MEASURING THE OUTCOME OF PSYCHIATRIC CARE USING CLINICAL CASE NOTES

Marco Akerman MD, MSc

A report submitted as part of the requirements
for the degree of Doctor of Philosophy
in the Faculty of Medicine, University of London

Department of Epidemiology and Public Health
University College London

1993

ProQuest Number: 10016721

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10016721

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

ABSTRACT

This thesis addresses the question: 'Can case notes be used to measure the outcome of psychiatric treatment?'

Patient's case notes from three psychiatric units (two in-patient and one day hospital) serving Camden, London (N=225) and two units (one in-patient and one day- hospital) in Manchester (N=34) were used to assess outcome information of psychiatric care in terms of availability (are comparable data present for both admission and discharge?), reliability (can this data be reliably extracted?), and validity (are the data true measures?). Differences in outcome between units and between diagnostic groups were considered in order to explore the possibility of auditing the outcome of routine psychiatric treatment using case notes.

Availability: For the Camden patients, 57% of all symptoms ratings made from their case notes were considered available information on admission and 55% of the symptoms reported as present on admission were also reported on discharge and could be assessed for change. However, the proportion of behaviour (8%) and social (32%) items reported in the case notes was very much lower.

Reliability: Two observers compared abstracting for 15 Camden case notes and showed kappa coefficients ranging from 0.65 to 0.85.

Validity: Symptom and syndrome ratings in the Manchester case notes were compared with independent Present State Examination (PSE) ratings and showed overall measures of sensitivity (0.43 and 0.56), of specificity (0.91 and 0.95) and of positive predictive value (0.62 and 0.69), respectively. Scores of change at discharge in the Manchester notes were compared with independent Comprehensive Psychopathological Rating Scale (CPRS) measurements and showed weak association.

The results indicate that routine case notes do not contain complete information to assess the outcome of psychiatric treatment and prospective collection would be needed to improve availability of data.

There was no difference in the average improvement of Camden patients between the three settings when symptom change between admission and discharge were assessed, a finding which is supported by studies in the literature comparing outcome of different sites of treatment.

"The medical record is regarded as an object of awe, much as historians revere the Dead Sea Scrolls or Harvard graduates worship their diplomas. Although it contains the distilled wisdom of past physicians and scholars and the resulting truth of prior tests and studies, these revelations are obscured by strange idioms and are buried in an avalanche of bureaucratic landfill. Like Indiana Jones, the astute physician must take an archaeologic approach to becoming a reader of the lost chart: plunge courageously into these dark tomes in search of the grail of knowledge. To find the grail, the only option is to examine the grisly remains of each chart one-by-one while suffocating in the heavy odour of bureaucracy. The medical chart, mother lode of medical history, contain data beyond the wildest dreams of the most compulsive diagnostician. Although exploring and interpreting the chart is dangerous the rewards are great."

Brancati (1992)

"All retrospective surveys run the risk of error in identification because the investigator is dealing with data collected in an unstandardised way. This is very true for clinical case notes where the phenomena are described by diverse reporters"

Feinstein (1969)

To Suzana

My companion in this odyssey.

ACKNOWLEDGEMENTS

CONTENTS

13

LIST OF TABLES

15

LIST OF MATRICES

18

LIST OF FIGURES

ABBREVIATIONS

BPRS	- Brief Psychiatric Rating Scale
CPRS	- Comprehensive Psychopathological Rating Scale
DH	- Day hospital
FSC	- Final Score of Change
GSC	- General Statement of Change
HPC	- History of present complaint
IA	- Inferred as absent
ICD	- International Classification of Diseases
IMPS	- In-patient Multidimensional Psychiatric Scale
INP	- In-patient
MSE	- Mental State Examination
NN	- Nursing notes
NR	- Not reported
PEF	- Psychiatric Evaluation Form
PPV	- Positive Predictive Value
PSE	- Present State Examination
PSS	- Psychiatric Status Schedule
RA	- Reported as absent
RCT	- Randomised Controlled Trial
RP	- Reported as present

S1	- Setting 1
S2	- Setting 2
S3	- Setting 3
SBAS	- Social Behaviour Assessment Scale
SCL	- Syndrome Check List
SP	- Specificity
SPSS	- Statistical Package for Social Science
SS	- Sensitivity
UCH	- University College Hospital
WHO/ DAS	- World Health Organization / Disability Assessment Scale

CHAPTER 1

INTRODUCTION AND
BACKGROUND LITERATURE

1.1. Introduction

Mental Health is one of the key areas of England's *The Health of the Nation*

document (Secretary of State for Health, 1992) and contributes to a large extent

to increase the burdens of mortality, morbidity and cost of care. Each year, 6

million people in the U.K. suffer from mental illness and it is estimated that

20,000 individuals die each year as a result of mental illness; four times as many

as die in road accidents (Mental Health Foundation, 1989). In terms of costs,

psychological problems account for 14% of days lost from work, 20% of the total

NHS expenditure, 23% of in-patient costs and 25% of pharmaceutical charges. It

is estimated a total annual spending on mental illness at £4.6 to £5.6 billion

(Thornicroft and Strathdee, 1991).

Such information suggests that mental health should be investigated more fully,

demanding knowledge about how effectively services and programmes are dealing

with this problem. In other words, to assess whether the resources are

appropriately deployed to achieve optimal benefits. This has led to the recognition

of **evaluative research** in the mental health field as a welcome line of enquiry

(Sartorius and Harding, 1983).

Evaluation research in general is defined by Rossi and Freeman (1989) as "the systematic application of social research procedures for assessing the conceptualization, design, implementation and utility of social intervention programmes". In particular, in health services research, it has been understood as the formal determination of the effectiveness, efficiency, and acceptability of a planned service or programme (Cochrane, 1971; Warren, 1979; Holland, 1983). Evaluation is not distinct from any other scientific approach except that its results and methods may sometimes be immediately useful to clinicians and administrators (Wing, 1972).

Donabedian (1966) suggested the approaches to assessment of a health service or a programme under three headings. The **"structure"** describes "the setting in which the intervention takes place and the instrumentalities of which it is the product". The **"process"** represents the activities triggered by any patient who enters the health care system. And finally, the **"outcome"** is the impact that the service or programme has produced on the health of the patient and on the health status of the community.

The fact that outcome evaluation is a demanding, expensive and risky procedure has led to health service process (resources consumed) being used as a proxy measure for outcomes, when a clear relationship between them has been

established (Holland, 1983; Shaw, 1989, Black, 1990; Jenkins, 1990). Measuring

outcomes is also inhibited by lack of consensus on which are the actual objectives

of the care provided. However, it has been said that outcome of care by and large

"remains the ultimate validator of the effectiveness and quality of care"

(Donabedian, 1977) and that most politicians, patients, doctors and managers

would prefer health services to be evaluated by the results achieved rather than by

the resources consumed (Shaw, 1989).

