,J., Johnson,S., Shore,M. (1981) “Overview: Deinstitutionalisation of psychiatric patients, a critical review of outcome studies” *American Journal of Psychiatry* vol.138(6) pp736-749,

Brekke,J. (1987), "The Model-Guided Method for Monitoring Program Implementation", *Evaluation Review*, vol. 11,(3) pp. 281-299.

Brekke,J. & Test, M. (1992), "A model for measuring the implementation of community support programs: results from three sites", *Community Mental Health Journal*, vol. 28, no. 3, pp. 227-247.

Brekke,J. & Long,J. (1997) "The impact of service characteristics on functional outcomes from Community Support Programs for persons with schizophrenia: a growth curve analysis" *Journal of Consulting and Clinical Psychology* vol.65 no.3 pp 464-475

Brekke,J., Ansel,J., Long,J. Slade,E., Weinstein,M. (1999) "Intensity and continuity of services and functional outcomes in the rehabilitation of persons with schizophrenia" *Psychiatric Services* vol.50 no. 2 pp 248-256

Brewin,C.; Wing,J.; (1987) “Principles and practice of measuring need in the long-term mentally ill: The MRC Needs for Care Assessment” *Psychological Medicine* vol. 17 pp 971-982

Brook,B. (1973) "Crisis hostel: an alternative to psychiatric hospitalization for emergency patients", *Hospital and Community Psychiatry*, vol. 24, no. 9, pp. 621-624.

Brugha,T. & Glover,G. (1998), "Process and Health Outcomes: Need for Clarity in Systematic Reviews of Case Management for Severe Mental Disorders", *Health Trends*, vol. 30, no. 3, pp. 76-79.

Brugha,T. & Lindsay,F. (1996), "Quality of mental health service care: The forgotten pathway from process to outcome", *Social Psychiatry and Psychiatric Epidemiology*, vol. 31, no. 2, pp. 89-98.

Bryman, A. (2004), *Social Research Methods*, 2 edn, Oxford University Press.

Burns,T. & Priebe,S. (1996), "Mental health care systems and their characteristics: A proposal", *Acta Psychiatrica Scandinavica*, vol. 94, no. 6, pp. 381-385.

Burns,T., Fiander,M., Kent,A., Ukoumunne,O., Byford,S., Fahy,T., & Kumar,K. (2000), "Effects of case-load size on the process of care of patients with severe psychotic illness. Report from the UK700 trial", *British Journal of Psychiatry*, vol. 177, pp. 427-433.

Burns,T.; Catty,J.; Dash,M.; Roberts,C.; Lockwood,A.; Marshall,M. (2007) “Use of intensive case management to reduce time in hospital for people with severe mental illness: systematic review and meta-regression” *British Medical Journal* vol.335 pp 336-340

Burns,T. “End of the road for treatment as usual studies?” *British Journal of Psychiatry* vol.195(1) pp 5-6

Calton,T., Ferriter,M., Huband,N., Spandler,H. (2008) “A systematic review of the Soteria paradigm for the treatment of people diagnosed with schizophrenia” *Schizophrenia Bulletin* vol. 34(1) 181-192

Campbell,M.; Fitzpatrick,R.; Haines,A.; Kimmonth,A.; Sandercock,P.; Spiegelhalter,D. (2000) “Framework for the design and evaluation of complex interventions to improve health” *British Medical Journal* vol. 321 pp 694-6

Care Services Improvment Partnership (2005) *“Report of the Acute Care Collaborative”* www.londondevelopmentcentre.org.uk/silo/files/458.pdf

Care Services Improvement Partnership (2008) *“Laying the foundations for better acute mental health care”* www.nimhe.csip.org.uk/silo/files/csip-finalpdf.pdf

Catty,J.; Winfield,H.; Clement,S. (2007) “The therapeutic relationship in secondary mental health care: a conceptual review of measures” *Acta Psychiatrica Scandinavica* vol. 116(4) pp 238 – 252

Chicchetti,D. & Feinstein,A. (1990) “High agreement but low kappa 2: resolving the paradoxes” *Journal of Clinical Epidemiology* vol. 43(6) pp 551-558

Ciompi,L., Dauwalder,H., Maier,C., Aebi,E., Trutsch,K., Kupper,Z., Rutishauser,C. (1992) “The Pilot Project “Soteria Berne” Clinical Experiences and Results” *British Journal of Psychiatry* vol 161 (suppl. 18) pp145-153

Ciompi,L., Kupper, Z., Aebi,E., Dauwalder,H., Hubschmid,T ., Trutsch,K., & Rutishauser,C. 1993, "The pilot project "Soteria Berne" in treatment of acute schizophrenic patients. II. Results of a comparative prospective follow-up study over 2 years", *Nervenarzt*, vol. 64, no. 7, pp. 440-450.

Ciompi,L. & Hoffmann,H. (2004) “Soteria Berne. An innovative milieutherapeutic approach to acute schizophrenia based on the concept of affect-logic” *World Psychiatry* vol 3, pp 140–6.

Clifford,P. (1988) “M is for outcome: the CORE outcomes initiative” *Journal of Mental Health* vol. 7 pp 19-4

Cochrane Collaboration (2003) *“Review Manager (RevMan) [computer program] Version 4.2 for Windows”* Copenhagen, The Nordic Cochrane Centre

Cochrane Collaboration (2006) *“Cochrane handbook for systematic reviews of interventions version 4.2.6”* (www.cochrane.org.resources.handbook)

Cohen,J. (1988) “*Statistical power analysis for the behavioural sciences”* (2nd ed.) L.Erlbaum Associates, New Jersey

Cook,N. Phillips,B., Sadler,D (2005) “The Tidal Model as experienced by patients and nurses in a regional forensic unit” *Journal of Psychiatric and Mental Health Nursing* vol.12(5) pp 536-540

Craig,P., Dieppe,P., Macintyre,S., Mitchie,S., Nazareth,I., Petticrew,M. (2008) “Developing and evaluating complex interventions: the new Medical Research Council guidance” *British Medical Journal* vol. 337 pp 979 - 983

Crawford,M.; deJong,E.; Freeman,G.; Weaver,T. (2004) “Providing continuity of care for people with severe mental illness: a narrative review” *Social Psychiatry and Psychiatric Epidemiology”* vol. 39 pp 265-272

Davidge,M.; Elias,S.; Hayes,B. (1993) *“Survey of English mental illness hospitals”* Birmingham, Inter-Authority Comparisons and Consultancy, University of Birmingham

Davies, S., Presilla, B., Strathdee, G. and Thornicroft, G. (1994). Community beds: the future for mental health care. *Social Psychiatry and Psychiatric Epidemiology* vol. 29, pp 241–3.

Dean,R. & Proudfoot,R. (1993), "The Quality of Interactions Schedule (QUIS): Development, Reliability and Use in the Evaluation of Two Domus Units", *International Journal of Geriatric Psychiatry*, vol. 8, pp. 819-826.

Deeks, J., Dinnes, J., D'Amico, J., Sowden, A., Sakarovitch, C., Song, F., Petticrew, M., & Altman, D. (2003) "Evaluating non randomised intervention studies" *Health Technology Assessment*, vol. 7, no. 27.

De Jong,A. (2000), "Development of the International Classification of Mental Health Care (ICMHC)", *Acta Psychiatrica Scandinavica*, vol. 102, Supplement 405, pp. 8-13.

De Jong,A., Giel,R., & Tentlom,G. (1991), *"International classification of mental health care: a tool for classifying services providing mental health care Part 1"* Dept Social Psychiatry, University of Groningen and World Health organisation Regional office for Europe, Groningen.

Department of Health (1998) *“Modernising Mental Health Services”* London, The Stationery Office

Department of Health (1999) *“A national service framework for mental health”* Department of Health

Department of Health (2005) *“Mental Health Policy Implementation Guide: Adult Acute Inpatient Care Provision”* Department of Health

Department of Health (2006) “*Hospital Activity Statistics”* (web publication) http://www.performance.doh.gov.uk/hospitalactivity

Department of Health (2009) “Performance data and statistics: beds” (web publication)http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/Performancedataandstatistics/Beds/index.htm

Dietzen,L. and Bond,G. (1993) "Relationship between case manager contact and outcome for frequently hospitalised psychiatric clients" *Hospital and Community Psychiatry* vol.44 no. 9 pp 839-843

Dodds, P. & Bowles, N. (2001) "Commentary. Dismantling formal observation and refocusing nursing activity in acute inpatient psychiatry: a case study", *Journal of Psychiatric and Mental Health Nursing*, vol. 8, no. 2, pp. 183-188.

Donabedian,A. (1966) “Evaluating the quality of medical care” *Milbank Memorial Fund Quarterly* vol.44(3) pp 166-203

Donabedian, A. (1980) “*Explorations in Quality Assessment and Monitoring, Volume 1: the Definition of Quality and Approaches to its Assessment”* Health Administration Press.

Donabedian, A. (1992) “The role of outcomes in quality assessment and assurance” *Quality Review Bulletin*, November pp 356-360

Dunn, G. (1989) *“Design and analysis of reliability studies”* London: Arnold.

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*, vol. 45, pp. 242-247.

Eisen, S. V., Wilcox, M., Idiculla, T., Speredelozzi, A., & Dickey, B. (2002) "Assessing consumer perceptions of inpatient psychiatric treatment: the perceptions of care survey", *Joint Commission Journal of Quality Improvement*, vol. 28, pp. 510-526.

Ellis, R. H. & Wilson, N. C. (1973) "Evaluation of treatment effectiveness using a Goal-oriented Automatic Progress Note", *Evaluation*, Special Monograph No. 1, pp. 6-11.

Ellwood, P. (1998) “Outcomes management – a technology of patient experience” *New England Journal of Medicine* vol. 318 pp1549-1556

Endicott,J. (1976) “The Global Assessment Scale” *Archives of General Psychiatry* vol. 33 pp 766 - 771

Endicott, J. & Spitzer, R. L. (1972) "What! Another rating scale? The Psychiatric Evaluation Form", *Journal of Nervous and Mental Disease*, vol. 154, pp. 88-104.

Engel,G.L. (1977) “The need for a new medical model: a challenge for biomedicine” *Science* vol. 196(4286) pp129-136

Fakhoury,W. & Priebe,S. (2000) “The process of deinstitutionalisation: an international overview” *Current Opinion in Psychiatry* vol. 15 pp 187-192

Fakhoury,W. & Priebe,S. (2007) “Deinstitutionalisation and reinstitutionalisation: major changes in the provision of mental health care” *Psychiatry* vol.6(8) pp 313-316

Falloon, I. & Fadden, G. (1995), “*Integrated Mental Health Care: a comprehensive community-based approach”* Cambridge University Press.

Faulkner, A. Petit-Zeman,S. Sherlock,J. Wallcraft,J. (2002) *“Being there in a crisis: a report of the learning from eight mental health crisis services”* The Mental Health Foundation and Sainsbury Centre for Mental Health

Feinstein,A. & Chicchetti,D. (1990) “High agreement but low kappa 1: the problems of two paradoxes” *Journal of Clinical Epidemiology* vol. 43(6) pp 443-449

Fenton, W. S., Mosher, L. R., Herrell, J. M., & Blyler, C. R. (1998) "Randomized trial of general hospital and residential alternative care for patients with severe and persistent mental illness", *American Journal of Psychiatry* vol. 155, no. 4, pp. 516-522.

Fenton, W. S., Hoch, J. S., Herrell, J. M., Mosher, L., & Dixon, L. (2002) "Cost and cost-effectiveness of hospital vs residential crisis care for patients who have serious mental illness", *Archives of General Psychiatry* vol. 59, no. 4, pp. 357-364.

Fiander, M. & Burns, T. (2000), "A Delphi approach to describing service models of community mental health practice", *Psychiatric Services* vol. 51, no. 5, pp. 656-658.

Fisher, G., Landis, D., & Clark, D. (1988), "Case Management Service Provision and Client Change", *Community Mental Health Journal*, vol. 24, no. 2, pp. 134-142.

Flannigan, C., Glover, G., Wing, J. et al (1994) “Inner London Collaborative Audit of admissions in two health districts. III Reasons for acute admission to psychiatric wards” *British Journal of Psychiatry* vol. 165 pp 750-759

Ford,R., Durcan,G., Warner,L., Harvey,P., Muijen,M. (1998) “One day survey by the Mental Health Act Commission of acute adult psychiatric wards in England and Wales” *British Medical Journal* vol. 317 pp 1279 – 1283

Freeman,C. & Tyrer,P. (1992) *“Research methods in Psychiatry”* Gaskell, London

Garcia,I., Kennett,C., Quraishi,M., Durcan,G., (2005) *“Acute care 2004: a national survey of adult psychiatric wards in England”* Sainsbury Centre for Mental Health

Gilbody,S.; House,A.; Sheldon,T. (2002) “Routine administration of health-related quality of life (HRQOL) and needs assessment instruments to improve psychological outcome: a systematic review” *Psychological Medicine* vol. 32 pp 1345-1356

Gilbody, S. & Whitty, P. (2002) "Improving the delivery and organisation of mental health services: beyond the conventional randomised controlled trial." *British Journal of Psychiatry*, vol. 180, pp. 13-18

Gilburt,H., Rose,D., Slade,M. (2008) “The importance of relationships in mental health care: a qualitative study of service users’ experiences of psychiatric hospital admission in the UK” *Biomed Central Health Services Research e-journal*: 8:92

Gilliland, A. & Colgin, R. (1951) "Norms, reliability and forms of the MMPI", *Journal of Consulting Psychology*, vol. 15, no. 5, pp. 435-438.

Glick,I., Hargreaves,W. & Raskin,M. (1975) “Short versus long hospitalization 2: results for schizophrenia in-patients” *American Journal of Psychiatry* vol. 123 pp 385-390

Glick,I., Burti, L., Suzuki, K., & Sacks, M. (1991), ""Effectiveness in Psychiatric Care: 1. A cross-national study of the process of treatment and outcomes of major depressive disorder"", *Journal of Nervous and Mental Disease*, vol. 179, no. 2, pp. 55-63.

Glover G, Arts G, Babu KS. (2006) Crisis resolution/home treatment teams and psychiatric admission rates in England. *British Journal of Psychiatry* vol. 189: 441-445

Goffman,E. (1961) *“Asylum”* London, Penguin

Gordon,W., Morton,T., & Brooks,G. (2005) "Launching the Tidal Model: evaluating the evidence", *Journal of Psychiatric and Mental Health Nursing*, vol. 12, no. 6, pp. 703-712.

Goveia,L. & Tutko,T. (1969) *"Psychiatric Rehabilitation in a Community Center"* Rehabilitation Mental Health Services, Inc., Santa Clara, Calif.

Grawe,R., Ruud,T., & Bjorngaard,J. (2005) "Alternative akuttilbud i psykisk helsevern for voksne. [Alternative emergency interventions in adult mental health care]", *Tidsskrift for den Norske laegeforening*, vol. 125, no. 23, pp. 3265-3268.

Grol,R. & Grimshaw,J. (2003) “From best evidence to best practice: effective implementation of change in patients’ care” *The Lancet* vol. 362(9391) pp. 1225 - 1230

Hamilton,M. (1960) "A rating scale for depression", *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 23, pp. 56-62.

Hansson,K., Allebeck,P., & Malm,U. (2001), "Event recording in psychiatric care: Developing an instrument and 1-year results", *Journal of Nordic Psychiatry*, vol. 55, no. 1, pp. 25-31.

Harris,E. and Barrowclough,B. (1998) “Excess mortality of mental disorder” *British Journal of Psychiatry* vol. 173 pp 11-53

Hawks,D. (1975) “Community care: an analysis of assumptions” *British Journal of Psychiatry* vol. 127 pp. 276-285

Hawthorne,W., Green,E., Lohr,J., Hough,R., & Smith,P. (1999) "Comparison of outcomes of acute care in short-term residential treatment and psychiatric hospital settings", *Psychiatric Services* vol. 50, no. 3, pp. 401-406.

Hawthorne,W., Green,E., Gilmer,T., Garcia,P., Hough,R., Lee,M., Hammond,L., & Lohr,J. (2005) "A randomized trial of short-term acute residential treatment for veterans", *Psychiatric Services* vol. 56, no. 11, pp. 1379-1386.

Haycox,A., Unsworth,L., Allen,K., Hodgson,R., Lewis,M., & Boardman,A. (1999) "North Staffordshire Community Beds Study: longitudinal evaluation of psychiatric in-patient units attached to community mental health centres. 2: Impact upon costs and resource use", *British Journal of Psychiatry* vol. 175, pp. 79-86.

Healthcare Commission (2003) *“What CHI has found in mental health trusts: sector report”* Commission for Healthcare Improvement

Healthcare Commission (2007) *“Count me in 2007: Results of the 2007 national census of inpatients in mental health and learning disability services in England and Wales”* London: Commission for Healthcare Audit and Inspection

Hermann,R.; Leff,H.; Palmer,R. (2000) “Quality measures for mental health care: results from a national inventory” *Medical Care Research and Review* vol. 57 supplement 2 pp 136-154

Herz,M., Endicott,J., & Spitzer,R. (1975) "Brief hospitalization of patients with families: Initial results", *The American Journal of Psychiatry*, vol. 132, no. 4, pp. 413-418.

Herz,M., Endicott,J., & Spitzer,R. (1977) "Brief hospitalization: a two-year follow-up", *American Journal of Psychiatry*, vol. 134, pp. 502-507.

Higgins,R., Hurst,K., & Wistow,G. (1999), "Nursing acute psychiatric patients: A quantitative and qualitative study", *Journal of Advanced Nursing*, vol. 29, no. 1, pp. 52-63.

Higginson,I. (1994) “Quality of care and evaluating services” *International Review of Psychiatry* vol. 6 pp5-14

Hirsch,S., Platt,S., Knights,A., & Weyman,A. (1979) "Shortening hospital stay for psychiatric care: effect on patients and their families"", *British Medical Journal*, vol. 1(6161), pp. 442-446.

Hoehler,F. (2000) “Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity” *Journal of Clinical Epidemiology* vol. 53 pp499-503.

Hoult,J., Reynolds,I., Charbonneau-Powis,M., Weekes,P. and Briggs,J. (1983) “Psychiatric hospital versus community treatment: the results of a randomized trial” *Australia and New Zealand Journal of Psychiatry* vol. 17(2) pp 160–7

Howard,L.; Leese,M.; Byford,S.; Killaspy,H.; Cole,L.; Lawlor,C.; Johnson,S. (2009) “Methodological challenges in evaluating the effectiveness of women’s crisis houses compared with psychiatric wards: findings from a pilot patient preference RCT” *Journal of Nervous and Mental Disease* vol. 197(10) pp 722-727

Ianzito,B., Fine,J., Sprague,B. & Pestana,J. (1978) "Overnight admission for psychiatric emergencies", *Hospital and Community Psychiatry* vol. 29(11) pp 728-730.

Jepson,R.; DiBlasi,Z.; Wright,K. (2000) *“The Scoping Review of the Effectiveness of Mental Health Services”* York: NHS Centre for Reviews and Dissemination

Jeste,D.; Gladsjo,J.; Lindamer,L.; Lacro,J. (1996) “Medical comorbidity in Schizophrenia” *Schizophrenia Bulletin* vol. 22 pp 413-427

Johnson,S. & Salvador-Carulla,L. (1998), "Description and Classification of mental health services: a European perspective", *European Psychiatry*, vol. 13, pp. 333-341.

Johnson,S., Kuhlman,R., & The EPCAT Group (2000), "The European Service Mapping Schedule (ESMS): development of an instrument for the description and classification of mental health services", *Acta Psychiatrica Scandinavica*, vol. 102, supplement 405, pp. 14-23.

Johnson,S., Billings,J., Bingham,C., Pilling,S., Bebbington,P., Morant,N., Dalton,J., & McNicholas,S. (2001) *"An Evaluation of Drayton Park Women's Crisis Service, a community-based alternative to the psychiatric ward"*, unpublished report.

Johnson,S, BinghamC, BillingsJ, PillingS, MorantN, Bebbington,P., McNicholas,S. and Dalton,J. (2004) “Women's experiences of admission to a crisis house and to acute hospital wards: a qualitative study.” *Journal of Mental Health* vol. 13 pp 247-62

Johnson,S., Gilburt,H., Lloyd-Evans,B. and Slade,M. (2007) “Acute inpatient psychiatry: residential alternatives to hospital admission” *Psychiatric Bulletin* vol. 31 pp 262-264

Johnson,S., Gilburt,H., Lloyd-Evans,B., Boardman,J., Leese,M., Osborn,D., Shepherd,G., Thornicroft,G., Slade,M. (2009) “The Alternatives Study 1: Alternatives to inpatient care in England: emerging models and populations served” *British Journal of Psychiatry* vol.194 pp 456 – 463

Johnstone P. & Zolese G. (1999) Systematic review of the effectiveness of planned short hospital stays for mental health care. *British Medical Journal,* 318, 1387-1390.

Jones,S., Thornicroft,G., & Coffey,M. (1995) "A brief mental health outcome scale: Reliability and validity of the Global Assessment of Functioning (GAF)", *British Journal of Psychiatry*, vol. 166, pp. 654-659.

Jorgensen,K.; Romma,V.; Rundmo,T. (2009) “Associations between ward atmosphere, patient satisfaction and outcome” *Journal of Psychiatric and Mental Health Nursing* vol. 10 pp113-120

Joy,C. & Saylan,M. *"Mother and Baby Units for schizophrenia"* (2007) Cochrane Database of Systematic Reviews

Katschnig,H., Konieczna,T. and Cooper,J. (1993). *“Emergency Psychiatric and Crisis Intervention Services in Europe: A Report Based on Visits to Services in Seventeen Countries”* Geneva: World Health Organization.

Kay,S., Opler,L., & Fishbein,A. (1992) “*A Positive and Negative Syndrome Scale (PANSS) Rating Manual”* Multihealth Systems, Toronto.

Kaziz,L. (1998) “*SF-36V: The Short Form Health Survey for Veterans: Technical Report and Manual*” Department of Veterans Affairs HSR and D, Boston.

Killaspy,H., Dalton,J., McNicholas,S. and Johnson,S. (2000) “Drayton Park, an alternative to hospital admission for women in acute mental health crisis” *Psychiatric Bulletin* vol. 24 pp 101–14.

Kitwood,T. (1997), *Evaluating Dementia Care: The DCM method*, 7th edn, University of Bradford.

Kovess,V. & Lafleche,M. (1988), "How do Teams Practice Community Psychiatry?", *Canada's Mental Health*, vol. 36, no. 2, pp. 1-28.

Lafferty,S. & Davidson, R. (2006) "Putting the person first", *Mental Health Today*, March issue pp. 31-33.

Lantz,C., & Nebenzahl,E. (1996). Behavior and interpretation of the k statistic: Resolution of the two paradoxes. *Journal of Clinical Epidemiology* vol. 49 pp 431–434.

Leese,M.; Johnson,S.; Slade,M.; Parkman,S.; Kelly,F.; Phelan,M.; Thornicroft,G. (1998) “User perspectives on needs and satisfaction with mental health services: PriSM Psychosis Study 8” *British Journal of Psychiatry* vol. 173 pp 409 – 415

Leff,J.; Thornicroft,G.; Coxhead,N.; Crawford,C. (1994) “The TAPS Project 22: a five year follow-up of long-stay psychiatric patients discharged to the community” *British Journal of Psychiatry*vol. 165 Suppl. 25 pp13-17

Lloyd-Evans,B., Gilburt,H. and Johnson,S. (2008) “Integrating crisis residential care and crisis resolution teams” in Johnson,S., Needle,J., Bindman,J. and Thornicroft,G. eds *“Crisis Resolution and Home Treatment in Mental Health”* Cambridge University Press

Lorr,L., Klett,C., & McNair,D. (1963) “*Syndromes of Psychoses”* Macmillan, New York.

Machin,D.; Campbell,M.; Walters,S. (2007) *“Medical Statistics: a textbook for the health sciences”* Wiley

Marshall,M., Lockwood,A., Bradley,C., Adams,C., Joy,C., & Fenton,M. (2000) "Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia.", *British Journal of Psychiatry*, vol. 176, pp. 249-252.