There is a general agreement that outcome in the psychiatric field is

multidimensional (Fontana and Dowds, 1975) resulting in the existence of many

indicators to measure final results of treatment (Avison and Speechley, 1987).

Measuring psychiatric outcome is very difficult (Garden, 1989). The concept of

"cure" in psychiatry is elusive and treatment results may, in many instances, reflect

the characteristics of the patient and of the illness as much as they do of the care

provided (Zusman, 1989). For example, variables like severity of the illness, age

of patient and social factors may influence the final outcome. These risk factors

must be controlled in any outcome assessment in order to   determine how much

of the outcome is explained by the processes of care and how much is due to risk

factors. Large cohort studies examining diagnostic groups could help to answer the

question which patient benefit most from which treatment.

Fauman (1989) believes that even if there is no consensus on standards auditing

psychiatric care is still possible, if quality of care is not equated with complete

success. Moreover, standards are locally decided and agreed between management and professionals (Department of Health, 1989). Zusman (1988) recommends outcome measurement as the gold standard in the evaluation of health care, and Jenkins (1990) claims that a "system of outcome indicators for mental health care is urgently needed".

The present study sets out to propose and test a method of assessing outcome of psychiatric treatment for acutely ill patients admitted to day hospital and in-patient units. Acutely disturbed patients form the bulk of the treated population in mental health services and day and in-patient services are the common modalities of treatment in every District in England and Wales.

Differently from previous outcome studies that based their assumptions on primary collection of data, this thesis aims to make maximum use of existing clinical information. Given that case notes are the most common instrument of the routine clinical contact, they have been adopted as the source of outcome data for this research. That case notes are a suitable instrument is still a matter for investigation as this subject has not yet been well explored (Siegel and Fischer, 1981: 8-9).

Outcome data collected from case notes may also be used to audit psychiatric care. Indeed case note review could be the best initial audit package to be adopted by health districts and clinicians (McKee, 1989; Garden, 1989; Glover,1990; Heath, 1990; Shaw, 1990; Gulliford et al, 1991; Hopkins, 1991; Robinson, 1991).

Clements (1992) points out the "crucial role" of medical records emphasizing that good and accessible notes are the foundation for medical audit.

This research has a "problem-solving strategy" (Cahn and Richman, 1985), starting from a particular research problem in the real world - how to measure outcome of psychiatric treatment - and bringing together all the intellectual resources that can be brought to bear on its solutions. "The problem has to be defined and the method of solution has to be discovered" (Philips and Pugh, 1987: 45).

1.2. An overview of the thesis

First, an overview of the evolution of psychiatric treatment and research on outcome is given. We concentrate the review on studies that assessed day hospitals versus in-patient units as they are the modalities assessed by this present study. In addition, these modalities form the bulk of studies on service effectiveness presented by the mental health literature. Studies that used clinical case notes (medical records) as their main data source to assess different characteristics of medical treatment as well as studies on instruments of psychiatric research are also reviewed. Later, medical audit papers published mainly after 1989 when the White Paper *Working For Patients* (Department of Health, 1989) came out, are examined in order to see if the proposed method can enhance the current process of auditing

psychiatric care.

Next, chapter 2 presents the aims, objectives and hypotheses of the thesis followed by the main body of the present study, chapters 3, 4 and 5. These chapters, where the methods and results are presented and discussed, are based on the framework given by the main attributes of the outcome information tested in this study - availability, reliability, and validity.

Finally, we conclude the study, chapter 6, with its main findings and the potential applicability of the proposed method to the current process of medical audit.

1.3. Background literature

1.3.1. Sources

The studies examined were initially selected from Compact Cambridge: Medline, 1983-92, and from the relevant reviews and books on the subject: (1) for outcome studies - Test and Stein, 1978; Braun et al, 1981; Greene and De la Cruz, 1981; Wallace and Hass, 1983; Talbot, 1985; Pang, 1985; Schene and Gersons, 1986; Rosie, 1987; Avison and Speecheley, 1987; Goldstein and Horgan, 1988 and Creed et al, 1989b; (2) for case notes - Feinstein et al, 1969; Siegel and Fischer, 1981; (3) for the instruments of psychiatric research - Freeman and Tyrer, 1989;

Thompson, 1989; (4) for medical audit - Crozier & Spiby, 1989; Mckee et al,

1989; Shaw, 1990; Hopkins, 1991. The King's Fund Centre Information Resources

were also consulted as a complementary source of bibliographical references.

1.3.2. Evolution of psychiatric treatment and research on outcome

The philosophy of psychiatric treatment has changed over the last three decades.

The major aspect of this transformation has been the gradual replacement of huge

psychiatric hospitals, the basic modality of treatment, by a wide range of

community based services. Several interrelated factors[1] have been put forward to

explain this trend, and there has been a great deal of debate about the relative

contribution of each factor (Test and Stein, 1978). These factors have had a

profound influence on the mental health research area and, in particular, in the

search for new alternatives. The aim being to reduce institutional dependency and

to be less disruptive to patients' lives thereby minimising stigmatisation and

reducing costs.

---

[1]These factors are: (1) the discovery of neuroleptic drugs in the middle 1950s giving to professionals and relatives greater sense of control and increasing administrative flexibility; (2) the post-World War II explosion in the number of mental health professionals of all disciplines; (3) the civil rights and civil liberties movements which extended their claims for the rights of the mentally ill demanding better conditions of treatment, specifically the conception of treatment near home; (4) the explosion of health costs which put pressure on psychiatric hospitalisation, a major share of mental health costs (Mechanic and Aiken, 1987; Mosher, 1983; Test and Stein, 1978).

In the 1950s, the research was still concentrated on the hospital (Shepherd, 1957; Norris, 1959; Rice et al, 1961). In the 1960s, studies seeking alternatives to continued long-term hospitalisation were carried out (Wing, 1960; Brown, 1966) and sociological research on the negative effects of fully-fledged psychiatric hospitalisation (Goffman, 1961) gave impetus to community care in the 1970s and in the first half of the 1980s.

Studies in community care (Pasamanick et al, 1967; Grad and Sainsbury, 1966; Mosher et al, 1975; Polack and Kirby, 1976; Stein et al, 1975; Test and Stein, 1978; Fenton et al, 1979; Stein and Test, 1980; Hafner and Klugh, 1982; Wing, 1982; Tansella, 1986) concluded that community care generated better results than the treatment provided by standard hospitals just so long as it is comprehensive and ongoing (Stein and Test, 1980). Mosher (1983) suggested that comprehensive community care has not been adopted either because of professional or public vested interests. He is contradicted by Tantam (1985) who raised methodological problems in community care studies in order to explain the slow rate of adoption of this alternative. Creed et al (1989b) pointed out that these studies cannot be generalised to the acutely ill since they include a high proportion of patients with chronic illness.

Still on the question of traditional hospitalisation, Herz et al (1975,1977), Glick et al (1974, 1976a, 1976b, 1977), Hargreaves et al (1977), Hirsch et al (1979) and Knights et al (1980) compared the traditional longer hospitalisation (3 to 4

months) with brief admission (1 to 4 weeks). For non-schizophrenic patients, Glick and Hargreaves found no differences in outcome on any measure at discharge, after one year and two years. However, for schizophrenic patients with a good pre-admission level of functioning, they showed that at discharge and after one year the long-term group had fewer symptoms. This made them to recommend long-term hospitalisation for some schizophrenic patients. Herz et al, however, failed to show any differences between the two groups, irrespective of diagnostic, and concluded that brief admission is preferable for the majority of patients. Neither did Hirsch et al (1979) and Knights et al (1980) show any difference between the groups on clinical measures, but clinical outcome was not related to psychiatric diagnosis.

As regards deinstitutionalisation - the contraction of traditional institutional settings, with the concurrent expansion of community-based services (Bacharah, 1976) - its consequences are considered an important research topic in the 1980s. Schene et al (1992) point out that there have been few studies in this subject. One of these studies, started in 1985 by TAPS (Team for the Assessment of Psychiatric Service - North East Thames Regional Health Authority), is unique in its scale and detail. This project, an ongoing research, is an attempt to answer the question if the move from hospital to community care has proved better for long-term patients. Despite some problems - inertia of the planning programme, the selection of the best patients for discharge, accumulation of 'new long stay patients', and neighbourhood opposition to local services for the deinstitutionalised patients, the

programme is "going reasonable well" (Leff, 1992).

1.3.3. Day hospital care versus in-patient treatment

1.3.3.1. Early Studies

Studies on the feasibility and efficacy of day hospital treatment have been done because of the growth of units providing this kind of care "became astronomic" (Pang, 1985). Early studies comparing day hospital with in-patient care (Smith and Cross, 1957; Craft, 1958; Lystad, 1958; Kris, 1965) found that day treatment was as effective as or better than in-patient treatment. However, Herz et al (1971) criticized these studies for using few outcome measures, limited types of diagnoses, small sample size and no use of standardized instruments of measurement.

In one such study, Wilder et al (1966) randomly allocated patients to either day treatment or in-patient care and found no significant differences in outcome between the two groups, although economic measures and patient satisfaction favoured day treatment. Herz et al (1971) argued that Wilder and Zwerling had demonstrated that day treatment *could* be an alternative to in-patient care for a large number of acutely disturbed patients, but their study did not address the question of whether day treatment *should* replace in-patient care. Herz's own

study addressed this question directly and he found better outcomes for day patients at short term, and yet considerable selection of patients prevented him from asking the question raised in the first place. Hogarty et al (1968) in an observational study concluded that day hospital population was not representative of in-patient admissions, thus not representing a true alternative to in-patient hospitalisation.

1.3.3.2. More recent studies: expanding the use of outcome measures

In connection with this area of study Washburn et al (1976), Fink et al (1978), Penk et al (1978), Dick et al (1985) and Creed et al (1990) did not find any long term difference in symptomatology when they compared patients treated by day or in-patient care, despite the differences in their study population in terms of diagnosis and sex. For psychosocial measures, the findings tended to favour day treatment. For example, Potomianos et al (1986) found a reduction of alcohol intake for day patients. And Herz et al (1971), Michaux et al (1973), Washburn et al (1976) and Penk et al (1978) found better long term social adjustment within day patients. On the other hand, Creed et al (1990) observed better social role performance for in-patients at 3 months but no difference at 12 month follow up. Although, Dick et al (1985) showed that patient satisfaction was greater among day patients, Tantam (1985) argued that this greater satisfaction may have been generated by staff enthusiasm with the new venture. Test and Stein (1978)

32

suggested that perhaps community options showed better results for psychosocial

measures due to the fact that they concentrated their efforts on managing mainly

these aspects of patient functioning.

Service utilisation rate was also used as an outcome measure. Sappington and

Michaux (1975) and Kecmanovic (1985) did not find any significant difference for

day and in-patient samples using re-admission rate as their main outcome measure,

but Edwards et al (1979) showed lower rates for day patients.

Some studies also looked into the financial aspect of in-patient care versus day

treatment. Wilder et al (1966), Washburn et al (1976), Fink et al (1978), Guillete

et al (1978) and Dick et al (1985) found day treatment more cost- effective than

in-patient care. By contrast, Creed et al (1990) pointed out that day care while

feasible for some acutely ill patients, would require a well staffed day unit and a

relatively long admission. This kind of facility is unlikely to be cheaper, but the

reduction in re-admissions would compensate, making day care an attractive form

of treatment.

However, because of (1) different types of research design (Test and Stein, 1978),

(2) very restricted entry criteria (Braun et al, 1981), (3) different measures of

outcome (Creed et al, 1989b), (4) different rating scales (Test and Stein, 1978;

Snaith, 1981), (5) different characteristics of treated population, (6) many patients

lost to follow-up (Creed et al, 1989b), (7) different findings (8) failure to caution

the reader about the possible circumstantial aspect of the findings (Greene and De La Cruz, 1981) and (9) 'successful' treatments not sustained over time, it has not been possible to demonstrate which is the most 'efficacious' alternative.

1.3.3.3. Assessing effectiveness

The field literature presents a heated debate on the issue of effectiveness. Guy et al (1969) observed that research assessing different modalities of psychiatric care is not a matter of comparing an effective treatment with a relatively ineffective one but an exercise in comparing two potent treatments and searching for specific effects which would help to discriminate between the treatments. Michaux et al (1973), comparing psychiatric day treatment with full hospitalisation, concluded that each option was effective in different and potentially complementary areas. Tansella (1989) has suggested that the outcome of community care might depend more upon the way in which the various units of the system are interrelated rather than on the efficacy of each unit separately. Rosie (1988) states that day hospital treatment is a well established, effective and economical alternative for a large number of acutely ill patients and adds that "further research on the general validity of day hospitals would be superfluous". Braun et al (1981), Vaughan (1983), Wilkinson (1984) and Creed et al (1989b) disagree with this idea on the grounds that methodological weaknesses in day treatment research would not support this conclusion.

Moreover, effectiveness is a dynamic concept and health services are not static. They tend to change their structure in reaction to historical transformations and even to changes in other services. For instance, the in-patient and day units studied by Wilder et al in 1966 are probably very different from the in-patient and day units in Creed's 1990 study. Therefore, we have to interpret cautiously their similar results of no significant difference in outcome between the units.

1.3.4. Research on outcome: alternative approaches

There are two different approaches that research on outcome can take. Chalmers (1982), in his book *What is this thing called Science?*, presents Lakato's theory on research programmes, which are structures that contain fairly clear clues and prescriptions as to how they should be developed and extended (Lakatos, 1974).

We may argue that there are two different research programmes with two different prescriptions to advance the knowledge of the study of psychiatric outcomes. Lawrence et al (1991) divide them into experimental and quasi-experimental studies.