Martin,J. (1984) *“Hospitals in trouble”* Oxford,Blackwell

Mathews,C., Glidden,D., Murray,S. *et al.* (2002) The effect on treatment outcomes of assigning patients to ethnically focused inpatient psychiatric units. *Psychiatric Services* vol. 53 pp 830-835

Mattes,J. (1982) "The optimal length of hospitalization for psychiatric patients: A review of the literature", *Hospital and Community Psychiatry* vol. 33(10) 824-828,

McCabe,L., Butterill,D., & Goering,P. (2004) "Residential crisis units. Are we missing out on a good idea?", *Canadian journal of community mental health; Revue canadienne de sante mentale communautaire*, vol. 23, no. 1, pp. 65-74.

McCrone,P. & Larusso,P. (1999) “A comparison of psychiatric casemix in the UK and Italy” *Social Psychiatry and Psychiatric Epidemiology* vol. 34(8) pp 432-436

McGrew,J., Bond,G., Dietzen,L., Salyers,M. (1994) "Measuring the fidelity of implementation of a mental health program model" *Journal of Consulting and Clinical Psychology* vol. 62 no. 4 pp 670-678

McGuire-Snieckus,R.; McCabe,R.; Catty,J.; Hansson,L.; Priebe,S. (2007) “A new scale to assess the therapeutic relationship in community mental health care” Psychological Medicine (37) pp85-95

McHugo,G., Drake,R., Teague,G., Xie,H. (1999) "Fidelity to Assertive Community Treatment and Client Outcomes in the New Hampshire dual disorders study" Psychiatric Services vol.50 no. 6 pp 818-824

McHugo,G.; Drake,R.; Whitley,R.; Bond,G.; Campbell,K.; Rapp,C.; Goldman,H.; Lutz,W.; Finelly,M. (2007) “Fidelity outcomes in the national Implementing Evidence Based Practices Project” *Psychiatric Services* vol. 58(10) pp 1279-1284

McLellan,A., Kushner,H., & Metzger,D. (1992) "The fifth edition of the Addiction Severity Index", *Journal of Substance Abuse Treatment*, vol. 9, pp. 199-213.

Mechanic,D. (1996) "Emerging issues in international mental health services research", *Psychiatric Services*, vol. 47, no. 4, pp. 371-375.

Mendel,W. (1966) "Effect of length of hospitalization on rate and quality of remission from acute psychiatric episodes" *Journal of Nervous and Mental Disease*, vol. 143, pp. 226-233.

Mental Health Act Commission (2008) *Risk, Rights and Recovery. Twelfth Biennial Report 2005-2007.* London: The Stationery Office

Mezzina,R. and Vidoni,D. (1995) “Beyond the mental hospital: crisis intervention and continuity of care in Trieste. A four year follow-up study in a community mental health centre.” *International Journal of Social Psychiatry* vol. 41, pp 1–20.

Moher,D., Cook,D., Eastwood,S., Olkin,J., Drummond,R., & Stroup,D. (1999) "Improving the quality of reports of meta analyses of randomised controlled trials: the QUOROM statement" *The Lancet*, vol. 354, pp. 1896-1900.

Moos,R. (1996) “*Ward Atmosphere Scale Manual*” 3rd edn, Mind Garden.

Morgan,C.; Mallett,R.; Hutchinson,G.; Bagalkote,H.; Morgan,K.; Fearon,P.; Dazzan,P.; Boydell,J.; McKenzie,K.; Harrison,G.; Murray,R.; Jones,P.; Craig,T.; Leff,J. (2005) “Pathways to care and ethnicity:1 Sample characteristics and compulsory admission: Report from the Aesop Study” *British Journal of Psychiatry* vol. 186 pp 281-289

Morley, S. & Snaith, P. (1992), "Principles of Psychological Assessment," in *Research Methods in Psychiatry*, 2nd edn, C. Freeman & P. Tyrer, eds., Gaskell, pp. 135-152.

Morse,G., Calsyn,R., Klinkenberg,W., Helminiak,T., Wolff,N., Drake,R., Yonker,R., Lama,G., Lemming,M., McCudden,S. (2006) "Treating homeless clients with severe mental illness and substance use disorders: costs and outcomes" *Community Mental Health Journal* vol.42 no.4 pp 377-404

Mosher,L.; Pollin,W.; Stabenau,J. (1971) "Identical twins discordant for schizophrenia: neurological findings“ *Archives of General Psychiatry* vol. 24(5) pp 422-430

Mosher,L., Menn,A. and Matthew,S. (1975). Soteria: evaluation of a home-based treatment for schizophrenia. *American Journal of Orthopsychiatry* vol. 45, pp 455–67.

Mosher,L. & Menn,A. (1978) "Community residential treatment for schizophrenia: two-year follow-up", *Hospital and Community Psychiatry*, vol. 29, no. 11, pp. 715-723.

Mosher,L., Vallone,R., & Menn,A. (1995) "The treatment of acute psychosis without neuroleptics: six-week psychopathology outcome data from The Soteria Project", *International Journal of Social Psychiatry* vol. 41, no. 3, pp. 157-173.

Mosher,L. (1999). “Soteria and other alternatives to acute psychiatric hospitalization: a personal and professional review” *Journal of Nervous and Mental Disease* vol. 187 pp 142–9.

Mowbray CT, Holter MC, Teague GB, Bybee D. (2003) “Fidelity criteria: Development, measurement, and validation” *American Journal of Evaluation* vol.24(3) pp 315-40.

Mueser,K.; Torrey,W.; Lynde,D.; Singer,P.; Drake,R. (2003) “Implementing evidence-based practices for people with severe mental illness” *Behaviour Modification* vol. 27 pp 387-411

Muijen,M. (1999) Acute hospital care: ineffective, inefficient and poorly organised. *Psychiatric Bulletin* vol. 23 pp 257-259

National Colaborating Centre for Mental Health (2009) *“Schizophrenia: core interventions in the treatment and management of schizophrenia in primary and secondary care (update): national clinical practice guideline Number 82”* National Institute for Health and Clinical Excellence

Office of National Statistics (2001) *“National Census – England Household Form”* www.census.ac.uk/CensusForms/2001_England_Household

Olfson,M., Plotke,G., Harding,C., Jones,J., & Mayfield,W. (1990) "A controlled evaluation of inpatient crisis treatment for acute schizophrenia episodes", *Psychiatric Quarterly*, vol. 61, pp. 143-154.

Oliver,J., Huxley,S., & Priebe,S. (1997) "Measuring the quality of life of severely mentally ill people using the Lancashire Quality of Life Profile", *Social Psychiatry and Psychiatric Epidemiology*, vol. 32, pp. 76-83.

Olusina,A., Ohaeri,J., & Olatawura,M. (2003), "The quality of interactions between staff and psychiatric inpatients in a Nigerian general hospital", *Psychopathology*, vol. 36, no. 5, pp. 269-275.

Oppenheim,A. (1992) *"Questionnaire Design, Interviewing and Attitude measurement"*, 2 edn, Pinter.

Overall,K. & Gorham,D. (1962) "The Brief Psychiatric Rating Scale", *Psychological Reports*, vol. 10, pp. 799-812.

Patmore,C. & Weaver,T. (1989), "A measure of care", *Health Services Journal* vol. 99, no. 5142 pp. 330-331.

Patrick,M.; Higgitt,A.; Holloway,F.; Silverman,M. (1989) “Changes in an inner city psychiatric inpatient service following bed losses: a follow up of the East Lambeth 1986 survey” *Health Trends* vol.21 pp 121-12

Paul,G. (1987), "Rational operations in residential treatment settings through ongoing assessment of staff and client functioning," in “*Assessment for Decision*” D. Peterson & D. Fishman, eds., Rutgers University Press, pp. 145-203.

Pawson,R. & Tilley,N. (1997) *"Realistic Evaluation"* Sage

Peele,R. & Chodoff,P. (1999) *“The ethics of involuntary treatment and deinstituionalisation”* pp 423-440 in Bloch,S.; Chodoff,P.; Green,s. eds *“Psychiatric Ethics 3rd edition”* Oxford, OUP

Perneger,T. (1998) “What’s wrong with Bonferroni adjustments?” *British medical Journal* vol. 316 pp 1236-8

Phelan,M.; Stradris,L.; Amin,D.; Isadore,R.; Hitrou,C.; Doyle,A.; Inglis,R. (2004) “The Physical Health Check: a tool for mental health workers” *Journal of Mental Health* vol. 13 pp 277 - 284

Polak,P. & Kirby,M. (1976) "A model to replace psychiatric hospitals", *Journal of Nervous and Mental Disease* vol. 162, no. 1, pp. 13-22

Polak,P., Kirby,M. and Deitchman,W. (1979). “Treating acutely psychotic patients in private homes” *New Directions for Mental Health Services* vol 1 pp 49–64

Popkin,M., Callies,A., Lurie,N., Harman,J., Stoner,T., Manning,W. (1997) "An instrument to evaluate the process of care in ambulatory settings" *Psychiatric Services* vol.48 no.4 pp 524-527

Popper,K. (1963) *“Conjectures and refutations: the growth of scientific knowledge”* Routledge, London

Priebe,S.; Badesconyi,A.; Fioritti,A. (2005) “Reinstituionalisation in mental health care: a comparison of data on service provision from six European countries” *British Medical Journal* vol. 350 pp 123-126

Quirk,A. & Lelliot,P. (2001) " What do we know about life on acute psychiatric wards in the UK? A review of the research evidence." *Social Science and Medicine* vol.53(12) pp 1565-1574

Quirk,A.; Lelliott,P.; Seale,C. (2006) ”The permeable institution: an ethnographic study of three acute psychiatric wards in London” *Social Science and Medicine* vol. 63 pp 2105-2117

Ramsay,R. & Holloway,F. (1998) *“Mental Health Services”* pp 1274-1333 in Stein,G. & Wilkinson,G. *“General psychiatry: college seminars in psychiatry”* London, Gaskell

Rappaport,M., Goldman,H., Thornton,P., Stegner,B., Moltzen,S., Hall,K., Gurevitz,H., & Attkisson,C. (1987) "A method or comparing two systems of acute 24-hour psychiatric care" *Hospital and Community Psychiatry*, vol. 38, no. 10, pp. 1091-1095.

Readhead,C., Henderson,R., Hughes,G., & Nickless,J. (2002) "Accredited accommodation: An alternative to in-patient care in rural north Powys", *Psychiatric Bulletin*, vol. 26, no. 7, pp 264-265.

Resnick,S., Neale,M., Rosenheck,R. (2003) "Impact of Public Support Payments, intensive pychiatric community care and program fidelity on employment outcomes for people with severe mental illness" *The Journal of Nervous and Mental Disease* vol.191 no. 3 pp 139-144

Rorshach,H. (1921) "*Psychodiagnostics"* Grune and Straton, New York.

Rose,D. (2001*) "Users' Voices, the perspectives of mental health service users on community and hospital care".* London: The Sainsbury Centre for Mental Health.

Rossberg,J. & Friis,S. (2004), "Patients' and staff's perceptions of the psychiatric ward environment", *Psychiatric Services*, vol. 55, no. 7, pp. 798-803.

Rossi,P., Freeman,H., & Lipsey,G. (1999) "*Evaluation: a Systematic Approach*" 6 edn, Sage.

Royal College of Psychiatrists (1992) "Report of the General Psychiatry Section Working Party on Post-Natal Mental Illness" *Psychiatric Bulletin* vol. 16 pp 519-22.

Royal College of Psychiatrists (2001) "Perinatal Maternal Mental Health Services: Recommendations for Provision of Services for Childbearing Women" *Psychiatric Bulletin* vol. 25 p.278.

Royal College of Psychiatrists (2007) *"Accreditation for acute inpatient mental health services: AIMS"* Royal College of Psychiatrists

Ruggeri,M., Dall'Agnola,R., & Agostini,C. (1994) "Acceptability, sensitivity and content validity of the VECS and VSSS in measuring expectations and satisfaction in psychiatric patients and their relatives", *Social Psychiatry and Psychiatric Epidemiology*, vol. 29, pp. 265-276.

Sainsbury Centre for Mental Health (1998) *"Acute problems: a survey of the quality of care in acute psychiatric wards"* SCMH, London

Sainsbury Centre for Mental Health (2006) *"Search for acute solutions: improving the quality of care in acute psychiatric wards"* SCMH, London

Salize,H., Kustner,B., Torres-Gonzalez,F., Reinhard,I., Estevez,J. & Rossler,W. (1999), "Needs for care and effectiveness of mental health care provision for schizophrenic patients in two European regions: a comparison between Granada (Spain) and Mannheim (Germany)", *Acta Psychiatrica Scandinavica*, vol. 100, no. 5, pp. 328-334.

Schmidt,L.; Garratt,A.; FitzPatrick,R. (2000) *"Instruments for mental health: a review"* Department of Health, London

Schneider,S. & Ross,I. (1996) "Ultra-short hospitalization for severely mentally ill patients", *Psychiatric Services* vol. 47(2) pp137-138.

Schulz,K.; Chalmers,I.; Hayes,R.; Altman,D. (1995) "Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials" *Journal of the American Medical Association* vol. 273(5) pp 408-412

Shepherd,G. & Richardson,R. (1979), "Organization and Interaction in Psychiatric Day Centres", *Psychological Medicine*, vol. 9, pp. 573-579.

Slade,M., Phelan,M., & Thornicroft,G. (1996) "The Camberwell Assessment of Need (CAN): comparison of assessments by staff and patients of the needs of the mentally ill", *Social Psychiatry and Psychiatric Epidemiology*, vol. 31, pp. 109-113.

Slade,M., Phelan,M., & Thornicroft,G. (1998), "A comparison of needs assessed by staff and by and epidemiologically representative sample of patients with psychosis", *Psychological Medicine*, vol. 28, no. 3, pp. 543-550.

Slade,M.; Thornicroft,G.; Loftus,S.; Phelan,M.; Wykes,T. (1999) *"The Camberwell Assessment of Need"* Gaskell, London

Slade,M. (2002) "Routine outcome measurement in mental health services" *Psychological Medicine* vol. 32 pp 1339-1343

Soergaard,K.; Nivison,M.; Hansen,V.; Oieesvold,T (2008) "Treatment needs and acknowledgement of illness – importance for satisfaction with psychiatric inpatient treatment" *BMC Health Services Research* 8:103

Spitzer, R. L., Endicott, J., & Fleiss, J. L. (1970) "The Psychiatric Status Schedule: a technique for evaluating psychopathology and impairment in role functioning", *Archives of General Psychiatry*, vol. 23, pp 41-55

Spitzer,R. and Endicott,J. (1971) "An integrated gtoup of forms for automated psychiatric case records: progress report" *Archives of General Psychiatry* vol 24. pp 540-547

SPSS (2006) *"Statistical Package for Social Sciences version 14.0"* Chicago IL, SPSS Inc

StataCorp (2007) *"STATA Statistical Software: Release 10"* College Station TX, StataCorp LP

Stein, L. (1991) "A systems approach to the treatment of people with chronic mental illness" In *"The Closure of Mental Hospitals"* ed. P. Hall and I. F. Brockington. London: Gaskell, pp. 91–106.

Stevenson, C., Barker, P., & Fletcher, E. (2002) "Judgement days: developing an evaluation for an innovative nursing model", *Journal of Psychiatric and Mental Health Nursing*, vol. 9, no. 3, pp. 271-276.

Stroul,B. (1988) "Residential crisis services: a review", *Hospital and Community Psychiatry*, vol. 39, no. 10, pp. 1095-1099.

Szmukler,G. & Holloway,F. (2001) *“Inpatient Treatment”* pp 321-333 in Thornicroft,G. and Szmukler,G. eds *“Textbook of community Psychiatry”* Oxford, OUP

Tansella,M. & Thornicroft,G. (2001) *“Mental Health Outcome Measures” 2nd edition*, Gaskell, London

Teague,G., Bond,G., & Drake,R. (1998), ""Programme fidelity in ACT: Development and Use of a Measure"", *American Journal of Orthopsychiatry*, vol. 68, pp. 216-232.

Thomas,H. (2003) *“Quality assessment tool for quantitative studies”* Effective Public Health Practice Project, McMaster University, Toronto

Thornicroft,G. & Bebbington,P. (1989) “Deinstituionalisation – from hospital closure to service development” *British Journal of Psychiatry* vol. 155 pp 739-753

Thornicroft,G. & Tansella,M. (1999), “*The Mental Health Matrix: a Manual to Improve Services”* Cambridge University Press.

Thornton,A., Hatton,A., & Tatham,A. (2004), "Dementia Care Mapping Reconsidered: Exploring the Reliability and Validity of the Observational Tool", *International Journal of Geriatric Psychiatry*, vol. 19, pp. 718-726.

Tilden,C.; Nelson,C.; May,B. (1990) ”The IPR Inventory; Development and Psychometric Characteristics” Nursing Research 39(6) pp 337-343

Timko,C., Chen,S., Sempel,J., & Barnett,P. (2006) "Dual diagnosis patients in community or hospital care: One-year outcomes and health care utilization and costs.", *Journal-of-Mental-Health.* vol. 15(2) pp 163-177

Tugwell,P. (1979), "A Methodological Perspective on Process Measures of the Quality of Medical Care", *Clinical and Investigative Medicine* vol. 2, pp. 113-121.

Tyson,G., Lambert,G., & Beattie,L. (1995), "The Quality of Psychiatric Nurses' Interactions with Patients: an Observational Study", *International Journal of Nursing Studies*, vol. 32, no. 1, pp. 49-58.

Vallis,T., Shaw,B., Dobson,K. (1986) “The Cognitive Therapy Scale: psychometric properties” *Journal of Consulting Clinical Psychology* vol. 54(3) pp 381-5

Venables,P. & O'Connor,N. (1959) "A short scale for rating paranoid schizophrenia", *Journal of Mental Science*, vol. 105, pp. 815-818.

Virtanen,M. (2003) “Phil Barker’s Tidal Model: does a chaos theory help in psychiatric nursing?” *Sairaanhoitaja* 76(8) pp 20-23

Voineskos,G., Morrison,M. & Jain,R. (1974) "The introduction of a crisis unit in a mental hospital", *Canadian Psychiatric Association Journal* vol. 19(5) pp445-452.

Wall,S.; Hotopf,M.; Wessely,S.; Churchill,R.; (1999) “Trends in the use of the Mental Health Act: England 1984-96” *British Medical Journal* vol. 316 pp 1520-1521

Ware,J., Snow,K., & Kosinski,M. (1993) “*SF-36 Health Survey Manual and Interpretation Guide”* New England Medical center Health Institute, Boston.

Welsh Assembly Government (2005) *“Raising the standard: the revised mental health adult National Service Framework and an action plan for Wales”* Welsh National Assembly

Wesson,M. and Walmsley,P. (2001). Service innovations: Sherbrook partial hospitalisation unit. *Psychiatric Bulletin* vol. 25 pp 56–8.

Whitley,R. & Crawford,M. (2005) "Qualitative Research in Psychiatry" *Canadian Journal of Psychiatry* vol. 50 (2) pp 108-114

Wing,J. & Brown,G. (1970), *“Institutionalism and Schizophrenia”* Cambridge University Press.

Wing,J., Cooper,J., & Sartorius,N. (1974) “*The measurement and classification of psychiatric symptoms: An instrument manual for the Present State examination and Catego Programme”* Cambridge University Press, Cambridge.

Wing,J., Curtis,R., & Beevor,A. (1996) “*HoNOS:Health of the Nation Outcome Scale: Report on Research and Development*” Royal College of Psychiatrists, London.

Worden,A. Challis,D.; Hancock,G.; Woods,R.; Orrell,M. (2008) “Identifying need in care homes for people with dementia: the relationship between two standard assessment tools” *Aging and Mental Health* vol. 12(6) pp 719-728

Wright,R., Sklebar,T., & Heiman,J. (1987), "Patterns of case management activity in an intensive community support programme", *Community Mental Health Journal*, vol. 23, no. 1, pp. 53-59.

Wykes,T. & Sturt,E. (1986) "The measurement of social behaviour in psychiatric patients: an assessment of the reliability and validity of the SBS schedule", *British Journal of Psychiatry*, vol. 148, pp. 1-11.

Young,A., Sullivan,G., Burnam,A., & Brook,R. (1998), "Measuring the quality of outpatient treatment for schizophrenia", *Archives of General Psychiatry*, vol. 55, no. 7, pp. 611-617.

Appendix 1: Outcome measures from Table 3.3

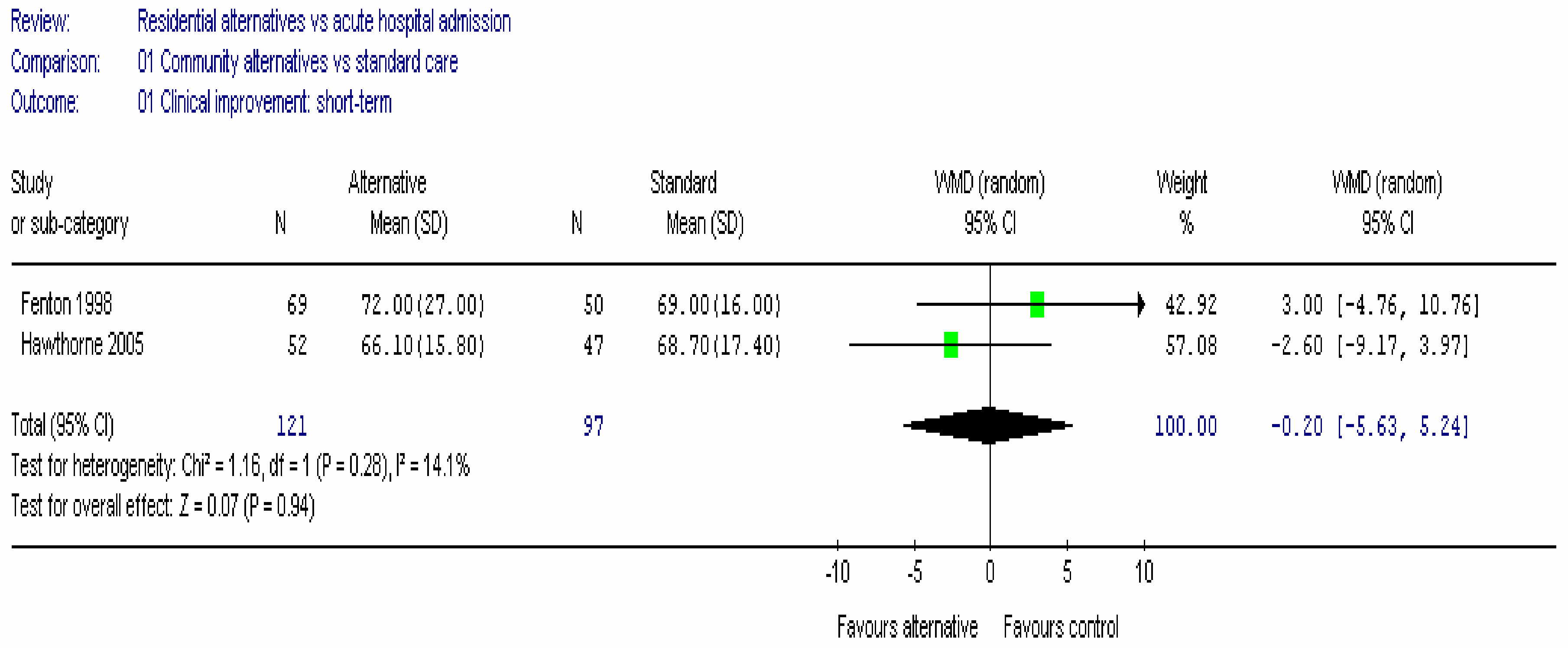
Table A1.1: References to outcome measures from Table 3.3