The experimental approach can be concerned with decisions about continuation and funding of programmes. It is policy directed and seeks to draw conclusions from a representative selection of patients and doctors. The quasi-experimental

method is usually concerned with improving and modifying local programmes, it is clinically directed and seeks to draw scientific inferences from the performance of individual doctors and services (Keenan, 1975; Burton, 1986; Fitz-Gibson and Morris, 1987; Russel and Wilson, 1992).

The first approach (randomised control trial) makes outcome inferences based upon different measurements over time. However, the lack of control over important variables like, for example, treatment between discharge and follow up, life events or social support, might interfere with measurement made after discharge giving a false impression of the service being assessed (Donabedian, 1977; Creed, 1989b; Fauman, 1989).

The second approach (before and after comparison) carries out measurement while the patient is in direct contact with the service and arrives at its outcome conclusions based upon measures taken between admission and discharge. Therefore, outcome inferences are based on a shorter period of time but less susceptible to the interference of unrelated treatment variables. The RAND publications are a good example of this kind of approach (Rubenstein, 1991).

The use of case notes to assess the outcome of service is a quasi-experimental study. Before and after clinical measurements are used to draw scientific inferences about the performance of individual doctors and services and to improve and modify local programmes.

1.3.5. Clinical case notes as a source of process and outcome information

1.3.5.1. Background

The main purpose of formal clinical case notes is to create a mechanism of continuity of health care by establishing means of communication between professionals (Thompson and Handelman, 1978, 82-83). Studies, particularly, from the U.S.A., present clinical notes in a continuously evolving process where "they have gone from a simple collection of physician observations, reflections, comments and recollections to a complex hospital medicolegal documentation system" (Soreff et al, 1990). Increasingly, case notes are being recognised by clinicians as evidence for accountability rather than mere instruments of communication between professionals (Vort and Mattson, 1989). In addition, the recent act to make health records accessible to patients (The Access to Health Records Act, 1990) has put clinical case notes in the spot light. This may make health professionals become more cautious when writing information (McShane et al, 1992)

1.3.5.2. Assessing the quality of case notes

Studies assessing the quality of the notekeeping process are well documented (Kiernan, 1976; Cunningham, 1991; Markantonakis and Weir, 1991; Barry and

Marsh, 1992; Gabbay and Layton, 1992). But to assess the quality of health care given by simply reviewing the inadequacies of the record can lead to error (Lapatra, 1975, 175-177), because there is no clear evidence that poor notekeeping means that satisfactory care has not been provided (McGuire, 1985; Edwards et al, 1987; Black and Creed, 1988). On the contrary, Lyons and Payne (1974), Kutcher et al (1985), Jacoboson et al (1987) and Mark et al (1991) believe that better record keeping improves patient management leading to higher quality care.

Other studies go one step further and propose strategies of intervention to improve the quality of such notes (Liptzin, 1974; Casper, 1987; Ellis et al, 1987; Vort and Mattson, 1989; Mark et al, 1991) but there are no institutional incentives for clinicians to provide organised and standardised clinical notes (Casper, 1987). On the other hand there are some reasons to believe in the importance of well-kept case notes. Firstly, good case notes can provide support when psychiatrists are called upon to justify their actions in a medicolegal setting (Gelder et al, 1989). Secondly, they have been important in historical research, when the effects of booms, slumps, unemployment and wars have been associated with the incidence, nature and prognosis of psychiatric illnesses (Kreitman, 1975). Thirdly, there are also some advantages of using case notes for teaching purposes (Taintor, 1981). Finally, given the extent to which medical audit relies on data recorded in the case notes, this will increase the pressure to improve the notekeeping process (Ellis, 1991).

1.3.5.3. Process and outcome studies using case notes

Studies using case notes as the main source of data to study aspects of care in different specialties are also documented. Such studies have looked at the predictors of length of day hospital attendance (McGrath and Tantam, 1987), at the assessment of self-poisoning patients by psychiatrists and junior medical staff (Black and Creed, 1988), at the patterns of attendance of child psychiatry (Stern et al, 1990), at the management of acute stroke (Williams et al, 1990), at changing disease patterns in patients with AIDS (Peters et al, 1991) and at the diagnosis of delirium (Johnson et al, 1992). These studies, however, restricted their assessment to the process of care rather than to outcome.

By linking process to outcome using a "structured implicit review method" [2], Rubenstein et al (1991) found convincing evidence that poor care of patients with congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident or hip fracture is associated with an increased frequency of bad outcomes (death rates). They showed that 17% with very good care died compared with 30% with very poor care.

Attempts to use case notes as the source of outcome information in the psychiatric field are scarce. Csernansky et al (1983) used the case notes of 20 schizophrenic

---

[2]Reviewers use their own unspecified criteria to judge care and to analyse specified parts of the medical records and to assess their quality using a rating scale.

patients to generate appropriate items for a retrospective chart review to assess treatment response to neuroleptics. With these items they constructed a Treatment Response Scale (TRS) and tested its concurrent validity and reliability. This lead them to conclude that "there is a surprising validity in assessing ordinary patient records". Curran and Pullen (1990) set out to assess a psychiatric liaison service and gain a measure of patient outcome from the use of GP case notes. They found that 60% of the notes contained some indications of patient outcome (deterioration, no change and improvement) but they judged GP case notes an unsatisfactory means of assessing patient outcome.

Feinstein et al (1969), in a comprehensive study on epidemiology of cancer, discuss collecting information from cancer patients' case notes. They start by discussing the validity of carrying out retrospective studies using case notes because "many clinical investigators have developed a fear that research done by retrospective surveys will produce scientific errors". They go on to say that all retrospective surveys run the risk of error in identification because the investigator is dealing with data collected in an unstandardised way. This is very true for clinical case notes where the phenomena are described by diverse reporters. However, if retrospection is devoted to prognosis and the study of the results of treatment (therapy studies), then errors of logic, which would occur in etiologic studies, do not apply (Feinstein et al, 1969).

Siegel and Fischer (1981) in a edited collection of papers discussing mental health

records, gave impetus to case note research and suggested that the only way to improve the quality of clinical records is to study them systematically. They did not agree with the complaint that clinical records are unscientific. Their findings support the idea that records reflect the care process so that this kind of complaint is a disguised criticism of the care process as being too intuitive and quasi-scientific. They see non-standardisation not as a characteristic of the psychiatric records themselves but rather as a by-product of the lack of standardisation in psychiatry. More generally, even when carefully recorded data are used, there may be shortcomings in drawing rigorous inferences and conclusions.

The structure of case notes in the health care field is sometimes assumed to be "an inevitable fact of life and therefore unalterable" (Tait, 1981). Historical accounts show that the quality of these records began to improve only when they were used as a tool to monitor the quality of care (Siegel and Fischer, 1981). It is paradoxical though, that proposals to assess the quality of care in the United States have come about as a result of administrative action on the part of the government and private insurers due to financial conditions rather than from firm action of clinicians interested in change (Grant, 1981). But Siegel and Fischer (1981) warn that any attempt to restructure records will fail unless clinicians' needs are considered because "people learn how to beat the system" if they are dissatisfied with new modifications. They conclude that any proposal to use and change case notes should be in support of clinical activities (see also Vort and Mattson, 1989).