Acronym	Title	Reference
ASI	Addiction Severity Index	McLellan et al. (1992)
BASIS-32	Behaviour and Symptom Identification Scale	Eisen et al. (1994)
BPRS	Brief Psychiatric Rating Scale	Overall and Gorham (1962)
CAN	Camberwell Assessment of Need	Slade et al. (1996)
CSQ	Client Satisfaction Questionnaire	Attkisson and Zwick (1982)
GAF	Global Assessment of Functioning	Jones et al. (1995)
GAS	Global Assessment Scale	Endicott (1976)
-	Goal Attainment System	Ellis and Wilson (1973)
HoNOS	Health of the Nation Outcome Scale	Wing et al. (1996)
HRSD	Hamilton Rating Scale for Depression	Hamilton (1960)
IMPS	Inpatient Multi-dimensional Scale for Rating Psychotic Patients	Lorr et al. (1963)
LQLP	Lancashire Quality of Life Profile	Oliver et al. (1997)
MMPI	Minnesota Multiphasic Personality Inventory	Gilliland and Colgin (1951)
MSER	Mental State Examination Record	Spitzer and Endicott (1971)
PANSS	Positive and Negative Symptom Scale	Kay et al. (1992)
PBAS	Patients’ Behaviour Assessment Scale	Hirsch (1979)
PEF	Psychiatric Evaluation Form	Endicott and Spitzer (1972)
POC	Perceptions of Care Questionnaire	Eisen et al. (2002)
PSE	Present State Examination	Wing et al. (1974)
PSS	Psychiatric Status Schedule	Spitzer et al. (1970)
-	Rorschach Test	Rorschach (1921)
SBS	Social Behaviour Schedule	Wykes and Sturt (1986)
SF-36	Health Survey – Short Form	Ware et al. (1993)
SF-36(V)	Health Survey – Short Form (Veterans Version)	Kaziz (1998)
TES	Treatment Effectiveness Scale	Bebeau (1971)
-	Untitled short scale for rating paranoia	Venables and O’Connor (1959)
VSSS	Verona Service Satisfaction Scale	Ruggeri et al. (1994)

Appendix 2 - Meta-Analysis Data from Chapter 3

Comparison 1: Community Alternatives vs Standard Acute Inpatient Care

Figure A2.1 Clinical improvement (symptoms) - short-term



Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 01 Clinical improvement: short-term

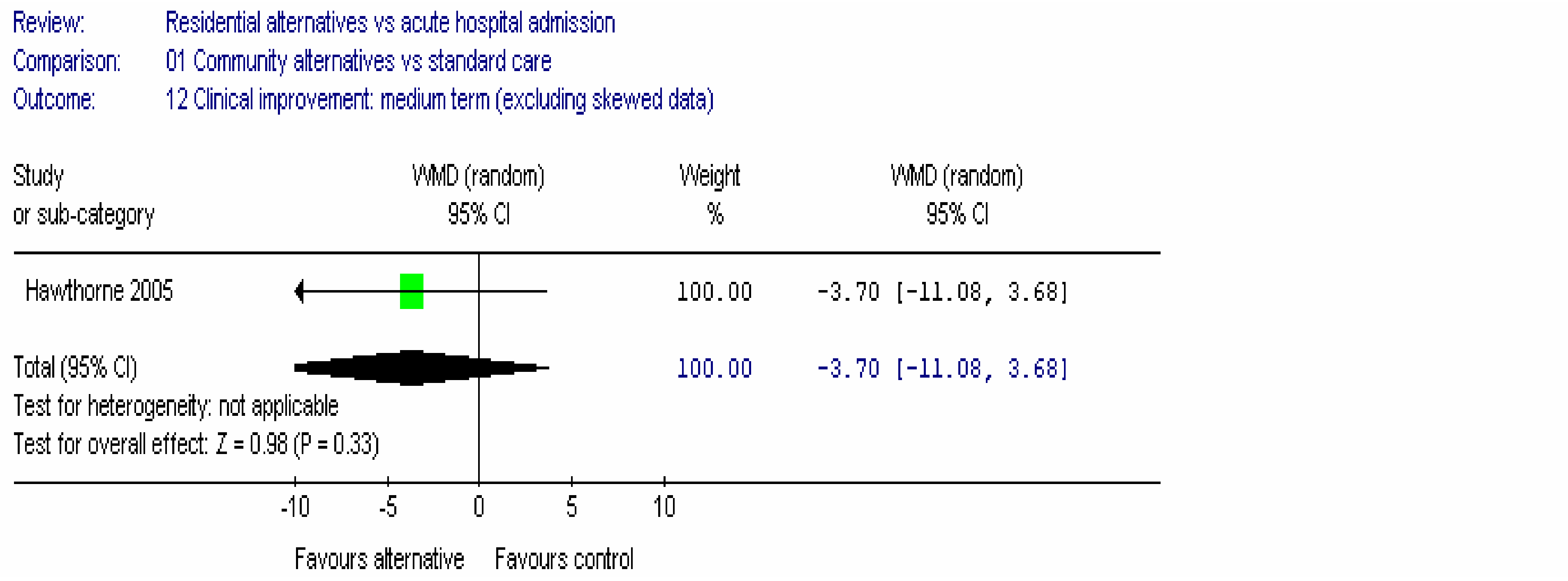
Study		Alternative		Standard	WMD (random)	Weight	WMD (random)
or sub-category	N	Mean (SD)	N	Mean (SD)	95% CI	%	95% CI
Fenton 1998	69	72.00(27.00)	50	69.00(16.00)		42.92	3.00 [-4.76, 10.76]
Hawthorne 2005	52	66.10(15.80)	47	68.70(17.40)		57.08	-2.60 [-9.17, 3.97]
Total (95% CI)	121		97			100.00	-0.20 [-5.63, 5.24]

Test for heterogeneity: Chi^2 = 1.16, df = 1 (P = 0.28), I^2 = 14.1%
Test for overall effect: Z = 0.07 (P = 0.94)

Table A2.2 Clinical improvement (symptoms) - medium term (some data skewed)

Study	Alternative			Standard service		
	N	mean	s.d.	N	mean	s.d.
Hawthorne 2005 (PANSS score)	40	63.60	18.80	40	67.30	14.60
Timko 2006 (ASI Psychiatric sub-scale score)	57	0.36	0.28	173	0.39	0.27

Figure A2.2 Clinical improvement (symptoms) - medium term (excluding skewed data)



Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 12 Clinical improvement: medium term (excluding skewed data)

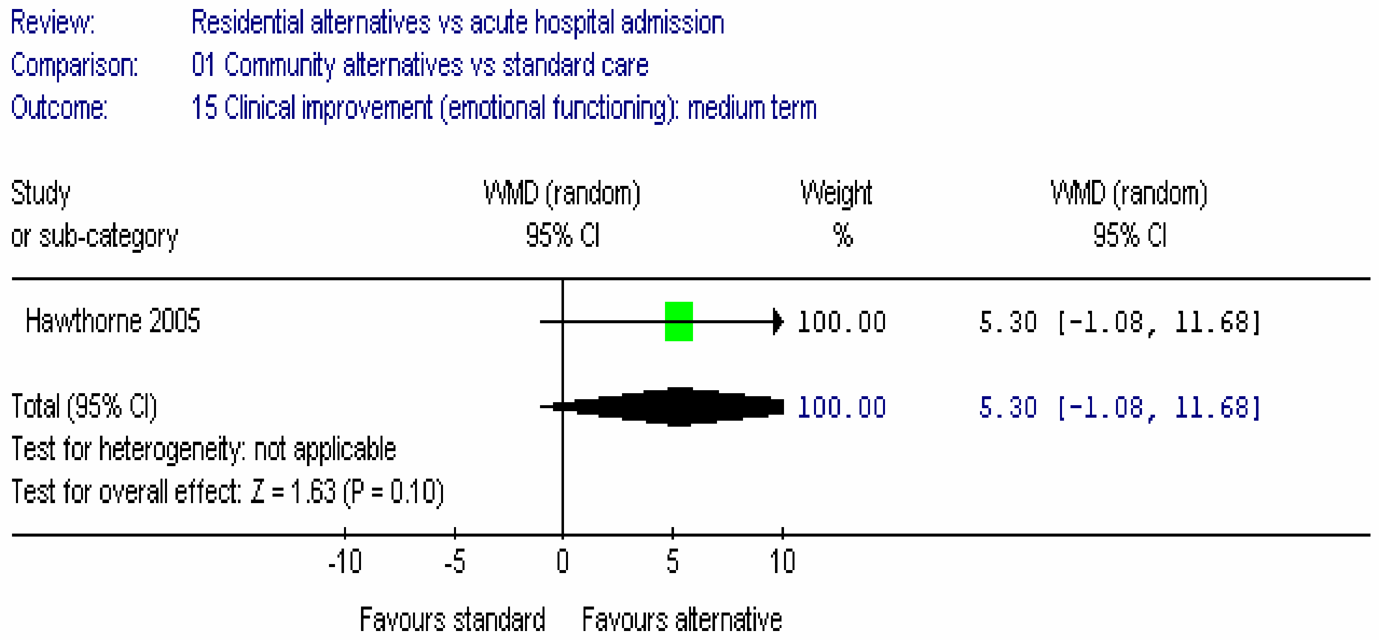
Study or sub-category	Weight %	WMD (random) 95% CI
Hawthorne 2005	100.00	-3.70 [-11.08, 3.68]
Total (95% CI)	100.00	-3.70 [-11.08, 3.68]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.98 (P = 0.33)

Table A2.3 Clinical improvement (emotional functioning): short term
No meta-analysis: data skewed

Study	Alternative			Standard service		
	N	mean	s.d.	N	mean	s.d.
Hawthorne 2005	52	25.20	11.4	47	22.40	13.3

Figure A2.3 Clinical improvement (emotional functioning): medium term

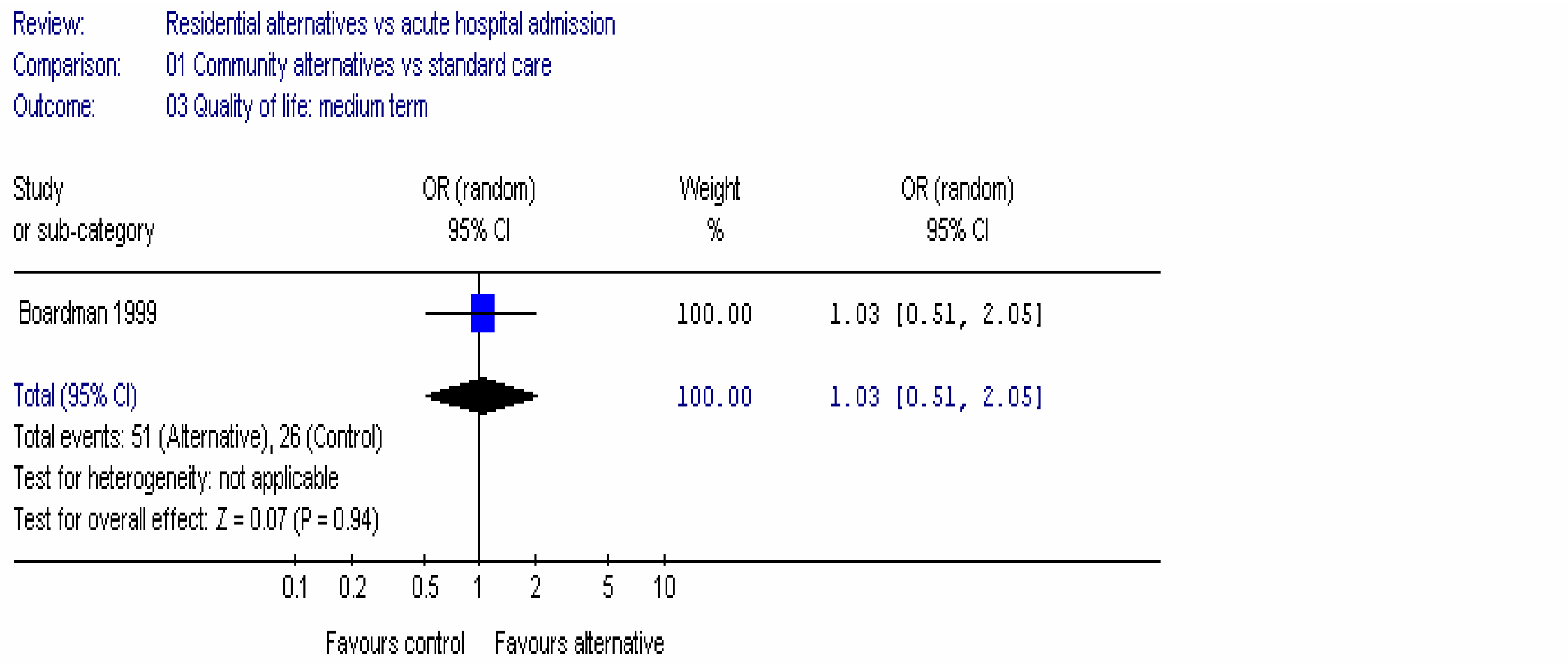


Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 15 Clinical improvement (emotional functioning): medium term

Study or sub-category	Weight %	WMD (random) 95% CI
Hawthorne 2005	100.00	5.30 [-1.08, 11.68]
Total (95% CI)	100.00	5.30 [-1.08, 11.68]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.63 (P = 0.10)

Figure A2.4 Quality of life - medium term

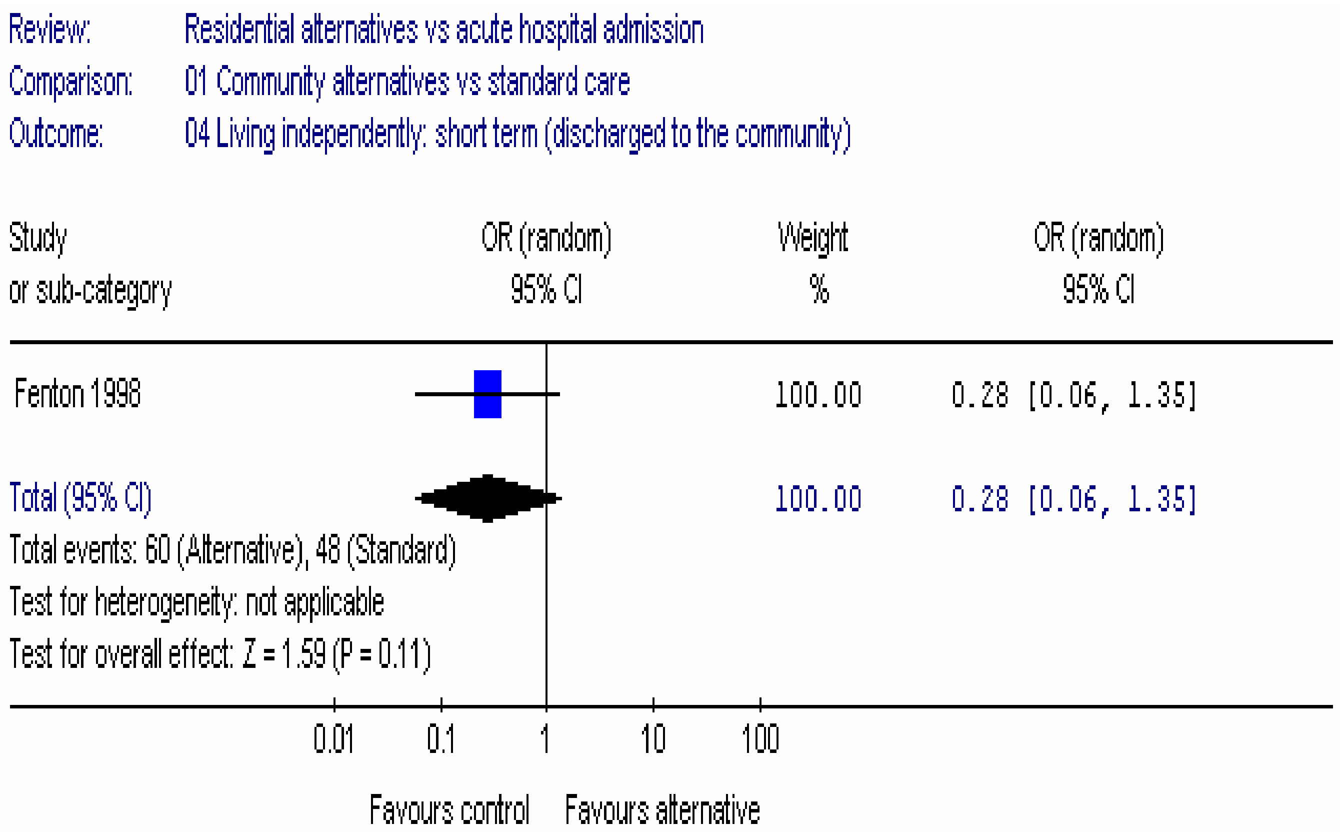


Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 03 Quality of life: medium term

Study or sub-category	OR (random) 95% CI	Weight %	OR (random) 95% CI
Boardman 1999		100.00	1.03 [0.51, 2.05]
Total (95% CI)		100.00	1.03 [0.51, 2.05]

Total events: 51 (Alternative), 26 (Control)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.07 (P = 0.94)

Figure A2.5 Living independently – short term (discharged to the community)

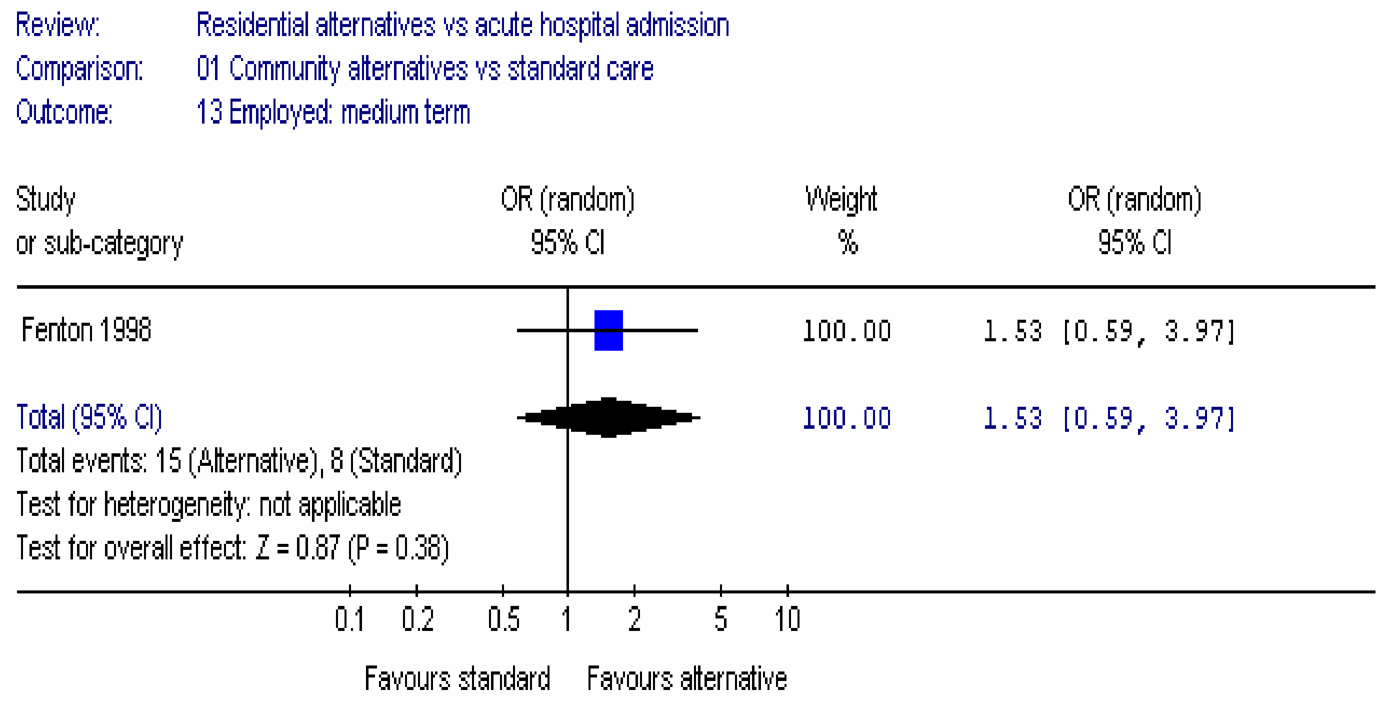


Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 04 Living independently: short term (discharged to the community)

Study or sub-category	OR (random) 95% CI	Weight %	OR (random) 95% CI
Fenton 1998		100.00	0.28 [0.06, 1.35]
Total (95% CI)		100.00	0.28 [0.06, 1.35]

Total events: 60 (Alternative), 48 (Standard)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.59 (P = 0.11)

Figure A2.6 Employment: Medium term



Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 13 Employed: medium term

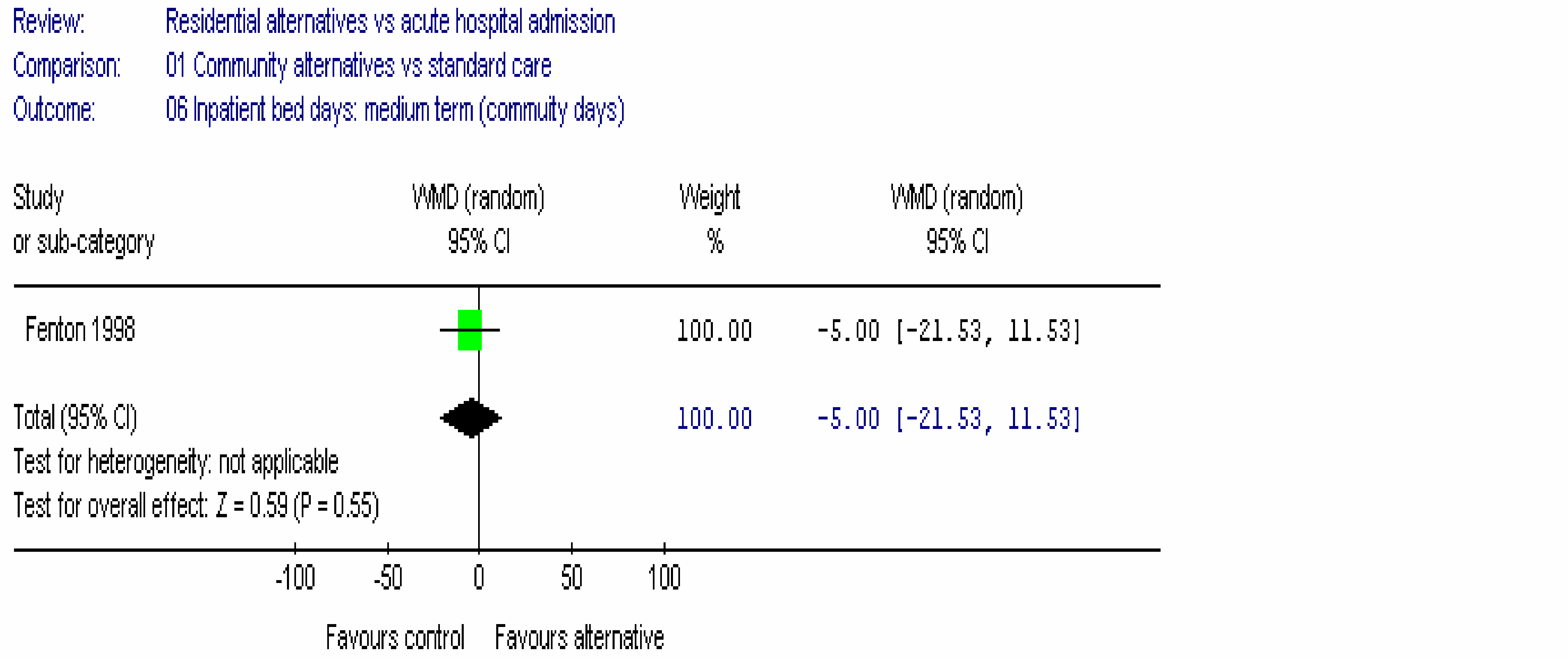
Study or sub-category	OR (random) 95% CI	Weight %	OR (random) 95% CI
Fenton 1998		100.00	1.53 [0.59, 3.97]
Total (95% CI)		100.00	1.53 [0.59, 3.97]

Total events: 15 (Alternative), 8 (Standard)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.87 (P = 0.38)

Table A2.4 Inpatient bed days – short term (no meta analysis: all data skewed)

Study	Alternative			Standard service		
	N	mean	s.d.	N	mean	s.d.
Fenton 1998	69	18.70	13.80	50	11.70	8.20
Hawthorne 2005	52	12.90	6.30	47	10.60	6.80
Timko 2006	57	55.39	51.30	173	26.01	21.35

Figure A2.7 Inpatient bed days – medium term (community days)

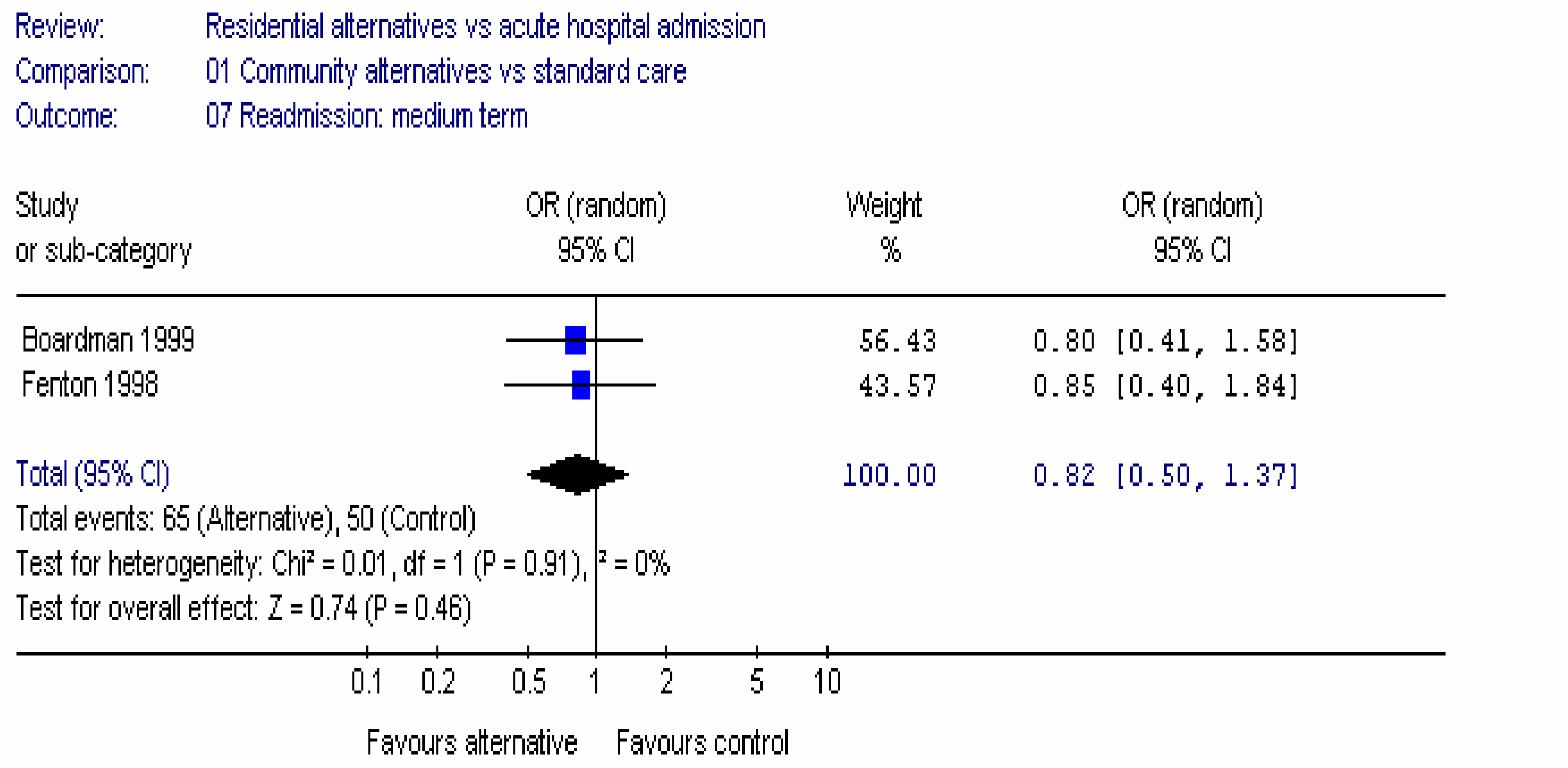


Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 06 Inpatient bed days: medium term (commuity days)

Study or sub-category	WMD (random) 95% CI	Weight %	WMD (random) 95% CI
Fenton 1998		100.00	-5.00 [-21.53, 11.53]
Total (95% CI)		100.00	-5.00 [-21.53, 11.53]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.59 (P = 0.55)

Figure A2.8 Readmission – medium term

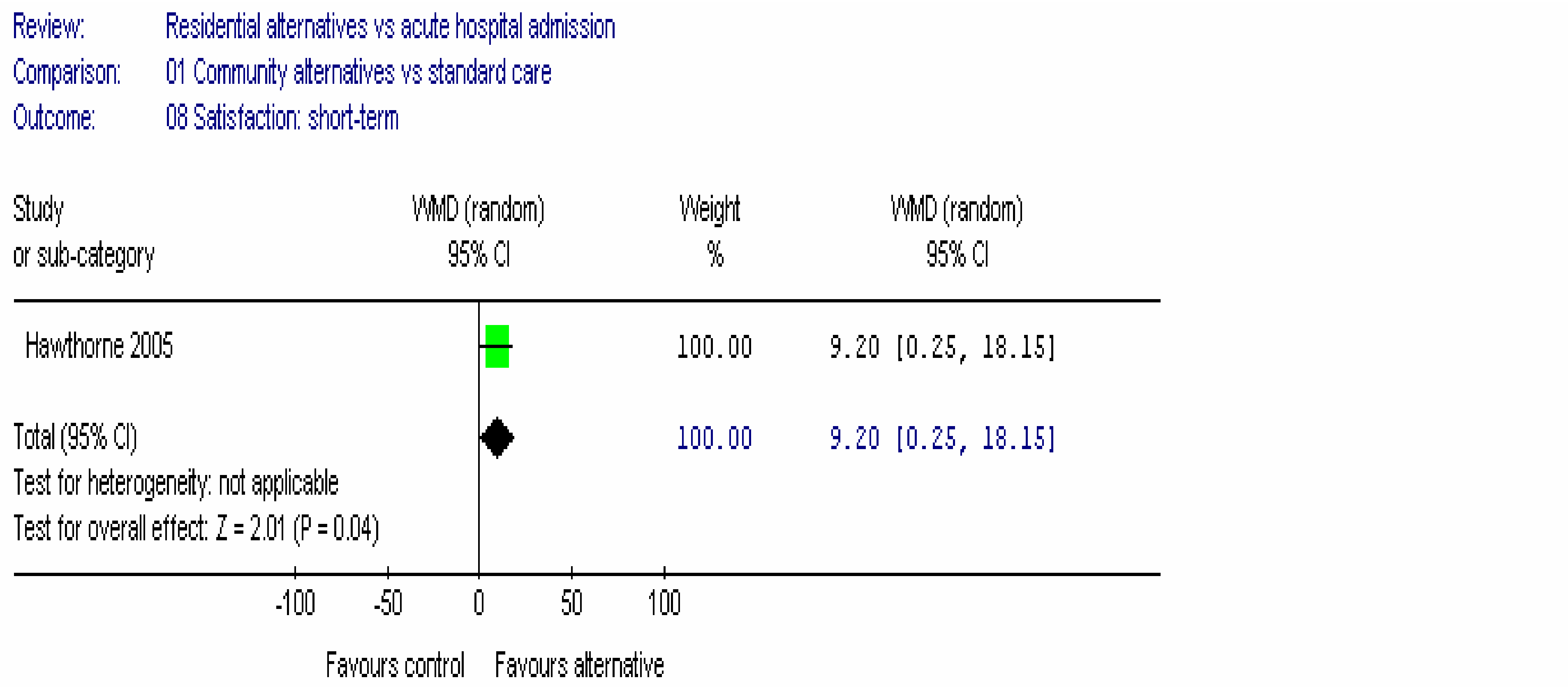


Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 07 Readmission: medium term

Study or sub-category	OR (random) 95% CI	Weight %	OR (random) 95% CI
Boardman 1999		56.43	0.80 [0.41, 1.58]
Fenton 1998		43.57	0.85 [0.40, 1.84]
Total (95% CI)		100.00	0.82 [0.50, 1.37]

Total events: 65 (Alternative), 50 (Control)
Test for heterogeneity: Chi² = 0.01, df = 1 (P = 0.91), I² = 0%
Test for overall effect: Z = 0.74 (P = 0.46)

Figure A2.9 Satisfaction – short term

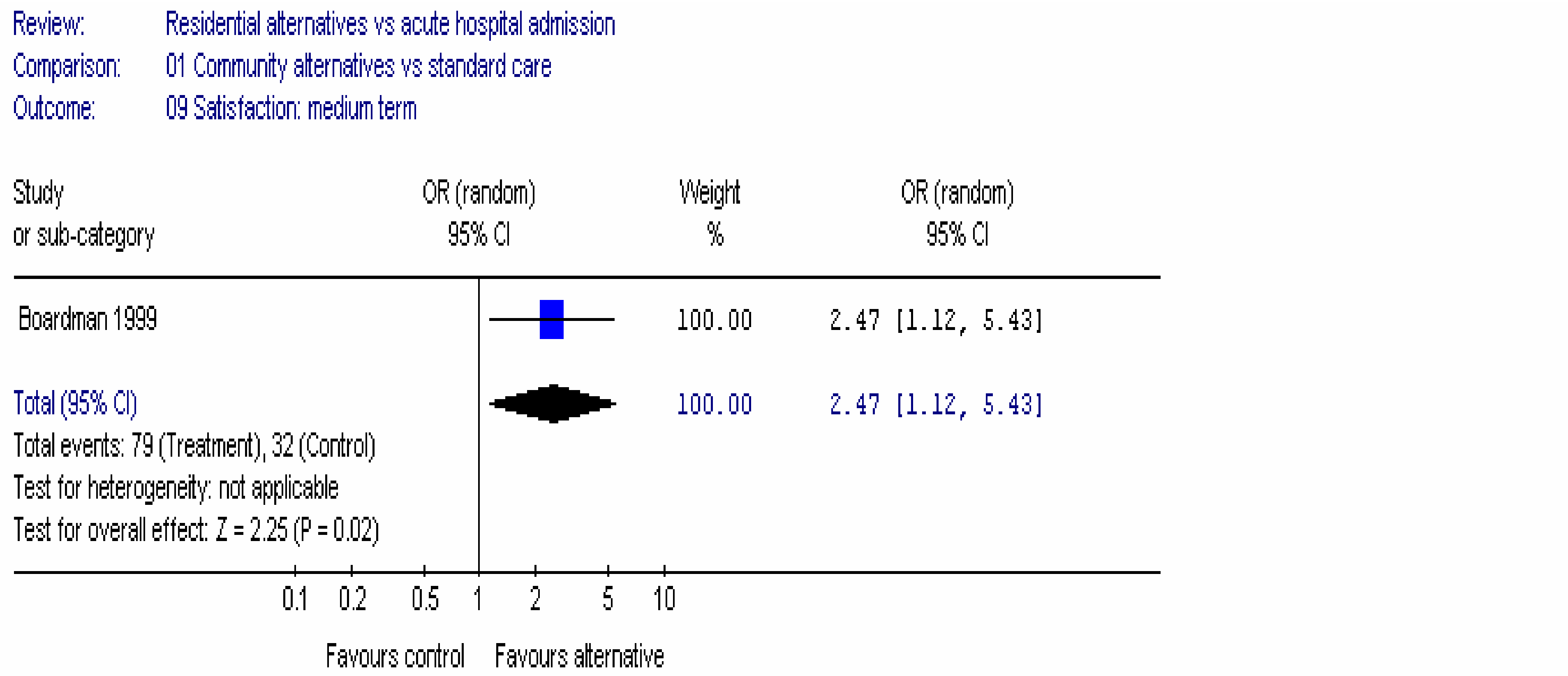


Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 08 Satisfaction: short-term

Study or sub-category	Weight %	WMD (random) 95% CI
Hawthorne 2005	100.00	9.20 [0.25, 18.15]
Total (95% CI)	100.00	9.20 [0.25, 18.15]

Test for heterogeneity: not applicable
Test for overall effect: Z = 2.01 (P = 0.04)

Figure A2.10 Satisfaction – medium term



Review: Residential alternatives vs acute hospital admission
Comparison: 01 Community alternatives vs standard care
Outcome: 09 Satisfaction: medium term

Study or sub-category	Weight %	OR (random) 95% CI
Boardman 1999	100.00	2.47 [1.12, 5.43]
Total (95% CI)	100.00	2.47 [1.12, 5.43]

Total events: 79 (Treatment), 32 (Control)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.25 (P = 0.02)

Table A2.5 Costs – short term (no meta analysis: all data skewed)
(Cost in $ of index admission)

Study	Alternative			Standard service		
	N	mean	s.d.	N	mean	s.d.
Fenton 1998	69	3046.00	2124.00	50	5549.00	3668.00
Hawthorne 2005	52	3234.00	1587.00	47	9136.00	5912.00

Table A2.6 Costs – medium term (no meta analysis: all data skewed)

Study	Alternative			Standard service		
	N	mean	s.d.	N	mean	s.d.
Fenton 1998 (service costs in $ over 6 month follow-up)	69	19941.00	19282.00	50	25737.00	21835.00
Timko 2006 (service costs in $ over 12 month follow-up)	57	21996.00	17559.00	173	33188.00	25473.00

Comparison 2: Time-Limited Alternatives vs Standard Acute Inpatient Care

Table A2.7 Time-limited alternatives -Inpatient bed-days: medium term (no meta analysis: data skewed)

Study	Alternative			Standard service		
	N	mean	s.d.	N	mean	s.d.
Olfson 1990 (inpatient bed days over 3 month follow-up)	8	29.13	14.72	18	30.61	17.92

APPENDIX 3

CaSPAR

Camden Staff-Patient Activity Record

CaSPAR

Camden Staff-Patient Activity Record

Name of Unit: ________________________________

Please record:
1) Date and time of the observation
2) Number of service users resident at the unit
3) Number of service users engaged in each of the categories below.

Please refer to attached guidance for scheduled recording times

Recorded by	**Day**	**Date**	**Time**	**Total Service users resident**	***In the Unit***		***Out of the unit***		**Not Known**
					With staff	**Not with staff**	**With staff**	**Not with staff**	
e.g. AA	***Monday***	***29/02/06***	***10.00am***	***20***	***5***	***8***	***2***	***4***	***1***
	Monday		10.45						
	Monday		13.15						
	Monday		15.45						
	Monday		17.15						
	Tuesday		11.15						
	Tuesday		13.45						

Recorded by	**Day**	**Date**	**Time**	**Total Service users resident**	***In the Unit***		***Out of the unit***		**Not Known**
					With staff	**Not with staff**	**With staff**	**Not with staff**	
e.g. AA	***Monday***	***29/02/06***	***10.00am***	***20***	***5***	***8***	***2***	***4***	***1***
	Tuesday		16.15						
	Tuesday		18.15						
	Wednesday		09.15						
	Wednesday		11.45						
	Wednesday		14.15						
	Wednesday		19.15						
	Thursday		09.45						
	Thursday		12.15						
	Thursday		14.45						
	Thursday		20.15						
	Friday		08.15						

Recorded by	**Day**	**Date**	**Time**	**Total Service users resident**	***In the Unit***		***Out of the unit***		**Not Known**
					With staff	**Not with staff**	**With staff**	**Not with staff**	
e.g. AA	***Monday***	***29/02/06***	***10.00am***	***20***	***5***	***8***	***2***	***4***	***1***
	Friday		10.15						
	Friday		12.45						
	Friday		15.15						
	Saturday		10.15						
	Saturday		13.45						
	Saturday		16.45						
	Saturday		18.45						
	Sunday		10.45						
	Sunday		13.15						
	Sunday		16.15						
	Sunday		19.15						

How to use CaSPAR

Who is resident?

- Service users who spent last night on the unit
- Service users who were admitted on day of recording
- **Include** current service users away from the unit except those on extended (overnight) leave

Who is with staff?

- Service users actively engaged (interacting) in any way with a member of staff (e.g. talking, playing a game, eating together, going out for a walk, receiving medication)

Who is not with staff?

- Service users alone
- Service users with other service users or family/friends
- Service users near staff but where staff are not interacting (e.g.) a service user queuing up to see staff or being in the same communal room but not interacting with staff)

Who is in?

- Service users in the unit itself
- Service users in another part of the building/hospital (e.g. participating in a group activity)
- Service users currently resident who are known to be in the unit's garden or grounds. *Record service users using periods of day leave or agreed time away from the service as out.*

Times of recordings

- 28 recordings will take place at each participating service at the times listed below
- A maximum of 2 recordings per day and 10 recordings altogether to be made in any one week

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
10.45	11.15	9.15	9.45	8.15	10.15	10.45
13.15	13.45	11.45	12.15	10.15	13.45	13.15
15.45	16.15	14.15	14.45	12.45	16.45	16.15
17.15	18.15	19.15	20.15	15.15	18.45	19.15

APPENDIX 4

CaRICE

Camden Record of Inpatient Care Events

CaRICE

Camden Record of Inpatient Care Events

All staff at this service are being asked to record your direct contact with patients during your working day. This is to measure all the care provided to patients at this service. Please use this form record contact they have with resident service users.

Please record when you have **any** face-to-face contact for five minutes or longer with a patient in your service. You do not need to record which patient(s) you saw. You will need to record:

- o Length of contact
- o Purpose of the contact (see purpose of contact sheet)

- Please record contacts with more than one service user at the same time as one contact (e.g. if you are running a group)

- If more than one member of staff sees a service user together, please could each member of staff record this as a contact (e.g. at a meeting)

Purpose of contact

Please record the main purpose of each contact with a service user as one of the categories of care on the attached list. If your contact involves more than one intervention, please record all the types of care you provided.

e.g. if you saw a service user at 10.15am for 25 minutes to help them with a housing application and to give them a depot injection, this would be recorded as below:

Time of contact	**Length of contact (in minutes)**	**Purpose(s) of contact** ***(put number(s) for type of contact)***
10.15am	*25*	*1, 16*

*** Please return your completed recording sheet before you leave work***

Thank you very much for your time and help with this research project.

CaRICE Recording Sheet

Name ____________________ **Service** ____________________

Job Title ____________________

Start of shift ____________**am/pm** **End of shift** ______________ **am/pm**

Date ______________________ **Ref. No.** **(researcher use only)** ________

Time of contact	**Length of contact (in minutes)**	**Purpose of contact** *(put number(s) for type(s) of contact)*
e.g. 10.15	***25***	***1,16***

Photocopy this page if further pages are needed

CaRICE: Categories of Care

Purpose of contact	Description
1	Help with **housing problems**, (e.g. help finding somewhere to live or making current housing more suitable)
2	Help with **financial problems** (such as claiming benefits or managing debts)
3	Help with **legal matters** (e.g. providing a letter or report for the court or his/her solicitor, providing information about his/her legal rights or help to access legal services)
4	**Current activity:** help to plan or engage in social, leisure, occupational or religious activities while staying at this service
5	**Future activity:** help to plan or access work, education, social, leisure or religious activities when he/she leaves this service
6	Help for him/her to practice or improve **skills in every day tasks** (e.g. managing shopping, cooking using a washing machine, self care)
7	Contact between staff and his/her **family, friends or carers** to help support him/her and them
8	**Assessment of his/her difficulties** (help focused mainly on asking him/her about the nature or severity of current problems or his/her life history)
9	Help (other than medication) focused mainly on finding **ways to cope with distressing feelings, thoughts and experiences** (e.g. low mood, panic attacks, intrusive or strange thoughts or hearing voices)
10	Help focused mainly on resolving or managing difficulties with **current relationships**
11	Help focused mainly on resolving or managing difficulties with negative or traumatic **past events**
12	Help focused mainly on problems he/she is experiencing with **drug or alcohol use**
13	Help focused mainly on **providing him/her with information or explanation** about the nature of his/her mental health or psychological problems or his/her diagnosis
14	Help with **concerns or questions he/she has about his/her medication** (e.g. providing information about treatment options or side effects, discussing advantages or disadvantages of medication or problems he/she is experiencing)
15	**A review or change of his/her current medication** *(only include medication for mental health problems or side-effects of mental health medication in this category)*
16	Practical help with **taking medication** (such as staff dispensing medication or giving an injection, or reminding him/her to take medication)
17	Help with his/her **physical health** (treatment, investigations, tests or help to access physical health services) *(Don't include prescribed medication for mental health problems or side-effects from mental health medication in this category)*
18	**A member of staff staying with him/her** continuously or at regular intervals to make sure he/she or others are safe at times when he/she has been distressed or disturbed.
19	**A member of staff physically restraining him/her** to make sure he/she or others are safe at times when he/she has been distressed or disturbed.
20	**A meeting** for him/her and everyone involved with his/her care to discuss his/her current needs and concerns
21	Help to plan or arrange **care from other mental health services once he/she leaves this service**

APPENDIX 5

CCCQ-P

Camden Content of Care Questionnaire (Patient version)

CCCQ-P
Camden Content of Care Questionnaire (Patient version)

Participant ID:

Participant D.o.B.

Service:

Date of admission:

Date questionnaire completed:

Researcher:

Please use this form to record the amount and types of care received by the service user from staff at this service during this admission.

Question	Type of help	**Have you received any help of this sort from staff at this service during your stay?** **0 = No** **1 = Yes**	**If yes, how frequently have you received this sort of help from staff at this service?** **1 = once only** **2 = less than once a week** **3 = once a week** **4 = two or three times a week** **5 = more than three times a week but less than every day** **6 = once a day** **7 = more than once a day**
1	Help with **housing problems**, (e.g. help finding somewhere to live or making current housing more suitable)		
2	Help with **financial problems** (such as claiming benefits or managing debts)		
3	Help with **legal matters** (e.g. providing a letter or report for the court or your solicitor, providing information about your legal rights or help to access legal services)		
4	**Current activity:** help to plan or engage in social, leisure, occupational or religious activities while staying at this service		
5	**Future activity:** help to plan or access work, education, social, leisure or religious activities for when you leaves this service		
6	Help for you to practice or improve **skills in every day tasks** (e.g. managing shopping, cooking using a washing machine, self care)		
7	Contact between staff and your **family, friends or carers** to help support you and them		
8	**Assessment of your difficulties** (help focused mainly on asking you about the nature or severity of current problems or your life history)		
9	Help (other than medication) focused mainly on finding **ways to cope with distressing feelings, thoughts and experiences** (e.g. low mood, panic attacks, intrusive or strange thoughts or hearing voices)		
10	Help focused mainly on resolving or managing difficulties with **current relationships**		
11	Help focused mainly on resolving or managing difficulties with negative or traumatic **past events**		
12	Help focused mainly on problems you are experiencing with **drug or alcohol use**		
13	Help focused mainly on **providing you with information or explanation** about the nature of your mental health or psychological problems or your diagnosis		
14	Help with **concerns or questions you have about your medication** (e.g. providing information about treatment options or side effects, discussing advantages or disadvantages of medication or problems you are experiencing)		
15	**A review or change of your current medication** *(only include medication for mental healh problems or side effects of mental health medication in this category)*		
16	Practical help with **taking medication** (such as staff dispensing medication or giving an injection, or reminding you to take medication)		
17	Help with your **physical health** (treatment, investigations, tests or help to access physical health services) *(Don't include prescribed medication for mental health problems or side-effects from mental health medication in this category)*		
18	**A member of staff staying with you** continuously or at regular intervals to make sure you or others are safe at times when you have been distressed or disturbed.		
19	**A member of staff physically restraining you** to make sure you or others are safe at times when you have been distressed or disturbed.		
20	**A meeting** for you and everyone involved with your care to discuss your current needs and concerns		
21	Help to plan or arrange **care from other mental health services once you leave this service**		

Appendix 6
Additional data from psychometric testing of study measures

Table A6.1: CCCQ-S reliability: comparison at different services

	Was care received: kappa values				How frequently was care received Spearman's Rho values		
	All data n=46	Alternative 3 only n=20	Standard 4 only n=26		All data n=46	Alternative 3 only n=20	Standard 4 only n=26
1 housing	.618	.565	.655		.681	.697	.655
2 finances	.506	.432	.570		.531	.507	.587
3 legal	.312	.118	.519		.348	.108	.586
4 current activity	.401	.500	.302		.352	.511	.156
5 future activity	.129	-.053	.231		.179	-.148	.459
6 adl	.605	.692	.527		.603	.643	.544
7 family	.207	.381	-.008		.307	.350	.436
8 assessment	.148	-----***	.152		-.047	-.390	.078
9 coping strategies	-.137	-.071	-.193		.189	.033	.158
10 relationships	.362	.400	.320		.336	.365	.315
11 past events	.191	.000	.316		.191	-.022	.449
12 drugs alcohol	.652	.694	.601		.661	.631	.637
13 illness education	.335	.286	.372		.578	.652	.525
14 meds concordance	.328	.341	.316		.386	.459	.326
15 meds review	.276	.406	.103		.127	.395	-.223
16 meds practical	.313	.059	.494		.399	.086	.530
17 physical	.306	.216	.308		.415	.232	.548
18 observations	.247	.138	.210		.238	.135	.212
19 restraint	.292	------***	.339		.329	----***	.400
20 care planning	.244	-.038	.152		.379	-.050	-.143
21 aftercare	-.027	-.167	.196		-.085	.186	-.133
Mean scores		.258	.309			.269	.338

*** no kappa value calculated: rater 1 or 2 is a constant

Table A6.2: Service clustering: CaSPAR (Duncan's multiple range test)

service name	N	Subset for alpha = .05	
		1	2
Alternative 3	28	4.55	
Standard 1	28	7.97	
Alternative 1	26	8.55	
Standard 4	28	8.77	
Alternative 4	28	10.75	
Standard 2	28	13.52	
Standard 3	28	14.05	
Alternative 2	28		21.60
Sig.		.107	1.000

Table A6.3: Service clustering: CaRICE contact per patient per day (Duncan's multiple range test)

service name	N	Subset for alpha = .05		
		1	2	3
Standard 4	5	82.56		
Alternative 4	5	109.59	109.59	
Standard 1	5	121.74	121.74	121.74
Standard 3	5		131.71	131.71
Alternative 1	5		133.33	133.33
Alternative 2	5		139.50	139.50
Standard 2	5			154.12
Alternative 3	5			160.37
Sig.		.055	.068	.077

Table A6.4: Service clustering: CaRICE staff time in contact with patients (Duncan's multiple range test)

service name	N	Subset for alpha = .05				
		1	2	3	4	5
Alternative 1	5	.1143				
Standard 4	5	.1586	.1586			
Standard 1	5		.2157	.2157		
Alternative 4	5			.2334		
Alternative 2	5			.2500	.2500	
Standard 2	5			.2671	.2671	
Alternative 3	5				.3119	.3119
Standard 3	5					.3404
Sig.		.142	.061	.120	.054	.341

Appendix 7: Staffing levels data

The effect of staffing levels on staff-patient contact: secondary analysis of CaRICE data

Method

CaRICE data were collected and stored as described in Chapter 7. Variables for the minutes of contact with staff per patient per day and the proportion of staff time spent in contact with patients were calculated for each service each day as described in Section 5.2.2. The number of staff per patient at each service each day was created as an additional variable: (the total minutes spent at work by all staff during the 24 hour period)/(1440 (the number of minutes in 24 hours) x number of patients resident).