1.3.5.4. Mental state examination (MSE): an important clinical action

The mental state examination is concerned with the symptoms and behaviour at the time of the psychiatric interview (Gelder et al, 1989). There should be some standardisation on which items of mental state should be recorded since textbooks present similar frameworks on how to report them (Hill et al, 1973; Gelder et al, 1989; Kaplan and Sadock, 1989). From the psychiatric interview clinicians obtain information which is important not only for diagnosis but also for assessing the course of a disorder and its subsequent response to treatment (Leon et al, 1989). The mental state examination is the section of the case notes which contains information capable of showing some degree of change during the course of treatment. In other words, it is the major section where the results of clinical action can be demonstrated.

Psychiatric treatment can involve a comprehensive set of long-term interventions and positive results are often not seen in a short period. If this is the case, a hospital admission is likely to represent only one segment of the whole process of treatment. However, Fontana and Dowds (1975) found that the substantial drop in symptomatology after admission to 1 month hospitalisation is maintained up to the 6-month point when the last assessment was made. If this indicates an association between clinical practice and outcome, measuring changes within the course of a hospitalisation might represent a proxy outcome measure for the whole process of the treatment (Zusman, 1988; Fauman, 1989; Fish, 1989; Rubenstein

et al, 1991). Therefore to assess change we need two different measurement points within the course of the treatment, a mental state at admission and one at discharge and we need to set standards on what is expected regarding reduction of symptoms between these two points.

1.3.5.5. Mental state examination and the real world

Ellis et al (1987) sent questionnaires to six consultants and nine junior doctors working in a psychiatric unit, asking them to rate the importance of 263 items in a case history and asking which of them should be recorded in detail by the end of the third day after admission. All the consultants and 92% of junior doctors agreed that the items of the mental state examination should be included. These figures can be contrasted with just 15% and 18% for the 'physical examination' items, 25% and 56% for 'presenting problems' items, and also 37.5% and 15% for 'diagnosis and formulation', respectively for consultants and junior staff. An educational programme was set up and the authors assessed three different sets of case notes in three different periods of time in order to review the degree of compliance with the case history outline developed. The mental state examination was the only section that showed significant rates of compliance and improvement on all three points of assessment. Recently, Craddock and Craddock (1990) examined 100 discharge summaries of patients from a psychiatric hospital and found that 71% of the summaries contained an admission mental state whereas

only 25% contained a 'physical examination' and only 42% a 'past psychiatric history'.

These findings let us think that the mental state examination possesses some attributes (consensus amongst clinicians of its importance, high frequency of reporting and susceptibility for improvement after an education process) that make it suitable to be the first point of rating.

The next step would be to rate symptom change between admission and discharge by looking at the details of the mental state at discharge. However, it surprisingly appears not to be routine practice to report the mental state at discharge: Craddock and Craddock (1990) found that only 26% of the discharge summaries presented the discharge mental state.

It is not apparent why the mental state examination at discharge is less frequently recorded. One could speculate on some reasons without empirical evidence to support them. First of all, classical psychiatric textbooks do not draw attention to the necessity of recording a mental state at discharge; second, legal procedures require the mental state at admission but do not demand the mental state at discharge; third, clinicians are more interested in the diagnostic information provided by the mental state at admission than they are in assessing change between admission and discharge; and finally, the effects of psychiatric treatment over a short period of time may be too small that this would not encourage

clinicians to record them.

1.3.6. The instruments of psychiatric research

1.3.6.1. General concepts

The instruments used to collect psychiatric information provide the means of quantifying aspects of the patient's symptoms, behaviour, social interactions for research, clinical care and teaching purposes (Grebb, 1989; Thompson, 1989). These patient attributes can be assessed using three distinct types of measurement (John et al, 1987): (1) measures of subjective wellbeing assessing happiness and satisfaction with life in quality of life research (such as The Nottingham Health Profile and The Sickness Impact Profile as cited in Bowling, 1991), (2) social and behavioral adjustment scales used to measure social impairment in clinical and community psychiatric research (Platt et al, 1980; Wykes and Sturt, 1986; WHO/Disability Assessment Scale, 1988) and, (3) psychiatric screening instruments to detect or estimate psychiatric disturbance in population and clinical studies (Goldberg, 1970; Wing et al, 1974; Edincott and Spitzer, 1978). The information needed to fill in these items, according to the requirements of each instrument, may be obtained by self-report (patient-rated) scales, informant-rated scales, professional-rated naturalistic observations, professional-rated interviews (Grebb, 1989) or even medical and other notes (Spitzer et al, 1967).

The collected data should be interpreted with caution because complex events are being measured and "diseases are names for theories rather names for things" (Wing, 1978). In this way we may overcome the dilemma of measuring phenomena that we do not fully comprehend as yet (Thompson, 1989).

1.3.6.2. Collecting symptoms

Spitzer et al (1967) present a list of 6 instruments for evaluating psychiatric status - Mental Status Schedule (MSS), Psychiatric Status Schedule (PSS), Psychiatric Evaluation Form (PEF), Psychiatric History Schedule (PHS), Social Background Record (SBR) and Diagnosis Recording Form (DRF). Four out of the six (PEF, PHS, SBR, DRF) could be used to obtain information from clinical case notes. Of these, two are recording forms (SBR and DRF) and one is designed to record aspects of the severity of psychiatric illness during one patient's lifetime (PHS).

Thompson (1989) and Ferguson and Tyrer (1989) in an attempt to "find a way through the rating scale-maze" present a comprehensive list of the main instruments used to collect symptom information. They divide them into (1) the detection of psychiatric morbidity, (2) psychiatric classification and diagnosis and (3) rating scales for measuring symptom change under specific conditions, for instance, schizophrenia, affective disorders and anxiety. Wallace and Hass (1983) point out that to recommend one instrument in particular is rather difficult, as the information required and the resources available to obtain it vary from situation

to situation and Snaith (1991) noted that the choice of an instrument is always an arbitrary decision.

Regarding the present study, the Psychiatric Evaluation Form - PEF (Spitzer et al, 1967) would be a potential choice for inclusion as it has a section to rate subjects who have been admitted to a psychiatric facility, and also a scale to measure role impairment. But, the fact that this instrument was devised in the United States 23 years ago would impair access to manuals and training courses.

The Present State Examination - PSE was intended as a clinical interview not as a 'questionnaire' to obtain degree of standardization in content and form in the diagnostic process in order to increase reproducibility (Wing, 1974;Thompson, 1989). The PSE has been used in more recent studies as an instrument to measure severity, including Hoult, 1986, Tyrer et al, 1987; Creed, 1989a; Creed, 1990, and this may facilitate comparison of results. The PSE includes a section, the Syndrome Check List - SCL, devised to collect information from case notes and other written reports.