Data from five days at four alternative and four standard services (n = 40) were used to compare scores at alternative and standard services and correlations for the following variables:

Mean scores for alternatives and standard services were compared using t-tests of:
i) the proportion of staff time spent in contact with patients
ii) staffing levels

The following correlations were investigated using Spearman's rho:
i) staffing level and minutes of contact per patient per day
ii) staffing level and proportion of staff time spent in contact with patients

Results

Descriptive data for staffing levels, minutes of contact per patient per day and proportion of staff time spent in contact with patients at services are presented in Table A7.1 below.

Table A7.1 Staffing levels and staff-patient contact: descriptive data

Service	Day	Staffing level (staff per patient)	Minutes of contact per patient	Proportion of staff time spent with patients
Alternative 1	Mon	.58	105.00	.13
	Tues	.85	133.75	.11
	Wed	1.10	161.00	.10
	Thurs	1.38	211.00	.11
	Fri	.72	133.33	.13
Alternative 2	Mon	.53	152.50	.20
	Tues	.42	145.00	.24
	Wed	.40	123.57	.22
	Thurs	.31	147.86	.33
	Fri	.34	128.57	.26
Alternative 3	Mon	.42	257.50	.43
	Tues	.44	141.00	.22
	Wed	.29	105.00	.25
	Thurs	.30	140.00	.32
	Fri	.32	158.33	.34
Alternative 4	Mon	.31	83.74	.19
	Tues	.33	109.00	.23
	Wed	.32	149.24	.33
	Thurs	.37	117.47	.22
	Fri	.31	88.53	.20
Standard 1	Mon	.35	94.11	.19
	Tues	.41	149.69	.25
	Wed	.40	116.94	.20
	Thurs	.39	130.41	.23
	Fri	.39	117.53	.21
Standard 2	Mon	.41	157.25	.27
	Tues	.33	113.52	.24
	Wed	.62	171.69	.19

Service	Day	Staffing level (staff per patient)	Minutes of contact per patient	Proportion of staff time spent with patients
Standard 2	Thurs	.39	178.53	.32
	Fri	.33	149.65	.32
Standard 3	Mon	.29	123.16	.30
	Tues	.33	144.58	.31
	Wed	.26	142.98	.39
	Thurs	.24	124.52	.35
	Fri	.24	123.32	.36
Standard 4	Mon	.42	78.33	.13
	Tues	.33	79.05	.17
	Wed	.35	80.90	.16
	Thurs	.37	82.38	.15
	Fri	.35	92.15	.18

Descriptive data for staffing levels, minutes of contact per patient per day and proportion of staff time spent in contact with patients at alternative and standard services are presented in Table A7.2 below.

Table A7.2: Staffing levels and patient contact: descriptives by service type

	Staffing level: staff per patient (s.d.)	Minutes of contact per patient (s.d.)	Proportion of staff time spent with patients (s.d.)
Alternatives	0.50 (0.30)	139.57 (39.84)	0.227 (0.89)
Standard Services	0.36 (0.08)	122.54 (31.00)	0.245 (0.76)
All services	0.43 (0.23)	131.05 (36.28)	0.236 (0.08)

The mean number of staff per patient at alternative and standard services was compared using a t-test. This is shown in Table A7.3.

Table A7.3: Comparison of staffing levels at alternatives and standard services

CaRICE staff per patient		t	d.f.	p	95% Confidence Intervals (of mean difference)
Alternatives (20 days of data)	Standards (20 days of data)				
0.50	0.36	2.07	21.91	.051	-0.00, 0.28

The proportion of staff time spent in contact with patients at alternative and standard services was compared using a t-test. This is shown in Table A7.4.

Table A7.4:

Comparison of staff time with patients at alternatives and standard services

CaRICE Proportion of staff time spent in contact with patients		t	d.f.	p	95% Confidence Intervals (of mean difference)
Alternatives (20 days of data)	Standards (20 days of data)				
0.227	0.245	-0.69	37.11	.494	-0.07, 0.04

The correlation between staffing levels and minutes of contact provided per patient and between staffing levels and proportion of staff time spent with patients were compared for the 40 days of data from CaRICE using Pearson's correlation. These correlations are shown in Table A7.5.

A6.5 Correlations between stafing level and staff-patient contact (Pearson's correlation)

Correlates	n	Correlation coefficient	p
Staff per patient and minutes of contact per patient per day	40	0.404	0.010
Staff per patient and proportion of staff time spent in contact with patients	40	-0.592	<0.001

The clinical implications of the secondary analyses presented here are discussed in Chapter 10, Section 10.3.3.

Appendix 8: Individual service item scores

Table A8.1 CaRICE social interventions subscale categories: individual service scores

	CaRICE minutes per patient per day: mean score (5 days' data per service)						
service name	Housing	Finances	Legal	Current activity	Future activity	Activities of daily living	Family
Alternative 1 (non clinical crisis house)	14.86	5.89	1.81	16.24	10.18	6.08	4.17
Alternative 2 (clinical crisis house)	0.00	0.14	1.24	87.20	3.06	5.97	0.43
Alternative 3 (crisis team beds)	1.75	3.19	0.00	5.10	0.19	4.75	9.94
Alternative 4 (Tidal Model ward)	3.20	1.39	0.24	25.77	2.89	8.33	1.47
Standard 1	3.07	4.04	4.46	12.55	6.88	9.78	2.41
Standard 2	1.44	0.22	2.10	35.10	4.72	11.20	2.20
Standard 3	0.83	1.48	2.19	14.20	3.09	6.39	2.90
Standard 4	0.76	0.97	2.17	17.75	1.22	6.90	2.79

Table A8.2 CCCQ-P social interventions subscale categories : individual service scores

	CCCQ-P data social interventions subscale categories: mean scores						
service name	Housing	Finances	Legal	Current activity	Future activity	Activities of daily living	Family
Alternative 1 (non clinical crisis house)	1.98	0.39	0.37	2.43	1.00	0.73	0.79
Alternative 2 (clinical crisis house)	0.00	0.00	0.30	3.45	0.60	1.30	2.25
Alternative 3 (crisis team beds)	0.77	0.56	0.38	1.05	1.16	0.72	2.21
Alternative 4 (Tidal Model ward)	0.58	0.60	0.80	2.50	0.76	0.88	2.55
Standard 1	0.69	0.40	0.24	2.52	1.31	1.05	1.81
Standard 2	0.53	0.45	0.75	3.28	1.10	0.98	1.75
Standard 3	0.51	0.57	0.28	2.45	0.74	0.86	1.42
Standard 4	0.50	0.52	0.20	2.30	1.00	0.61	2.48

Table A8.3: CaRICE psychological interventions subscale categories: individual service scores

	CaRICE minutes' psychological interventions per patient per day: mean score (5 days' data per service)				
service name	Coping strategies	relationships	past events	drugs and alcohol	illness education
Alternative 1 (non clinical crisis house)	9.28	8.34	7.31	11.55	5.18
Alternative 2 (clinical crisis house)	4.35	0.92	3.14	0.18	2.21
Alternative 3 (crisis team beds)	18.04	4.87	2.20	9.73	1.72
Alternative 4 (Tidal Model ward)	7.48	3.36	4.88	0.82	1.87
Standard 1	8.82	2.27	1.93	5.38	2.48
Standard 2	7.84	1.71	2.44	1.06	2.56
Standard 3	13.64	2.18	2.74	1.14	3.22
Standard 4	8.38	1.00	2.54	1.23	3.24

Table A8.4 CCCQ-P psychological interventions subscale categories : individual service scores

CCCQ-P data psychological interventions subscale categories: mean scores					
service name	coping strategies	relationships	past events	drugs and alcohol	illness education
Alternative 1 (non clinical crisis house)	1.39	0.76	1.20	0.78	0.54
Alternative 2 (clinical crisis house)	1.47	1.05	1.10	1.25	0.75
Alternative 3 (crisis team beds)	1.85	0.95	1.51	0.64	1.87
Alternative 4 (Tidal Model ward)	1.50	0.53	1.10	0.58	1.05
Standard 1	1.43	0.36	0.75	0.68	1.21
Standard 2	1.55	0.35	0.93	1.00	1.35
Standard 3	1.79	0.33	0.86	0.55	1.17
Standard 4	2.02	0.50	1.24	1.00	1.85

Table A8.5: CaRICE physical and pharmacological interventions subscale categories: individual service scores

	CaRICE minutes' physical and pharmacological interventions per patient per day: mean score (5 days' data per service)					
service name	medication concordance	medication review	medication practical help	physical health	observations	restraint
Alternative 1 (non clinical crisis house)	3.50	1.87	0.68	0.89	4.60	3.09
Alternative 2 (clinical crisis house)	0.63	2.48	13.09	0.71	0.00	0.00
Alternative 3 (crisis team beds)	3.83	4.61	9.26	3.00	42.80	0.67
Alternative 4 (Tidal Model ward)	3.62	2.23	5.03	2.68	14.71	0.47
Standard 1	3.58	4.34	11.65	4.06	3.32	3.33
Standard 2	2.33	2.00	11.33	10.73	32.29	1.68
Standard 3	2.05	1.80	4.40	3.08	42.61	0.30
Standard 4	2.93	4.72	2.65	2.62	7.39	0.05

Table A8.6 CCCQ-P physical and pharmacological interventions subscale categories : individual service scores

CCCQ-P data physical and pharmacological interventions subscale categories: mean scores

service name	medication concordance	medication review	medication practical help	physical health	observations	restraint
Alternative 1 (non clinical crisis house)	0.88	0.07	1.71	0.39	0.46	0.00
Alternative 2 (clinical crisis house)	0.89	1.30	6.55	2.90	0.79	0.00
Alternative 3 (crisis team beds)	1.32	1.05	6.28	0.56	0.54	0.00
Alternative 4 (Tidal Model ward)	1.65	1.60	6.45	1.60	1.00	0.23
Standard 1	1.60	1.43	5.40	2.23	2.13	0.26
Standard 2	1.38	1.43	5.32	3.13	1.18	0.65
Standard 3	0.65	1.09	5.81	1.72	2.44	0.33
Standard 4	1.74	1.69	6.30	2.24	1.87	0.22

Table A8.7: CaRICE general care organisation interventions subscale categories: individual service scores

CaRICE minutes' general care organisation interventions per patient per day: mean score (5 days' data per service)			
		Care Plannning	
service name	Assessment	Meetings	Care Coordination
Alternative 1 (non clinical crisis house)	15.66	16.22	1.43
Alternative 2 (clinical crisis house)	10.55	3.19	0.00
Alternative 3 (crisis team beds)	31.45	2.00	1.25
Alternative 4 (Tidal Model ward)	13.89	2.98	2.26
Standard 1	13.13	12.55	1.71
Standard 2	12.74	6.96	1.73
Standard 3	18.95	3.26	1.17
Standard 4	7.59	4.28	1.38

Table A8.8 CCCQ-P General care organisation interventions subscale categories : individual service scores

CCCQ-P data general care organisation interventions subscale categories: mean scores

service name	Assessment	Care Planning Meetings	Care Coordination
Alternative 1 (non clinical crisis house)	2.44	0.98	0.95
Alternative 2 (clinical crisis house)	2.15	1.16	0.84
Alternative 3 (crisis team beds)	3.46	0.56	1.87
Alternative 4 (Tidal Model ward)	2.88	1.90	1.13
Standard 1	2.73	1.77	1.65
Standard 2	2.83	1.13	0.73
Standard 3	3.49	1.56	0.82
Standard 4	2.74	1.87	1.46

Appendix 9

Residential alternatives to acute psychiatric hospital admission: systematic review

Article published in British Journal of Psychiatry
August 2009

BJPsych
The British Journal of Psychiatry (2009)
195, 109–117. doi: 10.1192/bjp.bp.108.058347

Review article

Residential alternatives to acute psychiatric hospital admission: systematic review

Brynmor Lloyd-Evans, Mike Slade, Dorota Jagielska and Sonia Johnson

Background
Reducing use of hospital wards and improving their quality are central aims of mental health service policy. However, no comprehensive synthesis is available of evidence on residential alternatives to standard acute psychiatric wards.

Aims
To assess the effectiveness and cost-effectiveness of and satisfaction with residential alternatives to standard acute in-patient mental health services.

Method
A systematic search identified controlled studies comparing residential alternatives with standard in-patient services. Studies were described and assessed for methodological quality. Results from higher quality studies are presented and discussed.

Results
Twenty-seven relevant studies were identified. Nine studies of moderate quality provide no contraindication to identified alternative service models and limited preliminary evidence that community-based alternatives may be cheaper and individuals more satisfied than in standard acute wards.

Conclusions
More research is needed to establish the effectiveness of service models and target populations for residential alternatives to standard acute wards. Community-based residential crisis services may provide a feasible and acceptable alternative to hospital admission for some people with acute mental illness.

Declaration of interest
None.

In-patient acute mental health services have been found to be unpopular with service users,[1,2] often failing to address individuals' needs or provide a safe and therapeutic environment.[3,4] Home treatment is not appropriate for all at times of acute illness because of levels of risk or adverse social circumstances, so residential acute services offering an alternative to standard in-patient care are of considerable interest. Several types of innovative residential acute service have been developed in Europe and North America over recent decades.[5] However, their effectiveness remains uncertain. Two relevant Cochrane reviews have been published. Johnstone & Zolese[6] review length of stay on acute wards, but include as brief-stay wards services with no maximum length of stay or admission for up to 4 weeks, arguably too similar to current standard care to be seen as an alternative to it. Joy & Saylan[7] have reviewed mother and baby units, finding no studies for inclusion. A number of reviews without meta-analyses identify some relevant studies.[8–14] All however are either old, reliant on search strategies too limited to be considered systematic, or narrowly focused on a specific service model. There is no systematic synthesis of the current evidence regarding all types of residential and in-patient alternatives to standard in-patient care.

A recent UK survey proposed criteria to define a service as a residential alternative to standard acute in-patient care.[15] This review uses these criteria and reviews the evidence for each type of alternative. The aims of the review are to examine the effectiveness and cost-effectiveness of and satisfaction with alternative services, and to identify the major research gaps.

Method

Inclusion criteria

Studies meeting the following criteria were included in the review.

(a) Study type: randomised controlled trials (RCTs), two-group non-randomised cohort studies or one-group interrupted time series studies providing a specific quantitative comparison of the effectiveness and/or acceptability of residential alternatives and standard acute in-patient services.

(b) Participants: adults aged 16–65 years assessed by a mental health professional as needing acute in-patient admission.

(c) Interventions: residential acute mental health services that offer an alternative to standard acute psychiatric wards in one of the following five ways: based in the community (non-hospital services such as crisis houses); time limited (services offering admission with a time limit or planned maximum stay of 14 days or fewer); dedicated to a specific diagnostic group (e.g. first-episode psychosis or borderline personality disorder); dedicated to a specific socio-demographic group (e.g. wards for specific ethnic groups); implementing a specific therapeutic model involving changes to the working practices of more than one professional group.

(d) Outcomes: any outcome relating to clinical improvement or social functioning, service use, costs or cost-effectiveness, satisfaction with services.

Search strategy

A systematic search was undertaken of seven electronic databases – Medline, PsycINFO, Web of Science, CINAHL, EMBASE, the Cochrane Library (Database of Abstracts of Reviews of Effects and Central Controlled Trials Register) and the National Health Service (NHS) Economic Evaluation Database – covering the period from January 1966 to February 2008. Search terms for acute, residential mental health services were combined with search terms for different types of alternatives: terms were searched in the title and abstract except where stated. No language restrictions were applied. The Medline search was conducted on PubMed as follows:

(a) crisis intervention[MeSH].exp OR crisis OR acute OR emergency

AND

(b) residential treatment[MeSH].exp OR residential OR hospital* OR inpatient* OR 'crisis house' OR 'community beds' OR 'crisis beds'

AND

(c) mental disorders[MeSH].exp OR mental OR psychiatr*

AND

(d) 'disorder specific' OR 'disorder-specific' OR specialist OR 'socio-demographic' OR sociodemographic OR ethnic* OR disabled OR disability OR 'visual* impair*' OR blindness OR deaf OR 'mother and baby' OR 'parent and child' OR 'therapeutic model' OR 'model of care' OR alternative OR innovative OR 'brief admission' OR 'brief stay' OR 'short stay' OR 'time limited' OR community OR 'non-hospital' OR 'sponsor homes' OR 'accredited accommodation' OR Soteria OR ('crisis intervention' AND Trieste) OR 'Tidal Model' OR 'refocusing model'.

Search terms were modified as necessary to search other databases. Reference lists of all included articles and review articles were also hand searched. 'Grey' literature was searched through directories of conference proceedings and additional unpublished studies or papers in press were sought by contacting experts within the field.

Data abstraction

Two of the authors (B.L.-E. and D.J.) independently scanned titles from all identified studies and from abstracts where relevant and available, then retrieved and read the full text of all potentially relevant studies. Queries about inclusion were discussed and any disagreement resolved by a third reviewer (S.J.). Study details, including type of study, service and participant characteristics, duration of study period and study outcomes, were collected using a standard data extraction form.

Quality assessment

The methodological quality of each study included in the review was assessed using a standard form adapted from the quality assessment tool of Thomas.[16] Studies were rated as strong, moderate or weak regarding potential selection bias, allocation bias, accounting for confounders, masking, data collection methods and withdrawals, according to the criteria set out in the dictionary accompanying the quality assessment tool.[16] The analysis strategy and intervention integrity of studies were also noted.

Operational criteria were created from quality ratings to distinguish studies of high, moderate and low quality overall. Studies that were rated as high quality reported allocation concealment during randomisation, analysed data based on intention-to-treat and rated strong in all domains in the Thomas tool bar masking. All other RCTs were rated as moderate quality, as were non-randomised studies which demonstrated no significant difference or adjusted in analyses for difference between experimental and control groups for the confounder of severity of illness at admission and rated at least moderate for all Thomas criteria assessed except masking.

Reporting of study results

This review includes outcomes only from studies that were rated as moderate or high quality. Outcomes are reported from all these studies: mean figures for service use and cost data are presented if reported; effect size and *P*-values are reported for all outcomes where there is a significant difference between alternative and standard services.

Results

Twenty-seven studies were identified for inclusion in the review, of which nine were rated as moderate or high quality. Figure 1 summarises the study flow through the review.

Types of service studied

The review identified studies of community-based[17–31] and time-limited[32–38] services, and services with a specific therapeutic model.[39–43] No studies of in-patient services for specific diagnostic or sociodemographic groups were found, although some studies of community-based services also included clinical or sociodemographic inclusion criteria for participants.

Community-based services

Fifteen studies of community-based services published from 1969–2006 were identified (Table 1; see online Table DS1 for a more detailed version of this table). Eleven of these were of US services.

A range of non-hospital service models have been evaluated. Four studies[21,25,26,29] concern Soteria houses. Developed by Loren Mosher in California in the 1970s,[44] Soteria services provided care for people with first- or second-episode psychosis in informal settings. Primary staff were not clinically trained and worked long shifts (36–48 h) designed to help them to attune to and engage with residents. Staff and residents shared responsibility for household tasks and there was minimal reliance on antipsychotic medication. The model has been replicated more recently in a number of European countries (Switzerland, Germany, Sweden, Hungary, Finland), and evaluated primarily by Luc Ciompi and colleagues in Bern, Switzerland.[25,26] Other crisis hostels described in studies in this review also provide care in small, homely settings, typically about 8- to 12-bedded. Services display variation in closeness of links with statutory services. Boardman *et al*[19] describe a residential unit embedded within a community mental health resource centre, staffed by clinically qualified mental health

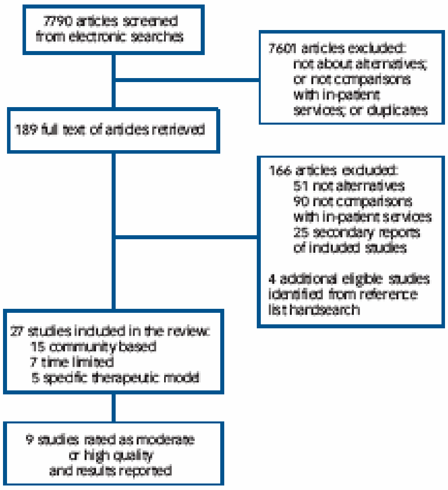


Fig. 1 Selection of studies for inclusion in systematic review.

Table 1 Characteristics of studies of community-based services[a]

Study reference	Service description	Study design and duration	Outcomes assessed[b]	Quality rating and main limitations[c]
Timko *et al*[17] (2006)	Veterans' community residential facilities, California, USA	RCT 30-day follow-up (from discharge)	1, 2, 4	Moderate 1, 2
Hawthorne *et al*[18] (2005)	6 crisis hostels (11–14 bedded), San Diego, USA	RCT 2-month follow-up	1, 2, 3, 4	Moderate 1, 2, 7 (some satisfaction data collected by service staff)
Boardman *et al*[19] (1999)	Community mental health centre beds, UK	Prospective non-randomised quasi-experiment 1-year follow-up	1, 2, ,3, 4	Moderate 1, 2
Fenton et al[20] (1998)	Crisis hostel (8 beds), Maryland, USA	RCT 6-month follow-up	1, 2, 3, 4	Moderate 1
Mosher *et al*[21] (1995) (Soteria study 2)	Soteria crisis hostel, California, USA	RCT 6-week follow-up	1	Moderate 1, 2, 6
Polak & Kirby[22] (1976)	Adult family placement, Colorado, USA	RCT 4-month follow-up	1, 3	Moderate 1, 2, 3
Readhead *et al*[23] (2002)	Adult family placements, UK	Interrupted time series study 1-year comparison period	2, 4	Low 2, 3
Hawthorne *et al*[24] (1999)	5 crisis hostels, San Diego, USA	Prospective non-randomised quasi-experiment 4-month follow-up	1, 2, 3	Low 1, 2, 4
Ciompi *et al*[25] (1993)	Soteria crisis hostel, Switzerland	Prospective non-randomised quasi-experiment 2-year follow-up	1, 2, 4	Low 2, 3
Ciompi *et al*[26] (1992)	Soteria crisis hostel, Switzerland	Non-randomised quasi-experiment (not stated if retrospective) 6-week follow-up	1, 4	Low 2, 3
Rappaport *et al*[27] (1987)	45 bed crisis hostel, California, USA	Retrospective non-randomised cohort study Assessment at discharge	1	Low 1, 2, 4
Bittle *et al*[28] (1986)	2 crisis hostels (10 bedded), Illinois, USA	Retrospective non-randomised cohort study 40-month follow-up	2	Low 2, 3
Mosher & Menn[29] (1978) (Soteria study 1)	Soteria crisis hostel, California, USA	Prospective non-randomised (pseudo-randomised) quasi-experiment 2-year follow-up	1, 2	Low 1, 2, 4
Brook[30] (1973)	Crisis hostel, Denver, USA (time limited to 7 days)	Non-randomised cohort study (not specified if retrospective) 6-month follow-up	1, 2	Low 2, 3
Goveia & Tutko[31] (1969)	Crisis hostel, California, USA	Prospective non-randomised quasi-experiment (some but not all participants randomised) 12-month follow-up	1, 2, 3	Low 1, 2, 3, 4, 5

RCT, randomised controlled trial.
a. A more detailed version of this table (Table DS1) is available online.
b. Domains: 1, improvement; 2, service use; 3, satisfaction; 4, cost.
c. Key to aspects of study quality: 1, analysis based on completer data not all intended to treat; 2, allocation concealment unclear (RCTs); not randomised (non-RCTs); 3, confounders (including severity of illness) not measured and if necessary adjusted for in analysis; 4, more than 40% of potential participants declined to participate or number not stated; 5, more than 40% participants lost at follow-up; 6, unspecified or previously unpublished outcome measure; 7, other.

professionals including supervision by psychiatrists; Timko *et al*[17] by contrast describe independently run services contracted by statutory agencies. Two studies in this review[17,18] describe services funded by American Veterans Associations, providing care mainly to veterans. Apart from Soteria hostels, crisis houses in studies in this review are not described as guided by a clear, manualised model of care: types of care provided are not described in detail.

Placement with individual families is the other model of community-based alternative to hospital identified in this review. Polak & Kirby evaluated family sponsor homes developed in the 1970s in Denver Colorado,[22] where host families supported one or two acutely ill people, with training and assistance from local crisis services. This scheme specifically aimed to divert people from acute hospital admission. Support included involving individuals in a normal family environment with participation in meals and other domestic tasks and activities. The model has been used on a small scale in Powys, Wales in the last decade.[23]

Time-limited services

Seven studies evaluated time-limited hospital in-patient services ranging from 1 day to 8 days maximum planned admission (Table 2). Studies dated from 1966 to 1996 and evaluated services in North America and the UK. All the brief-stay wards in studies in this review were located within larger hospital complexes and accepted general acute admissions. Brief-stay services report an aim to increase the intensity with which care was provided compared with standard acute wards (e.g. providing assessment, medication review, help solving psychosocial problems, after-care

Table 2 Characteristics of studies of time-limited services

Study reference	Service description	Study design and duration	Participants and total *n* (alternative *n*/comparison *n*)	Outcomes assessed[a]	Quality rating and main limitations[b]
Olfson[32] (1990)	5-day time-limited crisis admission ward, New England, USA	Prospective non-randomised quasi-experiment 3-month follow-up	Adults assessed as requiring acute admission: diagnosis of schizophrenia, one or more previous admissions, stable housing, no current substance misuse or major medical problems *n* = 26 (8/18)	1, 2	Moderate 2
Hirsch *et al*[33] (1979)	8-day target hospital admission ward, UK	RCT 1-year follow-up	Adults aged 16+ requiring acute admission No diagnosis of brain injury or major physical health problem *n* = 224 (115/109)	1, 2	Moderate 1, 2
Herz *et al*[34] (1975)	1-week target hospital admission ward, New York, USA	RCT 7-month follow-up	Adults aged 16+ requiring acute admission who live with a responsible adult, have a diagnosis of mental illness. Limitations on comorbidity *n* = 175 (51: brief hospital; 61: brief hospital and day care; 63: standard hospital)	1, 2	Moderate 2, 3
Schneider & Ross[35] (1996)	3-day crisis admission ward, Connecticut, USA	Retrospective non-randomised cohort 30-day follow-up	Adults assessed as requiring acute admission *n* = 1370 (590/780)	2	Low 2, 3
Ianzito *et al*[36] (1978)	24-h admission ward, Massachusetts, USA	Retrospective non-randomised cohort study 2-week follow-up	Adults assessed as requiring acute admission *n* = 184 (83/101) (also 198 non-admitted patients evaluated)	1, 3	Low 2, 3, 6
Voineskos *et al*[37] (1972)	3-day crisis admission ward, Canada	Retrospective non-randomised cohort Duration of initial admission assessed	Adults assessed as requiring acute admission *n* = 868 (439/429)	2	Low 2, 3
Mendel[38] (1966)	7-day time-limited admission ward, California, USA	Prospective non-randomised quasi-experiment 18-month follow-up	Adults 18+ assessed as requiring acute admission: diagnosis of schizophrenia, admission from community, voluntary or on 72-h section *n* = 443 (114/329)	1, 2	Low 1, 2, 3, 5

RCT, randomised controlled trial.
a. Domains: 1, improvement; 2, service use; 3, satisfaction; 4, cost.
b. Key to aspects of study quality: 1, analysis based on completer data not all intended to treat; 2, allocation concealment unclear (RCTs); not randomised (non-RCTs); 3, confounders (including severity of illness) not measured and if necessary adjusted for in analysis; 4, more than 40% of potential participants declined to participate or number not stated; 5, more than 40% participants lost at follow-up; 6, unspecified or previously unpublished outcome measure; 7, other.

planning within the period of a brief admission), but distinctly different interventions are not described.

Services with a distinctive therapeutic model

Five studies of services with a specific therapeutic model involved one of two UK nursing-led models of care developed in the last decade, the Tidal model or the Refocusing model (Table 3). The Tidal model, developed by Phil Barker in the 1990s[45] seeks to avoid a perceived reductionist approach of relating to people purely as patients with symptoms that need to be treated, by valuing people's own narrative of illness and perception of problems. Frequent, collaborative contact between staff and patients is encouraged through regular assessment of problems and goals. This involves documenting individuals' expressed needs and problems verbatim. The Refocusing Model, developed on acute wards in Bradford, UK, in the late 1990s,[43] increased nurses' authority to take risk management decisions and thus minimise or eradicate formal observations on wards, with time instead spent on more collaboratively agreed contact between staff and patients.

Study characteristics

Twenty-seven studies included in this review comprised 7 RCTs, 7 non-randomised prospective quasi-experimental studies and 13 before/after comparison studies. Most were of small or medium size, only four studies having more than 250 participants.[24,28,37,38] Duration of studies ranged from the period of admission only up to 40-month follow-up.

No studies identified by this review were assessed as high quality. Nine studies were rated as moderate quality, including two well-designed quasi-experimental studies in addition to seven RCTs. Three of these studies were of brief-stay wards[32–34] and six of community-based services, residential crisis beds[17–21] or family placement.[22] None rated strong on all criteria of the Thomas assessment tool.[16] Only one study[20] clearly described allocation concealment procedures during randomisation. Where participants were lost to follow-up, no studies based analysis on intention-to-treat; all provided completer data only. Reported levels of loss of participants during the course of studies varied substantially for studies of moderate quality. Only two[22,34] included in outcomes data all potential participants assessed as eligible; reported overall attrition rates on individual outcomes

Table 3 Characteristics of studies of services with a distinct therapeutic model

Study reference	Service description	Study design and duration	Participants and total *n* (alternative *n* /comparison *n*)	Outcomes assessed[a]	Quality rating and main limitations[b]
Berger *et al*[39] (2006)	Tidal model ward, Canada	Interrupted time series study 6-month comparison period: outcomes during admission assessed	Adults on an acute admission ward who consent to participate *n* = 46 (not stated)	3	Low 1, 2, 3, 4
Lafferty & Davidson[40] (2006)	Tidal model ward, Glasgow, UK	Interrupted time series study 1-year comparison period	Adults on an acute ward *n* not stated (service level data only collected)	1/3	Low 2, 3
Gordon *et al*[41] (2006)	Tidal model ward, Birmingham, UK	Retrospective non-randomised cohort study 1-year comparison period: outcomes during admission assessed	Adults on an acute ward *n* not stated (service level data only collected)	1/3	Low 2, 3
Stevenson *et al*[42] (2002)	Tidal model ward, Newcastle, UK	Interrupted time series study 6-month comparison period Duration of initial admission assessed	Adults on an acute ward *n* = 150 (81/69)	2	Low 2 ,3
Dodds & Bowles[43] (2001)	Refocusing model ward, Bradford, UK	Interrupted time series study 6-month comparison period: outcomes during admission assessed	Adults on an acute ward *n* not stated (service level data only collected)	1	Low 2, 3

a. Domains: 1, improvement; 2, service use; 3, satisfaction; 4, cost.
b. Key to aspects of study quality: 1, analysis based on completer data not all intended to treat; 2, allocation concealment unclear (randomised controlled trials (RCTs)); not randomised (non-RCTs); 3, confounders (including severity of illness) not measured and if necessary adjusted for in analysis; 4, more than 40% of potential participants declined to participate or number not stated; 5, more than 40% participants lost at follow-up; 6, unspecified or previously unpublished outcome measure; 7, other.

in other moderate-quality studies range from 3 to 52% of potential participants. Service use data, gathered from routinely collected records, was generally more comprehensive than assessment of participants' functioning or satisfaction.

Of the 18 remaining studies assessed as low quality, 15 did not adequately measure or adjust for confounders. Five had high risk of selection bias (more than 40% of those eligible declining to participate); two had unacceptably high withdrawal rates (more than 40%) for all outcomes and one used only unpublished outcome measures. Full details of the quality assessment of all studies can be found in online Table DS2.

Study outcomes

Outcomes assessed and results from studies of higher (moderate) quality found in this review are presented in Table 4 (a more detailed version, including details of the outcomes assessed, can be found as online Table DS3).

All six moderate-quality studies of community-based alternatives reported measures of symptoms and/or global improvement: four[17,20–22] found no significant differences from standard services, whereas two[18,19] found some results favouring alternatives and some showing no significant difference. Three[18,19,22] out of four studies that assessed satisfaction reported moderately and significantly greater satisfaction with the alternative than the standard service. Three[17,18,20] out of four studies that assessed cost also reported results favouring the alternative service. The exception was the community mental health team community beds assessed by Boardman *et al*:[19] a statutory service staffed by qualified mental health professionals, in contrast to the other community-based services described in this review, this service cost the NHS more per patient than standard acute wards, although with no significant overall difference to public services. Only with regard to service use were any results favourable to standard care: two of three studies reported length of index admission greater at community alternatives than standard wards;[17,20] one of four found in-patient bed-days, including index admission, over the study follow-up period significantly fewer for standard services.[17]

Of the three moderate-quality studies of time-limited alternatives, only one found the brief-stay ward able to discharge individuals within the planned admission period and significantly more quickly than standard wards.[34] One study[32] found a majority of people required transfer to a standard ward at the end of the brief-stay period; one study[33] found the 8-day planned admission period not rigorously adhered to and mean length of stay not significantly shorter than standard care. None found any significant differences in clinical outcomes or readmission rates. All five low-quality studies of services using a distinctive therapeutic model reported some outcomes favourable to alternative services regarding patient satisfaction or levels of untoward incidents on wards.

Discussion

Findings of the review

Current research is insufficient to provide convincing evidence about the effectiveness or acceptability of residential alternatives to standard acute in-patient mental health services. No studies of services for specific demographic groups or people with specific diagnoses were identified, despite literature providing descriptions of some such services as promising service models, e.g. mother and baby units[46] or psychosis, affective disorder and personality disorder wards.[47] Studies of services with a distinctive therapeutic model were all of low quality, failing to account for differences between groups in analysis. The before and after comparison provided by most studies of services with a distinctive therapeutic model and their lack of stated primary outcomes also exacerbate risks of reporting and publication biases. The feasibility of brief-stay acute wards is brought into question by the fact that in three moderate-quality studies, in only one[34] was the alternative service able to discharge a majority of individuals within the planned admission period. The applicability of this finding to contemporary

Table 4 Results from studies of moderate or high quality[a]

Study reference	Results
Community-based services	
Timko *et al*[17] (2006)	Favours alternative: total out-patient visits (104 *v.* 130: $P<0.001$), 1-year cost ($22 000 *v.* $33 000: $P=0.002$) Favours standard service: length of index admission (26 *v.* 55 days: $P<0.001$), 1-year total in-patient bed days (78 *v.* 86 days: $P<0.01$) No significant difference: ASI psychiatric subscale score at 1 year
Hawthorne *et al*[18] (2005)	Favours alternative: discharge SF–36V ($P=0.02$) and POC ($P=0.05$) scores, cost of index admission ($P=0.001$), homelessness ($P=0.001$) at discharge No significant difference: discharge PANSS and ASI scores, PANSS, SF–36V, ASI scores, homelessness and number of readmissions at 2-month follow-up
Boardman *et al*[19] (1999)	Favours alternative: GAF ($P=0.02$), HRSD ($P=0.01$), PSE ($P=0.001$), VSSS overall satisfaction ($P=0.02$) No significant difference: HoNOS, SBS, CAN, length of index admission, number readmitted in 12-month follow-up, cost to all public services (although cost to NHS significantly higher at alternative services)
Fenton *et al*[20] (1998)	Favours alternative: cost of index admission significantly less ($3046 *v.* $5549: effect size 0.78, $P<0.001$) Favours standard service: length of index admission (12 *v.* 19 days: $P<0.002$) No significant difference: PANSS scores, satisfaction, 6-month costs, cost-effectiveness
Mosher *et al*[21] (1995)	No significant difference
Polak & Kirby[22] (1976)	Favours alternative: satisfaction: TES score (patient report) at discharge ($P<0.001$) and 4-month follow up ($P<0.01$) No significant difference: all measures of clinical improvement
Time-limited services	
Olfson[32] (1990)	No significant differences (only 3/8 participants discharged from brief-stay service within planned 5-day limit)
Hirsch *et al*[33] (1979)	No significant differences (median length of stay but not mean length significantly shorter at alternative)
Herz *et al*[34] (1975)	Favours alternative: length of index admission (9 days *v.* 50 days: no *P* stated; in-patient days over 2-year follow-up (47 *v.* 115: $P<0.001$) No significant difference: PSS or GAS total scores at 3 months or 2 years, number of participants readmitted over 2-year follow-up

ASI, Addiction Severity Index; SF–36V, Health Survey – Short Form (Veterans Version); POC, Perceptions of Care Questionnaire; PANSS, Positive and Negative Symptom Scale; GAF, Global Assessment of Functioning; HRSD, Hamilton Rating Scale for Depression; PSE, Present State Examination; VSSS, Verona Service Satisfaction Scale; HoNOS, Health of the Nation Outcome Scale; SBS, Social Behaviour Schedule; CAN, Camberwell Assessment of Need; NHS, National Health Service; TES, Treatment Effectiveness Scale; PSS, Psychiatric Status Schedule; GAS, Global Assessment Scale.

a. A more detailed version of this table is available online as Table DS3.

mental health service settings may be limited however, as the moderate-quality studies of time-limited services identified in this review all pre-date the advent of modern community resources such as home-treatment teams. The most recent study of time-limited services included in the review, from 1996,[35] found that 69% of those admitted to a 3-day admission ward could be discharged into the community within this period, but the comparability of individuals with those admitted to general acute wards was unclear.

Despite the larger number of studies of community-based alternatives, the strength of evidence provided is nevertheless limited by the quality of included studies and the heterogeneity of services and participants studied. A crucial question is to what extent community-based alternatives can admit a population comparable with standard acute wards. All six moderate-quality studies of community-based services imposed different inclusion criteria for participants and all but one[22] included some criteria (such as veterans only, consenting to participate, no admissions in the previous 12 months) beyond those normally required for real-life acute admission. Two of the studies specifically excluded detained patients;[18,20] the remaining four did not report whether or how many participants were detained. The systematic exclusion from studies of some people who require acute admission and the not insignificant drop-out rates reported by some studies limit the strength and applicability of their results. Findings from studies of community-based services may only be applicable to a subgroup of people requiring acute admission, excluding some of those who are most severely unwell or least cooperative.

Notwithstanding its limitations, the evidence overall for community-based services retains the possibility that non-hospital residential crisis services are a useful service model: no clinical outcomes from studies in this review were worse than standard wards for community-based alternatives; the few differences reported tend to favour community-based models. That five studies successfully randomised participants assessed as requiring acute admission to either community alternatives or standard wards does suggest some similarity in populations served. As a whole, the studies provide preliminary evidence that for some people with acute mental health problems, community-based alternatives may be as effective and potentially less costly and more acceptable than standard in-patient wards. More generally, current research provides no contraindication to any of the types of alternatives included in this review.

Methodological issues and limitations

Residential alternatives are not clearly defined or described by a consistent terminology, providing a challenge for comprehensive retrieval of relevant studies. This review defined alternative services consistently with a current UK study of residential alternatives.[15] Over-inclusive search terms were used in the initial search to minimise the risk of missing relevant studies, but this search found no studies that appeared to describe an innovative acute residential service and did not meet the review's inclusion criteria. This suggests the list of studies included in the review was relatively comprehensive.

In order to avoid ignoring available evidence in an under-researched area, non-randomised studies were included in the review. Assessment of study quality was conducted to inform consideration of the strength of evidence provided. The tool used in this review[16] was recommended by Deeks *et al*[48] in a review of quality assessment tools and recommended as suitable for randomised and non-randomised studies. It assesses the domains of quality identified as important in the most recent Cochrane handbook[49] (although with less detailed assessment of randomisation procedures) and additionally allows higher and lower quality non-randomised trials to be distinguished. Study ratings for masking were not used to assess overall study quality: masking of participants and guarantee of consistent concealment from raters of participants' care status were not considered possible for health services research of this type. In order to rate high overall, in addition to rating strong on all other assessment tool criteria, studies were required to report adequate allocation procedures during randomisation and analyse data based on intention-to-treat, as these two dimensions of quality have been identified as most associated with estimates of treatment effects.[50] Only from studies rated moderate quality or above, based on the aspects of quality rated with the Thomas tool,[16] were results presented in this paper. This provides some safeguard from creating a false impression of the weight and strength of evidence concerning alternative services by collating results from numerous poor-quality studies. However, the absence of any studies rated as highest quality in this review indicates the need for additional caution about the precision of results.

A meta-analysis of data from studies included in this review was initially planned. However, only a minority of moderate-quality studies, describing services of considerable heterogeneity, could provide any data usable in meta-analyses. In these circumstances, a meta-analysis was not considered appropriate. Data from moderate-quality studies which were potentially usable or not usable for meta-analyses are however presented in online Table DS4.

Clinical implications

Current research evidence provides clinicians and commissioners with only very limited guidance about effective models of acute in-patient mental healthcare. Several service models identified in this review – Soteria houses, adult family placements, time-limited wards – have been developed in more than one country or time period without ever becoming a well-established part of a national acute service system. This suggests some doubt about their sustainability and/or usefulness, although also a persisting perception of a need to seek alternatives to standard acute care. Residential services that can only cater for a proportion of people requiring acute admission may be perceived by service planners and commissioners as a luxury and be vulnerable to losing funding. The failure of alternative service models to endure may also reflect a reliance of innovative services on charismatic leaders and local champions, without whom they may not thrive. The community beds embedded in a community mental health resource centre evaluated by Boardman and colleagues[19] allay some of these concerns. They were able to admit a reasonably high proportion (65%) of people assessed as requiring acute admission during the study period, can accept detained patients and are still running currently, a decade later. Drayton Park, a women's crisis house in North London, has also been established for more than a decade and evaluated in qualitative studies as providing a valuable role in local acute care.[51,52] This suggests that in a contemporary UK context, community crisis beds can constitute an important and sustainable part of local acute in-patient provision.

The dissatisfaction of many service users with standard psychiatric wards[1,2] suggests a need for alternatives. The potential for emergency residential accommodation outside the hospital setting to improve service user choice and thus the acceptability of services, while relieving bed pressure on acute wards, was identified 10 years ago.[53] Even if alternative service models can only divert a subgroup of people requiring acute admission, the increased scope this might bring for focusing appropriate facilities and expertise in in-patient services for a higher risk, predominantly detained client group is potentially useful. This review found no evidence against alternative models of care and, consistent with previous qualitative research,[52,54] provides an indication that crisis beds in non-hospital settings may increase satisfaction with acute residential services for users. Certainly, this review provides no discouragement to service managers and commissioners to consider innovation in the provision of acute in-patient care.

Future research

A recent service mapping study indicates that alternatives, as defined in this review, constitute almost 10% of current acute in-patient beds in England, with each type of alternative included in this review represented.[15] Many of the service models are far from new, with some to which papers included in the review relate dating back to the 1960s. Yet this review found only limited evidence for any and none for some types of alternative. This means that, where alternatives have been established, clinical practice is running ahead of the research evidence base. It goes some way towards explaining why some service models first described as promising several decades ago, such as crisis placements in family homes and brief-stay admission wards, have yet to be widely adopted despite the wish among service planners, clinicians and service users to develop alternatives to standard acute wards. The conclusion that there is a need and an opportunity for more research is an inescapable one in this area. A further research question beyond the scope of this review was how users of residential alternatives compare with people receiving crisis home treatment: information about the extent to which alternatives accommodate people who cannot be adequately treated at home would also illuminate their role and potential utility in the acute care system.

Only 9 of 27 studies identified for inclusion in this review were assessed as moderate quality and none as high quality, limiting the certainty with which any conclusions about the services being evaluated can be drawn. This highlights the need for research to be of good quality if it is to be useful. Key quality indicators that were absent from studies of moderate quality included in this review were arranging adequate allocation concealment and conducting analysis based on intention-to-treat and stating primary outcomes in advance. Insufficiently thorough description of participants' characteristics (e.g. whether people were detained or not) and inadequate reporting of variance in outcomes data were also common shortcomings. Only two studies[17,32] included in the review provided detailed quantitative comparison of the content of care provided at alternative and standard services, although five more,[21,25–27,29] including all the studies of Soteria houses, compared participants' medication use. More information about care provided would help identify differences in alternative service models and the extent and nature of difference from standard services, open the black box of service provision[55] and aid understanding of service outcomes. The studies included in this review evaluate young services or recently established service innovations. Evaluation of more established, enduring alternatives would also be desirable, in order to investigate whether outcomes, perhaps especially satisfaction, are sustainable and not merely a function of service novelty.

There are particular challenges to conducting RCTs in acute mental health settings, where the need for immediate intervention makes both the logistics of randomisation and the process of informing participants and obtaining consent problematic. In such circumstances, quasi-experimental studies may be more feasible and have strong real-world applicability by evaluating outcomes for cohorts that include all service users.[56] Two non-randomised, natural experiment studies were assessed as moderate quality and reported results broadly congruent with those from

RCTs included in the review. This indicates that a well-designed non-randomised study, which accounts for important confounders, may have a useful place in acute mental health service research. The development of clear protocols for ethically acceptable recruitment in mental health crises would also be very helpful, addressing issues such as how to conduct urgent randomisation out of hours when researchers are not available and how to deal with the often transient loss of capacity experienced by many people at the time of a crisis.

Compared with many models of community care, developments in acute in-patient care remain unevaluated. The treatment of people at times of crisis is clearly a crucial part of mental healthcare: establishing effective models of providing residential acute care should be a priority for future mental health services research.

Brynmor Lloyd-Evans, MSc, Department of Mental Health Sciences, University College London; **Mike Slade**, PhD, King's College London, Health Services and Population Research Department, Institute of Psychiatry, London; **Dorota Jagielska**, MA, **Sonia Johnson**, MD, Department of Mental Health Sciences, University College London, UK

Correspondence: Brynmor Lloyd-Evans, Department of Mental Health Sciences, University College London, Charles Bell House, 67–73 Riding House Street, London W1W 7EJ, UK. Email: b.lloyd-evans@ucl.ac.uk

First received 11 Aug 2008, final revision 11 Aug 2008, accepted 3 Dec 2008

Acknowledgements

This report presents independent research commissioned by the National Institute for Health Research Service Delivery and Organisation Programme. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

References

1 Quirk A, Lelliott P. What do we know about life on acute psychiatric wards in the UK? A review of the research evidence. *Soc Sci Med* 2001; 53: 1565–74.

2 Rose D. *Users' Voices, The Perspectives of Mental Health Service Users on Community and Hospital Care*. The Sainsbury Centre, 2001.