The coding instrument devised to collect symptom data for the purpose of this thesis is based upon the SCL. The fact that case notes have been selected as the source of data restrains the choice of instrument, as the majority of instruments available have been devised to collect information directly from patients and relatives and not from case notes.

47

There are few studies that describe data collection using the SCL. McGuffin et al (1986) in a family-genetic study of affective illness used the SCL to retrospectively collect information from case notes and then compared their ratings with PSE interviews carried out blindly and independently from the SCL ratings. Agreement between the SCL and interview diagnosis for the 'worst episode' was satisfactory with agreement of 89% and weighted kappa of 0.61.

Johnstone et al (1981) studied institutionalisation and the handicaps of schizophrenic patients after 5-9 years they had been discharged from hospital. They were assessed in terms of mental state and cognitive, behavioral and neurological functioning. The results of these assessments were related to information obtained from case notes using the SCL. Patients with diagnosis of schizophrenia or schizoaffective disorder were studied 2-20 years since onset of illness by Kolakowska et al (1985). They rated from the patients' case notes the presence of psychiatric symptoms at early stages of the illness using the SCL in order to reconstruct the full clinical history of each patient. In both studies the SCL is only a supplementary instrument.

1.3.6.3. Collecting behaviour and social information

Behaviour and social functioning are two attributes commonly measured in social psychiatry research in general (Wing, 1989) and in outcome studies in particular

(Platt, 1980). They express the interplay between individuals and their social environment (Weissman, 1975). Although there is some debate about the independence of symptoms from social functioning, behaviour and social role are measured independently from symptoms by the scales concerned.

The scales reviewed by Weissman (1975) and Wing (1989) are mainly follow-up instruments intended to measure community interactions after discharge. While symptom remission may be relatively quicker, the effects of treatment on social and behavioural aspects are only noticed later (Herz et al, 1971; Michaux et al, 1973). Many social adjustment scales count on 'significant others' to report and case notes are not very common sources of this kind of information. In addition, for hospital admissions it is not common practice to report on the behavioural dimensions independently from symptomatology, and social aspects do not tend to be reported in a comprehensive way. As a result, there are no relevant studies which report on the collection of these items from case notes.

Bearing these limitations in mind the present research explores the collection of behavioral and social items from case notes. Items from the section 'Ward Behaviour' of the WHO/Disability Assessment Scale - DAS (1988) and the main social topics of the scales reviewed by Weissman (1975) and Anthony and Cairns (1978) are put together to form the coding instrument for this study. The DAS was specially designed for use in a collaborative study (WHO, 1988) of patients with severe psychiatric disorders. The main objective in planning the scale was to

produce an instrument compatible with the PSE.

1.3.6.4. Assessing symptom change

The assessment of symptom change over short periods of time is essential for the evaluation of any psychiatric treatment (Tress et al, 1987). There are several instruments devised to assess change, but many are cumbersome and relatively insensitive (Ferguson and Tyrer, 1989). Tress et al (1987) present five characteristics that a rating instrument should embrace in order to be able to assess symptom change: (1) to generate data for clinical classification, (2) to be sensitive to change in symptomatology, (3) to contain mutually exclusive symptoms with clear definitions, (4) to use a standardised clinical interview and, (5) to be acceptable both to interviewers and subjects.

A review by Manchanda and Hirsch (1986) of the most common scales in use showed that only four were appropriate for all psychiatric diagnoses. The Brief Psychiatric Rating Scale (BPRS) (Overall and Gorhan, 1962) and the In-patient Multidimensional Psychiatric Scale (IMPS) satisfy only characteristics 2 and 5 above.

The Comprehensive Psychopathological Rating Scale (CPRS) (Asberg et al, 1978) meets characteristics 2,3 and 5. As a sensitive instrument to measure change it

may be particularly useful in the evaluation of new psychotropic compounds (Thompson, 1989). The scale is recommended for use by trained professionals from different fields and reliability studies between different disciplines showed satisfactory coefficients of agreement (Montgomery et al, 1978)

Although the PSE satisfies characteristics 1,3,4, and 5, its 3-point scale for most symptoms was found to be insensitive to symptom change in comparison with the 7-point scale of the BPRS.

Based on the PSE, Tress et al (1987) addressed this problem and devised a new PSE-Change Rating Scale with an 8-point scale. To begin with, they administered the original PSE interview to patients in order to determine a baseline for the ratings. Subsequent ratings were made only on positively rated symptoms using the PSE - Change Rating Scale, the other items were discarded. They concluded that the PSE - Change Rating Scale is sensitive to change, easy to use and reliable for all diagnostic groups.

1.3.7. Medical Audit

The White Paper *Working for Patients - Working Paper 6* (Department of Health, 1989) sets the scene for regular and systematic assessment of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcome and quality of life for the patient. It states that every consultant should participate in a form of medical audit analyzing critically the care given to their patients. Standards are locally decided and agreed between management and professionals. Audit is concerned with change and improvement through a cyclical process of practice observation (Mitchell and Fowkes, 1985, Black, 1989a). It is the measurement of the actual quality of care against criteria or standards of excellence (Vuori and Roger, 1989).

After "fears and loathing on the White Paper trail" (Holman, 1989) medical audit is seen now as the least controversial proposal of the recent NHS reforms (Bowden and Walshe, 1991), despite disagreement between managers and clinicians in their assessment of the time required and the opportunity costs (Smith et al, 1992).

There is a debate on which method should be adopted and implemented, but "no single method is both robust to criticism and simple to implement" (Robinson, 1989). In Hopkins'(1991) view, the best audit package that a health district can currently implement maybe a regular case note review by an assessor. Ellis and

Sensky (1991) see analysis of routinely collected health service data - "basic clinical audit" - and "incident review" as the essential audit activities of any clinical department. Shaw (1990) points out that criterion based audit "offers a realistic method of audit", but Bowden and Walshe (1991) states that audit will really start to count when it is able to prove service effectiveness in terms of outcomes.

In mental health, the debate on which audit method should be adopted is also present and psychiatrists have proposed needs-based audit as an alternative approach. (Royal College of Psychiatrists, 1989; Gath, 1989). The College also supports audit projects in the administration of ECT, reasons for admissions and discharge plans in two London districts and epidemiological estimate of the needs of the mentally ill population of these two districts (Lelliott, 1992). Independent projects have focused on the quality of the notes (Cunningham, 1991; Markantonakis and Weir, 1991), on therapeutic drug monitoring (Fenton et al, 1990), on rapid readmission of patients (Jones, 1991), on compulsory detention (Porgourides et al, 1992), on monitoring behavioral psychotherapy (McDonald et al, 1988) and on consumer attitudes to the care received (Ballard, 1990).