3 Muijen M. Acute hospital care: ineffective, inefficient and poorly organised. *Psychiatr Bull* 1999; 23: 257–9.

4 Barker S. *Environmentally Unfriendly: Patients' Views of Conditions on Psychiatric Wards*. MIND, 2000.

5 Warner R. *Alternatives to the Hospital for Acute Psychiatric Treatment*. American Psychiatric Press, 1995.

6 Johnstone P, Zolese G. Length of hospitalisation for people with severe mental illness. *Cochrane Database Syst Rev* 2000; 2: CD000384.

7 Joy CB, Saylan M. Mother and baby units for schizophrenia. *Cochrane Database Syst Rev* 2007; 1: CD006333.

8 Mattes JA. The optimal length of hospitalization for psychiatric patients: a review of the literature. *Hosp Community Psychiatry* 1982; 33: 824–8.

9 Arce AA, Vergare M. An overview of community residences as alternatives to hospitalization. *Psychiatr Clin North Am* 1985; 8: 423–36.

10 Stroul BA. Residential crisis services: a review. *Hosp Community Psychiatry* 1988; 39: 1095–9.

11 Braun P, Kochansky G, Shapiro R, Greenberg S, Gudeman J, Johnson S, et al. Overview: deinstitutionalisation of psychiatric patients, a critical review of outcome studies. *Am J Psychiatry* 1981; 138: 736–49.

12 McCabe L, Butterill D, Goering P. Residential crisis units. Are we missing out on a good idea? *Can J Common Ment Health* 2004; 23: 65–74.

13 Grawe RW, Ruud T, Bjorngaard JH. Alternative emergency interventions in adult mental health care (in Norwegian). *Tidsskr Nor laegeforen* 2005; 125: 3265–8.

14 Calton T, Ferriter M, Huband N, Spandler H. A systematic review of the Soteria paradigm for the treatment of people diagnosed with schizophrenia. *Schizophr Bull* 2008; 34: 181–92.

15 Johnson S, Gilburt H, Lloyd-Evans B, Osborn DPJ, Boardman J, Leese M, et al. In-patient and residential alternatives to standard acute psychiatric wards in England. *Br J Psychiatry* 2009; 194: 456–63.

16 Thomas H. *Quality Assessment Tool for Quantitative Studies*. Effective Public Health Practice Project, McMaster University, 2003.

17 Timko C, Chen S, Sempel J, Barnett, P. Dual diagnosis patients in community or hospital care: one-year outcomes and health care utilization and costs. *J Ment Health* 2006; 15: 163–77.

18 Hawthorne WB, Green EE, Gilmer T, Garcia P, Hough RL, Lee M, et al. A randomized trial of short-term acute residential treatment for veterans. *Psychiatr Serv* 2005; 56: 1379–86.

19 Boardman AP, Hodgson RE, Lewis M, Allen K. North Staffordshire Community Beds Study: longitudinal evaluation of psychiatric in-patient units attached to community mental health centres. I: Methods, outcome and patient satisfaction. *Br J Psychiatry* 1999; 175: 70–8.

20 Fenton WS, Mosher LR, Herrell JM, Blyler CR. Randomized trial of general hospital and residential alternative care for patients with severe and persistent mental illness. *Am J Psychiatry* 1998; 155: 516–22.

21 Mosher LR, Vallone R, Menn A. The treatment of acute psychosis without neuroleptics: six-week psychopathology outcome data from The Soteria Project. *Int J Soc Psychiatry* 1995; 41: 157–73.

22 Polak PR, Kirby MW. A model to replace psychiatric hospitals. *J Nerv Ment Dis* 1976; 162: 13–22.

23 Readhead C, Henderson R, Hughes G, Nickless J. Accredited accommodation: an alternative to in-patient care in rural north Powys. *Psychiatr Bull* 2002; 26: 264–5.

24 Hawthorne WB, Green EE, Lohr JB, Hough R, Smith PG. Comparison of outcomes of acute care in short-term residential treatment and psychiatric hospital settings. *Psychiatr Serv* 1999; 50: 401–6.

25 Ciompi L, Kupper Z, Aebi E, Dauwalder HP, Hubschmid T, Trutsch K, et al. The pilot project 'Soteria Berne' in treatment of acute schizophrenic patients. II. Results of a comparative prospective follow-up study over 2 years. *Nervenarzt* 1993; 64: 440–50.

26 Ciompi L, Dauwalder HP, Maier C, Aebi E, Trutsch K, Kupper Z, et al. The pilot project 'Soteria Berne'. Clinical experiences and results. *Br J Psychiatry* 1992; 161(suppl 18): s145–53.

27 Rappaport M, Goldman H, Thornton P, Stegner B, Moltzen S, Hall K, et al. A method or comparing two systems of acute 24-hour psychiatric care. *Hosp Community Psychiatry* 1987; 38: 1091–5.

28 Bittle RG. Acute care treatment services (ACTS): a model program for providing acute psychiatric services in a homelike environment in an institutional setting. *Behavioral Resident Treat* 2006; 1: 275–88.

29 Mosher LR, Menn AZ. Community residential treatment for schizophrenia: two-year follow-up. *Hosp Community Psychiatry* 1978; 29: 715–23.

30 Brook BD. Crisis hostel: an alternative to psychiatric hospitalization for emergency patients. *Hosp Community Psychiatry* 1973; 24: 621–4.

31 Govela LH, Tutko TA. *Psychiatric Rehabilitation in a Community Center*. Rehabilitation Mental Health Services, 1969.

32 Olfson M, Plotke G, Harding C, Jones J, Mayfield W. A controlled evaluation of inpatient crisis treatment for acute schizophrenia episodes. *Psychiatr Q* 1990; 61: 143–54.

33 Hirsch S, Platt S, Knights A, Weyman A. Shortening hospital stay for psychiatric care: effect on patients and their families. *BMJ* 1979; 1: 442–6.

34 Herz M, Endicott J, Spitzer R. Brief hospitalization of patients with families: initial results. *Am J Psychiatry* 1975; 132: 413–8.

35 Schneider SE, Ross IM. Ultra-short hospitalization for severely mentally ill patients. *Psychiatr Serv* 1996; 47: 137–8.

36 Ianzito BM, Fine J, Sprague B, Pestana J. Overnight admission for psychiatric emergencies. *Hosp Community Psychiatry* 1978; 29: 728–30.

37 Voineskos G, Morrison MF, Jain RC. The introduction of a crisis unit in a mental hospital. *Can Psychiatr Assoc J* 1974; 19: 445–52.

38 Mendel W. Effect of length of hospitalization on rate and quality of remission from acute psychiatric episodes. *J Nerv Ment Dis* 1966; 143: 226–33.

39 Berger JL. Incorporation of the Tidal Model into the interdisciplinary plan of care – a program quality improvement project. *J Psychiatr Ment Health Nurs* 2006; 13: 464–7.

40 Lafferty S, Davidson R. Putting the person first. *Ment Health Today* 2006; March: 31–3.

41 Gordon W, Morton T, Brooks G. Launching the Tidal Model: evaluating the evidence. *J Psychiatr Ment Health Nurs* 2005; 12: 703–12.

42 Stevenson C, Barker P, Fletcher E. Judgement days: developing an evaluation for an innovative nursing model. *J Psychiatr Ment Health Nurs* 2002; 9: 271–6.

43 Dodds P, Bowles N. Commentary. Dismantling formal observation and refocusing nursing activity in acute inpatient psychiatry: a case study. *J Psychiatr Ment Health Nurs* 2001; 8: 183–8.

44 Mosher L. Soteria and other alternatives to acute psychiatric hospitalization: a personal and professional review. *J Nerv Ment Dis* 1999; 187: 142–9.

45 Barker P. The tidal model: a person-centred approach to psychiatric and mental health nursing. *Perspect Psychiatr Care* 2001; 37: 79–87.

46 Royal College of Psychiatrists. *Perinatal Maternal Mental Health Services* (Council Report CR88). Royal College of Psychiatrists, 2000.

47 Bonsack C, Borgeat F, Dubois A. Impact of mental health services' specialization by diagnosis in Lausanne. I. Qualitative Study. *Ann Med Psychol* 2001; 159: 645–51.

48 Deeks J, Dinnes J, D'Amico J, Sowden A, Sakarovitch C, Song F, et al. Evaluating non randomised intervention studies *Health Technol Assess* 2003; 27: 1–173.

49 Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions* (eds J Higgins, S Green). Wiley, 2008.

50 Schulz K, Chalmers I, Hayes R, Altman D. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273: 408–12.

51 Killaspy H, Dalton J, McNicholas S, Johnson S. Drayton Park, an alternative to hospital admission for women in acute mental health crisis. *Psychiatr Bull* 2000; 24: 101–4.

52 Johnson S, Bingham C, Billings J, Pilling S, Morant N, Bebbington P, et al. Women's experiences of admission to a crisis house and to acute hospital wards: a qualitative study. *J Ment Health* 2004; 13: 247–62.

53 Sainsbury Centre for Mental Health. *Acute Problems: A Survey of the Quality of Care in Acute Psychiatric Wards*. The Sainsbury Centre, 1998.

54 Faulkner A, Petit-Zeman S, Sherlock J, Wallcraft J. *Being There in a Crisis: A Report of the Learning from Eight Mental Health Crisis Services*. The Mental Health Foundation and Sainsbury Centre for Mental Health, 2002.

55 Mechanic D. Emerging issues in international mental health services research. *Psychiatr Serv* 1996; 47: 371–5.

56 Gilbody S, Whitty P. Improving the delivery and organisation of mental health services: beyond the conventional randomised controlled trial. *Br J Psychiatry* 2002; 180: 13–8.

Psychiatry
in the movies

Paging Dr Love

Peter Byrne

When women's roles in the war effort brought their empowerment, from *The Flame Within* (1936) to *Lady in the Dark* (1944), the institution of movie psychiatry reminded women of their place – as passive recipients of male wisdom and treatments. The female movie psychiatrist (*The Flame Within*) is frequently no different from the successful but unhappy career woman (*Lady in the Dark*) – their career will never bring the same fulfilment as a solid marriage. The female movie psychiatrist must be 'cured' by her love for her male patient. Dr Constance Peterson (Ingrid Bergman) has no difficulties helping her male amnesic patient, accused of murder, escape confinement. She marries him at the denouement of *Spellbound* (1945). Rather than list over a hundred films where girl (psychiatrist) falls for boy (patient), the challenge is to name those that deviate from this storyline. Classic Hollywood depicted women therapists as inadequate, personally and professionally: *Knock on Wood* (1954), *A Perfect Furlough* (1958), *Wild in the Country* (1961), *A Very Special Favour* (1965) and *A Fine Madness* (1966). Similar unhappy archetypes continue to yearn for their male patients in modern films: *Mr Jones* (1993), *12 Monkeys* (1995) and *The Jacket* (2005). Perfect psychiatrist Dr Lowenstein must be rescued from her miserable personal life by an affair with her patient's brother in *Prince of Tides* (1991). In all these films, the only effective treatment is love. The audience are encouraged not to dwell on the boundary violations.

Male movie psychiatrists have romantic liaisons less frequently when taking into account the proportion of films where they are represented, but their behaviours are no less ludicrous: *What's New Pussycat?* (1965), *Beyond Therapy* (1987) and *Color of Night* (1994). Freud's provocative quote may help explain why male psychiatrists break fewer boundaries on the silver screen: 'for women, the level of what is ethically normal is different from what it is in men' – making the films here quoted truly Freudian.

The British Journal of Psychiatry (2009)
195, 117. doi: 10.1192/bjp.195.2.117

Appendix 10

“Assessing the content of mental health services: a review of measures”

Article published in Social Psychiatry and Psychiatric Epidemiology
August 2007

Soc Psychiatry Psychiatr Epidemiol (2007) DOI 10.1007/s00127-007-0216-x

ORIGINAL PAPER

Brynmor Lloyd-Evans · Sonia Johnson · Mike Slade

Assessing the content of mental health services: a review of measures

Received: 5 February 2007 / Accepted: 29 May 2007 / Published online: 27 June 2007

Abstract *Background* Measurement of service content is necessary to understand what services actually provide and explain variation in service outcomes. There is no consensus about how to measure content of care in mental health services. *Method* Content of care measures for use in mental health services were identified through a search of electronic databases, hand searching of references from selected studies and consultation with experts in the field. Measures are presented in an organising methodological framework. Studies which introduced or cited the measures were read and investigations of empirical associations between content of care and outcomes were identified. *Results* Twenty five measures of content of care were identified, which used three different data collection methods and five information sources. Seven of these measures have been used to identify links between content of care and outcomes, most commonly in Assertive Community Treatment settings. *Discussion* Measures have been developed which can provide information about service content. However, there is a need for measures to demonstrate more clearly a theoretical or empirical basis, robust psychometric properties and feasibility in a range of service settings. Further comparison of the feasibility and reliability of different measurement methods is needed. Contradictory findings of associations between service content and outcomes may reflect measures' uncertain reliability, or that crucial process variables are not being measured. *Conclusion* Measures providing a greater depth of information about the nature of interventions are needed. In the absence of a gold standard content of care measure, a multi-methods approach should be adopted.

Key words content of care – mental health services – process measurement

Introduction

The importance of understanding what actually happens in mental health care delivery is increasingly recognised. There is considerable variation in practice even amongst similarly labelled mental health services [13], such as case management teams [9]. Detailed investigation of the content of service interventions is therefore needed if differences in outcomes between services are to be understood [31]. However, a lack of complete or consistent approaches to describing mental health services has been identified [13]. It has been argued that valid measurement of the content of care provided to patients may be more crucial than attending to service style, setting or organisation in understanding the links between service processes and outcomes [25].

The place within mental health services research of measurement of the content of care at services can be identified with reference to existing conceptual frameworks. The Mental Health Matrix, for example, proposes two dimensions to formulate mental health service aims and practice [49]. In a temporal dimension, it distinguishes the process of care at services from the inputs and resources of a service and service outcomes. In a geographical dimension, care provided to patients within a service, by a local service system, or at a regional/national level can be distinguished. Within this framework, content of care measurement concerns the process of care at a patient level.

Measurement of the content of mental health services, including both the quantity and nature of

B. Lloyd-Evans (✉) · S. Johnson
Dept. of Mental Health Sciences
University College London
Charles Bell House
67-73 Riding House Street
London W1W 7EJ, UK

M. Slade
Health Services Research Dept.
Institute of Psychiatry
London, UK

interventions delivered to patients, can be identified as one of four ways in which mental health services have been described and classified [25], distinguishing service content from the style, setting or organisation of services. Description of the content of mental health services can be separated into several elements [6]—the nature, frequency, duration, scope, setting and style of care or "how much of what is done to whom, when and in what manner" [6, p. 284].

The concept of service content can thus be used to refer to the sum of what staff provide for patients at a service. The term "content of care" will be used for this purpose in this review. It will include both direct care: what staff do when they see patients and indirect care: what staff provide for patients in their absence. Arguably, content of care is a harder concept to define or measure than broader variables such as service type or specific single interventions such as pharmacological treatment. The clarity provided by valid measurement tools is consequently particularly needed.

There are at least five reasons to assess content of care in mental health services:

(i) *To describe service content*. Measurement of the content of care in services can identify differences in content provided by a service over time, between services, or to different groups of patients within a service [6].
(ii) *To measure model fidelity*. If established guidelines or operational criteria exist regarding the model of care to which a service is seeking to work, measurement of service content can be used to assess model fidelity and programme implementation [43].
(iii) *To understand service outcomes*. While not providing certainty, it can help generate hypotheses to explain service outcomes and identify active ingredients of complex interventions [25].
(iv) *To understand variation in patient outcomes*. Attending to what works for whom—identifying variation in the effectiveness of different service interventions for different groups of patients—has been advocated [37]. Patient level data about care received can help investigate this by illuminating whether variation in outcomes for groups of patients within a service may be due to differences in responsiveness to interventions or differences in interventions received.
(v) *To assess service quality*. If an element of service content has already been clearly demonstrated to produce good outcomes, it can be used as a measure of effectiveness or service quality [18].

Qualitative measures of content of care and quantitative measures of related variables can provide information about service content, but only to a limited extent. Three qualitative methods of inquiry in mental health research—in-depth interviews, focus groups and participant observation—have been identified [53], all of which can provide rich information about what happens at mental health services and how care is experienced. However, qualitative methods are ill-suited to comparing differences, potentially small but significant, in the number or types of interventions provided to representative groups of service users at services. In order to investigate associations between care provided and service outcomes and provide an empirical basis for identifying active ingredients of care, quantitative data is required.

Quantitative outcome measures may be used to draw inferences about the care provided at services. Most pertinently, measures of need such as The Camberwell Assessment of Need (CAN) [47], can be used to measure whether a service user's needs in different areas are met during a period of care, from either the service user's, carer's or staff's perspective. As an outcome measure, however, the CAN is limited as a process measure of content of care for the following reasons:

(i) It measures the effectiveness of care, not its provision. If someone receives considerable help with psychotic symptoms, which are not alleviated, for instance, this would be recorded in CAN as an unmet need, offering no record that care has been provided.
(ii) It measures whether needs are met but not how. For example, it is unclear whether someone with a met need for psychotic symptoms has received pharmacological or psychological treatment, or of what sort.
(iii) It provides little scope for differentiating how much care has been provided to individuals or at services. Inpatient care is always recorded as high-level care for example.

Specific quantitative measures of content of care are therefore required to measure what is provided in mental health services. Two organising frameworks, drawn from social research literature, can be applied to describe ways of measuring content of care:

(1) *Source of information*. Four sources of data for process measurement of social programmes or health services have been identified [42]: (1) direct observation by the researcher; (2) information from service records; (3) data from service providers; (4) data from service users. Measures may use a combination of data sources and have; (5) a mixed information source.
(2) *Method of data collection*. Two ways of conceptualising how to record activity are: recording in terms of time or in terms of incidents [11]. *Time recording* involves recording whatever is happening over a given period of time (instants, short periods or longer, continuous time periods) to specified person(s) or in a specified area. *Incident Recording* involves preselecting particular event(s) of interest and recording if and when these happen over a given period of time.

A further distinction can be made between contemporaneous and retrospective incident recording.

Here, the term *Event Recording* is used to describe methods of recording incidents at or very near the time they happen. *Retrospective Questionnaires* (completed by staff, patients, or researchers based on interviews, observation or reference to case records) is used to describe information about events of interest gathered retrospectively.

This literature review aims to identify existing measures of the content of care in mental health services. The measurement methods they employ will be presented. The empirical associations between content of care and outcomes found using the measures will be summarised and how far existing measures are able to meet the goals of content of care measurement will be considered. What is known about how best to measure the content of care in mental health services and directions for future research will be discussed.

Method

Identifying measures of content of care

Inclusion criteria

Measures were included which provide quantitative data about the amount and types of care provided at any type of specialist inpatient, residential or community mental health service for adults. Measures, which provided this information, despite having a different primary purpose (e.g. to measure patient activity or model fidelity) were included. Measures providing service level information only and measures providing individual patient level information, which could be aggregated to provide service information, were both included. Measures of direct care only or direct and indirect care were included.

Measures were excluded which assess related process factors (e.g. psychotherapy or pharmacotherapy rating scales, measures of continuity of care, service style, model fidelity or service quality) but do not assess provision of the amount and types of care.

Search strategy

The literature involving content of care measures does not use a consistent terminology and thus does not lend itself to straightforward retrieval from bibliographic databases. This review therefore uses a variety of methods to identify relevant studies.

(i) Medical and nursing electronic databases (Pubmed, Embase, PsycInfo, Cinahl) were searched using a Medical Subject Heading of "mental health services" or equivalent, combined with (1) generic terms for the content of mental health services—"content of care" or "process of care" or "process measure" in title or abstract; or (2) terms for specific methods of process measurement identified from reference works —"time recording" or "time sampling" or "time budget" or "event recording" or "incident recording". Publications from 1966–2006 were included in the search.
(ii) Reference lists from relevant studies identified in the electronic search were hand searched.
(iii) A group of six accessible experts involved in previous studies of content of care was asked for information on current studies or methodological approaches to content of care measurement in mental health services.

Data abstraction

The following characteristics of measures identified in this review were collected:

(i) Data collection method
(ii) Information source
(iii) Level of information provided: patient = care provided to individual patients; service = overall care provided at a service
(iv) Service settings the measure has been designed for/used in
(v) Established psychometric properties of the measure

Identifying the use of measures in process/outcomes investigation

Inclusion criteria

Studies were included in this part of the review if they used of one of the measures of content of care identified in this review to investigate associations between a defined content of care variable and subsequent inpatient admissions, clinical or social functioning or patient satisfaction.

Search strategy

(i) Studies presenting the measures included in this review were read in order to identify whether the measure had been used to investigate associations between content of care variables and outcomes.
(ii) Articles citing the above studies were identified through an electronic database. (No single database provided citations for all studies: Web of Science, PsycInfo and GoogleScholar were used.). These articles were also read to find any investigation of content of care/outcome associations using identified measures.

Data abstraction

The following information was collected about identified studies investigating associations between content of care and outcome:

(i) Content of care variable measured
(ii) Outcome variable measured
(iii) Study setting
(iv) Was an association between content of care and outcomes identified?
(v) Study reference

Results

25 measures of content of care were identified for inclusion in this review. The methods used by these measures are summarised in Table 1.

Titles and references for the individual measures are provided in sections i–iii below. The characteristics of measures are also described in these sections, grouped by data collection method.

(i) Event recording measures

6 event recording measures were identified (see Table 2).

Measures ask individual staff to record only their own contacts with clients, with the exception of the

Table 1 Methods used in measures of content of care

	Data collection method		
	Event recording	Time recording	Retrospective questionnaires
Information source			
Staff	6 measures	3 measures	2 measures
Service users	None	None	None
Observation by researchers	None	8 measures	None
Records	n/a	n/a	2 measures
Mixed	None	None	4 measures

structured record described by Patmore and Weaver [35], which requires one respondent to record all interventions received by a client from any member of staff at a service during the recording period.

Event Recording measures have only been used in community services. They vary in terms of:

- Collection method: All measures use paper recording forms except The Event Report [23], which required staff to use a pocket computer to complete daily records.
- Scope of information: Measures provide information about the content of care within a single service, except for The Mannheim Service Recording Sheet [44], which provides information about patients' use of the whole local mental health system.
- Depth of information: Measures record either face-to-face staff/patient contacts only [44] or a variety of types of staff activity, e.g. face-to-face, telephone or failed contact with a patient, contact with a carer and contact with another professional [12]. The nature/purpose of an intervention is categorised in measures as one of between 5 and 11 defined types of care (e.g. help with housing, medication review etc).

The psychometric properties of Event Recording measures have not been examined thoroughly. Only the Daily Contact Log [6] has been investigated regarding inter-rater reliability, through clinicians' ratings of case note vignettes and use of the measure in vivo by staff and researcher, following direct observation of clinical practice. Face validity alone has been established for the variables used in the measures.

Three rationales have been identified for the categorisation of types of care in event recording measures: (1) consistency with another established measure: e.g. The Mannheim Service Recording Sheet [44] mirrors the categories used in the International Classification of Mental Health Services [16]; (2) consistency with an established model of care: e.g. the Event Report [23] measures elements of Integrated Care, a model of care for people with schizophrenia [19]; (3) describing actual service practice: e.g. The Event Record [12] categories are informed by a rigorous Delphi Process with Intensive Case Managers [20], to ensure adequate and accurate reflection of their work practices.

(ii) Time recording measures

11 time recording measures were identified (see Table 3).

Measures record activity at a service at specific moments, during short periods of between 5 and 15 min, or continuously over whole days or shifts. They all provide information about the number of staff-patient interactions, although the main purpose of the measure may be to measure these interactions [14, 36, 45]; all staff activity [24, 35, 52, 55]; or all patient activity [5, 24, 27, 54]. All measures employ a paper recording method, but display a variety of approaches regarding:

- Scope of information: Researcher-observation measures record activity within a defined, observable area within a residential or inpatient service. Staff-report measures provide information about all activity within a service.
- Depth of information: Only the staff-completed time recording measures categorise the types of care provided in similar detail to event recording measures. Measures of staff activity distinguish different types of activity: for example, direct patient contact, indirect patient care, administrative work (e.g., record keeping) and other activity [24]. A number of observational measures record information about the quality of staff contacts with patients: for example rating them as accepting, tolerating or rejecting [45].