Even though outcome audit is considered the most valid type of medical audit (Robinson, 1991), there is still paucity in the development of such type of audit in the psychiatric literature. As outcome measures in psychiatry are complicated (Roy, 1991), we have to invest more in methodological research to develop cost

effective methods of monitoring quality of care (Black, 1992; Frater and Costain 1992). The present study is an attempt to follow this advice.

The framework of this thesis meets one important precept of medical audit, namely the "use of explicit criteria for measurement rather than implicit judgment" (Shaw, 1990). It also uses the existing audit strategy of retrospectively assessing case notes (Hopkins, 1991), but goes one step further by providing proxy measures for outcome, which are better indicators of the quality of care.

1.3.8. Concluding remarks

The outcome of psychiatric care is still an open area for study. Findings from previous studies conflict and the approach of seeking the most effective alternative seems reductionist when one considers the current complexity of the mental health field.

We propose the adoption of a more clinically orientated research approach and we regard case notes as the most common instrument of the clinical contact. Previous studies have shown the applicability of case notes as a source of information for health service research in process assessment, but there is shortage of outcome studies using case notes.

The alleged non-standardization of case notes in psychiatry can be seen as reflecting the lack of standardization in the area as a whole.

The Mental State Examination is an important source of symptom information within psychiatric case notes, with the SCL being a suitable instrument for collecting symptom information. Assessing symptom change over short periods of time is relevant to the assessment of the results of psychiatric treatment.

Audit is essentially a retrospective assessment of the care provided. Measuring outcome from case notes could be an important method to audit psychiatric care.

CHAPTER 2

PURPOSE OF THE STUDY
AND HYPOTHESES

2.1. Purpose of the study

The main goal of this study is to explore a method for routine measurement of outcomes of psychiatric care using clinical case notes. The study aims to make maximum use of existing clinical information, avoiding expensive primary data collection (Alderson, 1983) and discuss the applicability of collecting outcome data from case notes to auditing psychiatric care. Also, concentrating the assessment on the information produced in the clinical notes may raise clinicians interest in evaluative exercises, and create incentives to improve notekeeping.

More specifically, the objectives were:

1. To examine case notes in terms of *availability* of outcome information by demonstrating that there are comparable data present for both admission and discharge.

2. To examine the *reliability* of extracting data from case notes by testing observer agreement.

3. To examine the *validity* of using information obtained from case notes by testing if the data considered are true measures.

2.2. Hypotheses

One main hypothesis is thus formulated:

1. Case notes can be used to measure the outcome of psychiatric treatment.

2.3. Implications

The results of outcome measurement using clinical case notes may be considered:

1. According to different settings and diagnoses

2. To assess the possibility of use for audit.

CHAPTER 3

METHODS

Introduction

This chapter is divided in the following subsections:

- 3.1. The methods used to demonstrate the availability of information and the procedures used to examine outcome measurement according to different settings and diagnoses on the basis of the results of the London case notes sample (London study).

- 3.2. The methods used to show the reliability of extracting information from case notes (London sample).

- 3.3. The methods used to the validation procedure of the concerned measures (Manchester Study).

Ethical permission for the research was given by University College Hospital (U.C.H.) ethical committee and formal consent guaranteed by the concerned consultants (Appendix A).

3.1. London Study

This describes the methods used to demonstrate the availability of information and the procedures used to examine differences between settings and between diagnostic groups. The analyses also included broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses.

3.1.1. Design type

This study is a *retrospective* review of patients' case notes admitted in 1990 to three different psychiatric settings in Camden, London. These settings represent typical different aspects of acute psychiatric services in the National Health Service in England and Wales.

3.1.2. Description of the settings

South Camden, an area with approximately 75000 inhabitants is part of the Bloomsbury and Islington Health Authority (Bloomsbury and Islington Health Authority, 1991). It has access to the following mental health services (Bloomsbury and Islington Health Authority, 1990):

* Day-treatment	-	Jules Thorn Day Hospital, St. Pancras Hospital
* In-patient care	-	Laffan Ward, U.C.H. Dunkley Ward, North Wing, St. Pancras Hospital Tredgold Ward, North Wing, St. Pancras Hospital Ward 3, Friern Hospital
* Intensive care	-	Noel Harris Ward, St. Luke's Hospital
* Out-patient services	-	U.C.H. St. Pancras Hospital Middlesex Hospital

One unit representing each modality was chosen (apart from out-patients). These were Jules Thorn Day Hospital, Laffan Ward and Noel Harris Ward.

Jules Thorn (Setting 1 - S1)

This unit offers day care (25 places) highlighting treatment and support in the community. This is achieved through a multi-disciplinary team (1 consultant psychiatrist, 1 registrar, nursing and occupational therapy staff, 1 social-worker). A new referral is received by a member of the team in a weekly rota and an initial interview data is offered. It aims to identify patient problems and provide them an opportunity to look around the Hospital. After a team discussion an assessment period is offered (2 weeks) and a meeting held after this period decides the following course of treatment.

Each patient has his or her own key-worker who negotiates with the patient the appropriate treatment. This is described as a "client-centred" strategy .

Laffan Ward (Setting 2 - S2)

This unit, an in-patient ward (12 beds) in a general hospital, offers 24 hour treatment based on drugs and group activities developed by nurses and occupational therapists under the supervision of 2 consultant psychiatrists. The Ward offers special services such as transcultural approach for ethnic minorities and care for patients presenting psychosomatic disease. The unit also liaises with other specialties when they need psychiatric support.

Noel Harris Ward (Setting 3 - S3)

This unit with 16 beds offers 24 hour in-patient treatment based on drugs and group activities coordinated by nursing staff under supervision of <u>3 consultant psychiatrists.</u>

The Ward is able to provide intensive care when close observation is required. It has a high admission rate in comparison with other units within the District and a high proportion of these admissions are under sections of the Mental Health Act.

3.1.3. Subjects and study inclusion/exclusion criteria

Subjects

A list of the names of all patients admitted to the 3 settings in 1990 was obtained from the medical record department of each setting and cross-checked with computer lists issued by the District information office. Then, the case notes of all patients admitted to these 3 settings in 1990 were reviewed (N=359) and basic information was collected from these case notes (Appendix C).

Inclusion/exclusion criteria

Patients included in the study were: (1) limited to the adult population aged between 18 and 74, (2) admitted and discharged by the same unit (to ensure the same clinicians were recording at both admission and discharge) and (3) treated by a consultant who gave formal permission to have their patients' case notes reviewed. With these criteria 69 patients were excluded. There were 290 patients (80% of all patients admitted) during the year who fulfilled these criteria.

Exclusions were: (1) those with organic brain disease or mental retardation (n=5) (the instruments selected for collecting clinical information do not discriminate symptoms of these illnesses - Wing, 1983), (2) those with chronic conditions (n=3) (acute services were studied), (3) patients with no evidence of psychiatric disease (n=6), (4) patients remaining in the ward while waiting to be transferred to other psychiatric units (n=30), (5) those who did not have their mental state recorded on admission (n=2) and (6) those without case notes or discharge summaries available at the time of the collection were also excluded (n=19). The availability study is therefore based on the case notes of 225 patients (78% of all admissions included).