Inter-rater reliability testing of several researcher-observation based time recording measures indicate

Table 2 Event recording measures

Measure	Information source	Level of data provided	Service settings used in	Established psychometric properties
Daily contact log [6]	Staff	Patient	Assertive community teams (ACTs) (USA)	Adequate inter-rater reliability established
Mannheim service recording sheet [44]	Staff	Patient	Community and inpatient services (Spain and Germany)	
Event record [12]	Staff	Patient	ACTs and case management (UK)	
Event report [23]	Staff	Patient	"Integrated care" community service (Sweden)	
Untitled (structured record) [35]	Staff	Patient	Community mental health teams (CMHTs) (UK)	
Service activity log [21]	Staff	Patient	Case managers (community) (USA)	

Table 3 Time recording measures

Measure	Type of time recording	Information source	Level of data provided	Service settings used in	Established psychometric properties
Direct observation schedule [45]	Momentary	Researcher observation	Service	Mental health day centres (UK)	Inter-rater reliability established
Untitled (staff activity measure) [52]	Momentary	Researcher observation	Service	Adult acute inpatient wards (Australia)	Inter-rater reliability established
Dementia care mapping [27]	Short period	Researcher observation	Patient	Residential dementia care (UK)	Inter-rater reliability unacceptably low [50]
Untitled (Patient observation) [24]	Short period	Researcher observation	Patient	Adult acute inpatient wards (UK)	
Untitled (Staff observation) [24]	Short Period	Researcher observation	Service	Adult acute inpatient wards (UK)	
Quality of interactions schedule (QUIS) [14]	Short Period	Researcher observation	Service	Elderly mentally ill inpatient units (UK) [16] and adult acute wards (Nigeria) [36]	Inter-rater reliability established (in EMI setting)
Staff-Patient interaction chronograph [36]	Short period	Researcher observation	Service	Inpatient mental health services (USA)	Inter-rater reliability established.
Time Budget [55]	Short period	Staff	Service	Intensive Community Support Programmes (USA)	
Continuous time sampling [5]	Continuous	Researcher observation	Patient	Elderly mentally ill inpatient wards (UK)	Inter-rater reliability established
Time budget [54]	Continuous	Staff	Patient	Long-term Adult inpatient wards (UK)	Construct validity tested, but reliability not tested.
Untitled (staff diary) [35]	Continuous	Staff	Service	Community mental health teams (CMHTs) (UK)	

that observers can reliably identify what constitutes a staff-patient contact and rate whether that contact is positive, negative or neutral in nature. The reliability of staff-report time recording measures has not been tested.

No empirical basis for choice of categories of staff activity has been reported for any time recording measure beyond basic face validity. Only Wing and Brown report testing the construct validity of their measure [54]: time spent doing nothing, not engaged with staff or others, as measured by the Time Budget, did correlate with four other measures of poverty of the social environment.

(iii) Retrospective questionnaire measures

8 retrospective questionnaire measures were identified (see Table 4).

Information about the amount and types of care is obtained from a variety of information sources, but all measures are completed by researchers, bar the staff-completed measure of Kovess and Lafleche [28]. Two measures [48, 56], are primarily designed to measure services' model fidelity and one measure [2] to measure service cost, but all can provide information about service content. Retrospective questionnaires recording content of care vary regarding:

- Recording period: Measures are completed retrospectively for time periods varying from 1 month [2, 26] to 18 months [22].
- Scope of information: Measures provide information about the content of care provided across a service system [2, 26], or within one service.
- Depth of information: Of the retrospective questionnaire measures providing individual patient-level information, only two [2, 38] assess the specific number of interventions received by individuals. All retrospective questionnaires provide a measure of the amount of care provided at services, except the IC-MHC [16], which however identifies 10 different types of care, the most detailed information provided by a retrospective measure about the nature of care at services.

Demonstration of the psychometric properties of retrospective questionnaire content of care measures has not been extensive. The I.C.M.H.C., which has been demonstrated to have good inter-rater reliability [15], provides service level information about types of care only.

Content of care and outcome

7 measures included in this review were identified as having been used to investigate the association between content of care variables relating to amount, setting or nature of care and patient outcomes. These investigations are summarised in Table 5.

Table 4 Retrospective questionnaire measures

Measure	Information source	Level of data provided	Service settings used in	Established psychometric properties
Client service receipt inventory [2]	Staff	Patient	Community settings (various) (UK)	
Untitled (staff activity questionnaire) [28]	Staff	Service	Community mental health teams (Canada)	
Quality care intervention checklist [22]	Mixed: Staff, patients, carers	Patient	Community mental health services (USA, Japan, Italy)	
Process of care review form [38]	Records	Patient	Community Mental Health Centres (USA)	
Untitled (serve receipt form) [56]	Records	Patient	Community Mental Health Centres (USA)	
Dartmouth assertive community treatment scale (DACTS) [48]	Mixed (unspecified)	Service	Assertive Community Teams (USA and UK)	Predictive and construct validity investigated. Mixed results from inter-rater reliability testing.
International classification of mental health care (ICMHC) [16]	Mixed (unspecified)	Service	Various European services	Inter-rater reliability established [15]
European service mapping schedule (E.S.M.S). [26]	Mixed (unspecified)	Service	European services/local service systems	

Of 13 studies described here, 11 involve community-based services, 9 are of American services and 9 involve Assertive Community Treatment or proto-ACT services. The effect of the amount of staff-patient contact has been most widely investigated.

Discussion

This review has identified 25 measures of content of care in mental health services which use 6 different measurement methods. 7 measures have been used to investigate empirical associations between service content and outcomes.

Measures of content of care have been developed and used in a variety of service settings and offer a way to understand what services actually provide. This would not be possible through outcome studies alone. Progress in developing measures of content of care has been far from linear however. There is variation in existing measures regarding what is measured (direct care only or direct and indirect care) and how it is measured. The methodological framework presented in Table 1 shows that only a minority of possible methods of measuring content of care have been used in measures described in this review. This review finds that many measures lack a clear theoretical or empirical basis and/or have not been tested for psychometric properties. Many measures have been developed and used for a particular study, but not applied or further developed in future studies or different settings.

Where the association between content of care variables and outcomes has been investigated, findings have varied. Conflicting evidence exists, for example, for the most widely examined questions: whether amount of care [7, 8, 12, 17, 29] or ACT fidelity [3, 29, 30] in community-based services affect inpatient bed use.

The lack of repeated, consistent demonstration of association between any content of care variable and patient outcomes in part reflects the inherent difficulties this type of investigation, where numerous confounding factors other than received care will affect patients' subsequent health status [10]. It is not implausible, for example, that severity of illness could be associated with increased amount of treatment and poorer health outcomes for patients at a service. It is possible however, that the uncertain reliability of content of care measures used has obfuscated associations with outcomes, or that appropriate content of care variables have not measured. This review found that the majority of studies of process and outcome associations concerned the link between amount of direct care and outcomes. Studies, which assess what staff actually do when they see patients, to investigate links between the nature of care provided and outcomes, remain rare.

Table 5 The use of content of care measures to investigate associations between content of care and outcome variables

Content of care variable	Content of care measure	Association found with outcome?	Outcome variable	Setting	Study reference
Amount of care (number of staff contacts received per patient per month/year)	Daily contact log [6]	No	Inpatient admissions and bed days (1 year follow up)	7 ACT services (USA)	[17]
		Yes	Inpatient bed use and social functioning (1 year follow up)	Community support program for adults with schizophrenia (USA)	[8]
		No	Symptoms (1 year follow up)		
		Yes	Inpatient bed use, employment, independent living status (1 year follow up)	3 community services for adults with schizophrenia (USA)	[7]
	DACTS*[29]	Yes	Inpatient bed use (1 year follow up)	18 ACT services (USA)	[29]
	DACTS [48]	No	Symptoms, housing status, substance use, client satisfaction (2 year follow up)	2 ACT and 1 standard community team working with homeless dual disorder clients (USA)	[33]
	Event record [12]	No	Inpatient admissions (2 year follow up)	4 Intensive and standard case management services—clients with psychotic illness (UK)	[12]
	Mannheim Service Recording Sheet [44]	Yes	Reduction in unmet needs (skills and abilities) 1 year follow up	Community services in 2 regions - adults with schizophrenia (Spain and Germany)	[44]
		No	Reduction in unmet needs (symptoms) 1 year follow up		
Amount of care (time spent by patients doing nothing)	Time Budget [54]	Yes	Rating of clinical improvement (4 year follow up)	3 Long-stay psychiatric hospitals (UK)	[54]
Amount of care (duration of staff-patient contacts)	DACTS*[29]	No	Inpatient bed use (1 year follow up)	18 ACT services (USA)	[29]
Setting of care (community versus office-based contacts)	DACTS*[29]	No			
Specificity of care (proportion of interventions with vocational focus)	Daily contact log [6]	Yes	Employment status (3 year follow up)	3 community services for adults with schizophrenia	[7]
		No	Independent living status (3 year follow-up)		
Specificity of care (no of interventions providing referral or advocacy)	Service activity log [21]	No	Reduction in severity and number of rated problems (6 months+ follow up)	1 region's case management services (USA)	[21]
Quality of care (proportion of staff-patient contacts rated positive by observer)	Staff activity measure [52]	Yes	Reduction in rates of conflict and containment (1 year follow up)	2 acute inpatient services pre and post introduction of "City Nurses" (UK)	[4]
ACT-fidelity (composite measure)	DACTS *[29]	Yes	Inpatient bed use (1 year follow up)	18 ACT services (USA)	[29]
	DACTS [48]	Yes	Employment status (1 year follow up)	7 ACT and standard community services for veterans with severe mental illness (USA)	[40]
		No	Inpatient bed use (1 year follow up)	10 ACT teams (USA)	[3]
		Yes	Housing status and client satisfaction (2 year follow up)	2 ACT and standard community teams – homeless dual disorder clients (USA)	[33]
		No	Symptoms and substance use (2 year follow up)		
	Daily Contact Log [6] (main source from which ACT fidelity rating derived)	Yes	Inpatient admissions and substance use (3 year follow up)	7 ACT and standard case manangement services (USA)	[30]
		No	Symptoms, social functioning and satisfaction with services (3 year follow up)		

*forerunner of DACTS: IFACT

The need for effective content of care measurement in mental health services research has been highlighted repeatedly [10, 13, 31]. Criteria for effective content of care measurement, encompassing psychometric robustness, comprehensiveness, clinical credibility and feasibility, have been proposed [18, 51]. However, current measures of content of care in mental health services only partially meet these criteria. The following are four challenges to more effective content of care measurement:

Psychometric robustness

Evidence of inter-rater reliability has been provided most clearly and consistently for researcher-completed direct observation measures, which, however, provide more limited information about the mature of care provided than most other measures in this review. Whether a greater depth of information, or information from sources other than researcher observation, can be obtained as reliably, remains unclear. The work of Brekke [6] suggests that staff-report event recording measures can provide reliable information about the nature and amount of staff-patient contact at services [6], but the reliability of his Daily Contact Log has yet to be similarly demonstrated for other staff-report measures.

There are also obstacles, whatever methodological approach is used, to creating a valid measure, which accurately assesses significant elements of content of care. Case note extraction measures may rely on incomplete or inaccurate source material, as found in a study comparing information obtained from patient interviews and case notes [56]. Other retrospective questionnaires may be compromised by respondents' recall bias. All contemporaneous measures, meanwhile, may generate reactivity [32], i.e. where the process of measurement changes what is being measured. Participating in a research study, for example, could lead to a temporary increase in staff activity for the duration of a study. Staff-completed measures may also be vulnerable to deliberate distortion, to present a service in a good light.

The extent or comparative impact of these factors on the validity of different methods or measures is difficult to assess. A multi-methods and measures approach to assessing content of care may therefore be helpful: consistent findings from different measures could afford each a degree of convergent validity. This review suggests such an approach is rare, however: in practice, a measure is often developed for a specific study or service setting and used in isolation. The demonstration of clear links between service content and expected outcomes would also increase confidence that valid process variables are being accurately measured, but has also been rare.

Depth of information

A reasonable depth of information about the nature of care and types of intervention provided at services is necessary to understand what services actually do and begin to investigate what works for whom. Of the measures identified in this review however, even a comparatively informative measure with a clear theoretical basis (a Delphi Process with intensive case managers [20]), such as The Event Record [12], contains categorisations of types of care whose meaning is hard to infer—e.g. "specific mental health intervention". Other examples of descriptions of types of care whose breadth compromises clarity include: "Support" [23]; "Follow up" [21]; "1:1" [6].

This review found that studies of content of care in inpatient mental health services have assessed the amount and quality of care, but no measure designed for and used in inpatient settings describes the types of intervention provided. The paucity of our understanding of what happens in UK inpatient mental health wards has been highlighted [39]: however, there is no measure of inpatient service content with sufficient depth to help address this issue. If feasible and reliable measures could be developed to provide a greater specificity and depth of information about care provided at services than is currently possible, this would aid attempts to describe and distinguish services.

Feasibility

Content of care measures need to generate adequate completion rates to provide high quality information. Researcher-completed measures may be assumed to pose fewest obvious problems regarding completion rates. An adequate response rate (66%) has been reported for a contemporaneous staff-report measure [35], but most studies of staff-report content of care measures do not report a response rate. A good response rate (85%) has been reported for a staff-completed momentary time recording measure in an HIV case management setting [1], indicating this could be a useful method for mental health settings.

The difficulty of obtaining contemporaneous, staff-report data could potentially be greater in residential settings than community services, owing to staff's more numerous, briefer interactions with patients. However, we currently lack evidence with which to compare the feasibility of different methods of measuring content of care in similar service settings, or of any one measure in different service settings. It is also uncertain whether there are trade-offs between duration and depth of data collected from staff or service user-completed measures, i.e. whether respondents would be prepared to complete a lengthier or more complex measure for a limited period of time.

The proposals of this review, that a multi methods approach including using staff and service-user completed data be adopted and measures providing a greater depth of information be developed, would only increase the challenge of retaining feasibility in content of care measurement. Existing measures of content of care have been used to a great extent during research studies rather than in normal clinical practice: it may not be possible to create a measure of content of care which provides sufficient depth of information to be useful but is brief and simple enough to be acceptable for routine use in clinical settings.

Accounting for different perspectives

Few measures identified in this review include any information gathered from patient-report and none exclusively. This seems hard to justify: the experience of care received has as much face validity as a measure of content of care as the perception of care provided. Glick and colleagues most explicitly seek to include different perspectives [22], collecting information about care provided from physician, patient and carer. However, they then seek to reconcile discrepancies between accounts, without reporting how this was achieved. It is not self-evident that differences in the perception of care provided between staff and patients can or should be reconciled. Measures of patients' needs [46], or the style of service [41], for instance, have identified significant differences between the views of staff and patients. Whether there are significant differences in consumers' and providers' perceptions of the content of care in mental health services and whether any such differences are constant in different services remain to be researched.

Conclusion

Measures have been developed which can help describe what happens in mental health services. However, despite identification of the issue a decade ago [13], there remains no consensus about ideal methods or measures of service content. Further research in the following areas could help to establish such a consensus:

- The development of measures which provide greater depth of information about the nature of care provided at services, especially inpatient services.
- More testing of the psychometric properties of measures across a range of service settings.
- More investigation of the feasibility of measures in different service settings, including routine reporting of completion rates in use of process measures in studies.
- The development of measures which include patients' perspective on the content of care at services.

In the absence of established ideal methods and gold standard measures, current measurement of the content of care in mental health services should use a multi-methods approach. Data from a variety of information sources and collection methods can maximise the breadth and depth of information available and, if consistent, increase confidence in its validity. Focus on the nature of interventions provided by services, not just their number or the type of service within which they are provided, can aid description and distinction of mental health services and the goal of understanding service outcomes.

References

1. Abramowitz S, Obten N, Cohen H (1998) Measuring HIV/AIDS Case Management. Soc Work Healthcare 27(3):1–28
2. Beecham J, Knapp M (1992) Costing psychiatric interventions. In: Thornicroft G, Brewin CWJ (eds) Measuring mental health needs, Gaskell, pp 163–183
3. Bond G, Salyers M (2004) Prediction of outcome from the Dartmouth Assertive Community Treatment fidelity Scale. CNS-Spectrums 9(12):937–942
4. Bowers L, Brennan G, Flood C, Lipang M, Oladapo P (2006) Preliminary outcomes of a trial to reduce conflict and containment on acute psychiatric wards: City Nurses. J Psychiatr Mental Health Nursing 13:165–172
5. Bowie P, Mountain G (1993) Using direct observation to record the behaviour of long-stay patients with dementia. Int J Geriatr Psychiatry 8:857–864
6. Brekke J (1987) The model-guided method for monitoring program implementation. Eval Rev 11(3):281–299
7. Brekke J, Long J (1997) The impact of service characteristics on functional outcomes from Community Support Programs for persons with schizophrenia: a growth curve analysis. J Consult Clin Psychol 65(3):464–475
9. Brekke J, Ansel J, Long J, Slade E, Weinstein M (1999) Intensity and continuity of services and functional outcomes in the rehabilitation of persons with schizophrenia. Psychiatr Services 50(2):248–256
9. Brugha T, Glover G (1998) Process and health outcomes: need for clarity in systematic reviews of case management for severe mental disorders. Health Trends 30(3):76–79
10. Brugha TS, Lindsay F (1996) Quality of mental health service care: the forgotten pathway from process to outcome. Social Psychiatry Psychiatr Epidemiol 31(2):89–98
11. Bryman A (2004) Social research methods, 2nd edn. Oxford University Press
12. Burns T, Fiander M, Kent A, Ukoumunne OC, Byford S, Fahy T, Kumar KR (2000) Effects of case-load size on the process of care of patients with severe psychotic illness. Report from the UK700 trial. Br J Psychiatry 177:427–433
13. Burns T, Priebe S (1996) Mental health care systems and their characteristics: A proposal. Acta Psychiatr Scand 94(6):381–385
14. Dean R, Proudfoot R (1993) The Quality of Interactions Schedule (QUIS): Development, Reliability and Use in the Evaluation of Two Domus Units. Int J Geriatr Psychiatry 8:819–826
15. De Jong A (2000) Development of the International Classification of Mental Health Care (ICMHC). Acta Psychiatr Scand 102(Suppl 405):8–13
16. De Jong A, Giel R, Tentlom G (1991) International classification of mental health care: a tool for classifying services providing mental health care Part 1. Dept Social Psychiatry, University of Groningen and World Health organisation Regional office for Europe, Groningen

17. Dietzen L, Bond G (1993) Relationship between case manager contact and outcome for frequently hospitalised psychiatric clients. Hospital Community Psychiatry 44(9):839–843
18. Donabedian A (1980) Explorations in Quality Assessment and Monitoring, Volume 1: the Definition of Quality and Approaches to its Assessment Health Administration Press
19. Falloon I, Fadden G (1995) Integrated mental health care: a comprehensive community-based approach Cambridge University Press
20. Fiander M, Burns T (2000) A Delphi approach to describing service models of community mental health practice. Psychiatr Serv 51(5):656–658
21. Fisher G, Landis D, Clark D (1988) Case management service provision and client change. Commun Mental Health J 24(2):134–142
22. Glick I, Burti L, Suzuki K, Sacks M (1991) Effectiveness in psychiatric care: 1. A cross-national study of the process of treatment and outcomes of major depressive disorder. J Nerv Mental Disease 179(2):55–63
23. Hansson KS, Allebeck P, Malm U (2001) Event recording in psychiatric care: Developing an instrument and 1-year results. J Nordic Psychiatry 55(1):25–31
24. Higgins R, Hurst K, Wistow G (1999) Nursing acute psychiatric patients: a quantitative and qualitative study. J Advanced Nursing 29(1):52–63
25. Johnson S, Salvador-Carulla L (1998) Description and classification of mental health services: a European perspective. European Psychiatry 13:333–341
26. Johnson S, Kuhlman R (2000) The EPCAT Group, "The European Service Mapping Schedule (ESMS): development of an instrument for the description and classification of mental health services". Acta Psychiatr Scand 102:14–23
27. Kitwood T (1997) Evaluating dementia care: the DCM method, 7th edn. Universwity of Bradford
28. Kovess V, Lafleche M (1988) How do teams practice community psychiatry?". Canada's Mental Health 36(2):1–28
29. McGrew J, Bond G, Dietzen L, Salyers M (1994) Measuring the fidelity of implementation of a mental health program model. J Consult Clin Psychol 62:670–678
30. McHugo G, Drake R, Teague G, Xie H (1999) Fidelity to assertive community treatment and client outcomes in the new hampshire dual disorders study. Psychiatr Service 50(6):818–824
31. Mechanic D (1996) Emerging issues in international mental health services research. Psychiatr Services 47(4):371–375
32. Morley S, Snaith P (1992) Principles of psychological assessment. In: Freeman C, Tyrer P (eds) Research methods in psychiatry, 2nd edn. Gaskell, pp 135–152
33. Morse G, Calsyn R, Klinkenberg W, Helminiak T, Wolff N, Drake R, Yonker R, Lama G, Lemming M, McCudden S (2006) Treating homeless clients with severe mental illness and substance use disorders: costs and outcomes. Commun Mental Health J 42(4):377–404
34. Olusina AK, Ohaeri JU, Olatawura MO (2003) The quality of interactions between staff and psychiatric inpatients in a Nigerian general hospital. Psychopathology 36(5):269–275
35. Patmore C, Weaver T (1989) A measure of care. Health Serv J 99(5142):330–331
36. Paul G (1987) Rational operations in residential treatment settings through ongoing assessment of staff and client functioning. In: Peterson D, Fishman D (eds) Assessment for decision, Rutgers University Press, pp 145–203
37. Pawson R, Tilley N (1997) Realistic Evaluation. Sage
38. Popkin M, Callies A, Lurie N, Harman J, Stoner T, Manning W (1997) An instrument to evaluate the process of care in ambulatory settings. Psychiatr Service 48(4):524–527
39. Quirk A, Lelliot P (2001) What do we know about life on acute psychiatric wards in the UK? A review of the research evidence. Soc Sci Med 53(12):1565–1574
40. Resnick S, Neale M, Rosenheck R (2003) Impact of Public Support Payments, intensive pychiatric community care and program fidelity on employment outcomes for people with severe mental illness. J Nerv Mental Disease 191(3):139–144
41. Rossberg J, Friis S (2004) Patients' and staff's perceptions of the psychiatric ward environment. Psychiatr Service 55(7):798–803
42. Rossi P, Freeman H (1989) Evaluation: a systematic approach, 4th edn. Sage
43. Rossi P, Freeman H, Lipsey G (1999) Evaluation: a systematic approach, 6th edn. Sage
44. Salize HJ, Kustner BM, Torres-Gonzalez F, Reinhard I, Estevez JFJ, Rossler W (1999) Needs for care and effectiveness of mental health care provision for schizophrenic patients in two European regions: a comparison between Granada (Spain) and Mannheim (Germany). Acta Psychiatr Scand 100(5):328–334
45. Shepherd G, Richardson R (1979) Organization and interaction in psychiatric day centers. Psychol Med 9:573–579
46. Slade M, Phelan M, Thornicroft G (1998) A comparison of needs assessed by staff and by and epidemiologically representative sample of patients with psychosis. Psychol Med 28(3):543–550
47. Slade M, Thornicroft G, Loftus S, Phelan M, Wykes T (1999) The camberwell assessment of need Gaskell, London
48. Teague G, Bond G, Drake R (1998) Programme fidelity in ACT: development and use of a measure. Am J Orthopsychiatry 68:216–232
49. Thornicroft G, Tansella M (1999) The mental health matrix: a Manual to Improve Services Cambridge University Press
50. Thornton A, Hatton A, Tatham A (2004) Dementia care mapping reconsidered: exploring the reliability and validity of the observational tool. Int J Geriatr Psychiatry 19:718–726
51. Tugwell P (1979) A methodological perspective on process measures of the quality of medical care. Clin Investigat Med 2:113–121
52. Tyson G, Lambert G, Beattie L (1995) The quality of psychiatric nurses' interactions with patients: an observational study. Int J Nursing Studies 32(1):49–58
53. Whitley R, Crawford M (2005) Qualitative research in psychiatry. Can J Psychiatry 50(2):108–114
54. Wing J, Brown G (1970) Institutionalism and Schizophrenia Cambridge University Press
55. Wright R, Sklebar T, Heiman J (1987) Patterns of case management activity in an intensive community support programme. Commun Mental Health J 23(1):53–59
56. Young A, Sullivan G, Burnam A, Brook R (1998) Measuring the quality of outpatient treatment for schizophrenia. Arch Gen Psychiatry 55:611–617