Behavioural data were collected from nursing notes. 80% (N=48) of the nursing notes of setting 2 were found whereas only 39% (N=45) of them were available for setting 3. For Setting 1 ward behaviour was not computed because in the Day Hospital there is no collection of traditional nursing information related to ward

63

behaviour. Social information was gathered from nursing notes and from the 'current social situation' section. The 'current social situation' section was found in 83% (N=186) of the case notes and provided social items to be computed.

Patients who discharged themselves without clinician permission (n=34) had data on admission analyzed for availability, but were excluded when discharge information was considered. This reduced 225 to 191 patients.

Patients without any symptom linkage between admission and discharge were excluded when the Final Score of Change, (see page 70), was calculated (n=15). This reduced the total to 176 patients (61% of all patients included).

Table 4.1. shows all the figures for the excluded patients and the sample size for each study phase.

3.1.4. Collecting Information

A Pilot study that gave support to the development of the methods to collect information at admission and at discharge was carried out. We collected information from 25 case notes of patients admitted in 1988 and 1989 to the settings above described and to an out-patient service. We found that the case notes could be used to link related information at admission and at discharge. A

full description of the Pilot study is given in the Appendix B.

3.1.4.1. On admission

<u>Variables and instruments considered</u>

The Syndrome Check List - SCL (Appendix D) and the 'Ward Behaviour' section of the Disability Assessment Scale - DAS (mentioned in section 1.3.6.3.) supported the construction of the two coding instruments for collecting symptoms and behaviour. The former devised by the Medical Research Council - Research Unit at the Institute of Psychiatry was formed out of the 140 symptoms rated in the 9th edition of the Psychiatric Status Examination (PSE). The latter was developed by WHO to assess behaviour of patients admitted to a hospital during follow-up studies. The main criterion for choosing them was their reported ability to collect information from case notes (Wing et al, 1974; WHO/DAS, 1988).

The 'coding' instrument for collecting social information, is a compilation of the main items from well known social rating scales as described in Weissman (1975) and Clare and Cairns (1978). Because there is no standard social schedule to collect information from case notes, we decided to use these 5 general headings. Appendices E and F show the 3 coding instruments and a sample of the actual data collected by each coding instrument.

The symptoms were grouped under 8 main headings as found in Mental State Examinations (MSE) recorded in case notes - *Appearance and General behaviour, Speech, Affect, Mood, Thought Content, Delusions, Perceptions, Hallucinations.* General symptoms, autonomic anxiety and memory and cognition were not considered as it was found to be very difficult to match SCL symptoms with information reported under these headings in case notes. Four codings were made by the rater reviewing the case notes - (1) reported as present - (RP) - when the admitting doctor wrote in the mental state section that a symptom was present at the time of the examination, these symptoms were used to fill in the SCL form deriving a score for each patient (see Appendix J), (2) for a symptom reported as absent - (RA) - when the admitting doctor wrote in the mental state section that a symptom was absent at the time of the examination, (3) for a symptom inferred as absent - (IA) - when the rater could infer from the written mental state at admission that a symptom was absent at the time of the examination, for example, under the symptom group appearance and general behaviour, the statement that the patient was "cooperative and pleasant" allowed symptoms included in this group to be 'inferred as absent' and, (4) for a symptom which is not reported - (NR) - when it was not possible to infer symptom absence or when there was no written statement (present or absent) related to symptoms.

The SCL was completed considering the symptoms reported as present and a total score derived to represent a severity measure on admission. It is important to stress that a symptom was used to fill in the SCL only when it was <u>explicitly</u>

reported as present. For example, a statement like "patient feels influenced" was not considered evidence for rating nuclear syndrome.

Behaviour data was recorded using the 12 items of the 'Ward Behaviour' section of the Disability Assessment Schedule. There were two possible codings - (1) for a behaviour item reported as a problem and, (2) when the behaviour item was not reported.

Finally, social data were recorded, housing, occupation, economic situation, social activities and, family relationship with two codings - (1) for a social item reported as a problem and, (2) when the social item is not reported.

Assembling data

The Pilot Study identified the suitable areas within case notes from where these variables could be collected: Symptom information from the 'history of present complaint' (HPC) section, from the 'mental state examination' (MSE), and, from 'nursing notes' (NN); behaviour items from 'nursing notes'; social data from the section 'current social situation' and 'nursing notes'.

When a case note was reviewed symptom information was first rated from the MSE before any other data was examined. This was to minimise the *halo effect,* "the tendency to make a global judgement early in the process of rating and then

apply that to all subsequent ratings" (Thompson 1989).

After examining the MSE section, HPC and NN were checked and symptoms re-rated in some cases. Whereas the MSE is a cross-sectional picture of patient symptomatology, the HPC could provide a more longitudinal view if symptoms occurring before admission are reported and considered for rating purposes. For example, if the symptoms appetite or sleeping disturbance were not reported in the MSE, but were reported as present in the HPC, these symptoms would be re-rated. Appendices G and H present examples of information related to symptom, behaviour and social items reported in the case notes and their respective codings.

Demographic and clinical variables considered and a sample of the actual data collected are shown in the Appendix I.

3.1.4.2 On Discharge

Having collected symptom information from the mental state at admission the next logical step would be to collect the same information of a mental state at discharge. However, it appears not to be a routine practice to report mental state at discharge as found by Craddock and Craddock (1990), the only study found in the literature that mention this fact.

The great majority of case notes reviewed in this study also did not present a structured and complete mental state at discharge like the one reported on admission.

There are three sections within case notes that judge patient mental, behaviour and social state at the time of discharge:

- 'clinical notes' where judgments are made during ward rounds regarding symptoms elicited on admission and sometimes considerations on social and behaviour conditions are also made.

- 'nursing notes' where by means of periodic reports the condition of patients is stated, including self-care, diet, sleeping patterns, mood, feelings, concentration, memory, relationship with others, work with nurses and general behaviour.

- 'treatment and progress' (discharge summary) which reports on progress patterns of major symptoms elicited on admission and gives a general impression on the mental state at the time of discharge.

Data was drawn from the last ward round and the last nursing report before discharge date and, the 'treatment and progress' section as a whole. Appendix J shows the coding sheet for assembling discharge information using data collected from a real research subject.

3.1.4.3. Statements of change at discharge

The sections were scanned to find statements of change (deterioration, no change, improvement) connected to information recorded on admission. Two types of statements of change were considered.

The first one related to a general assessment of the patient's mental state at the time of discharge that was called the **General Statement of Change (GSC)**. The second statement at discharge was related to each item reported as present at admission. The connecting relationship between items reported as present at admission and the respective discharge information is here named <u>linkage</u>. This linkage was worked out to derive an overall score of change for each patient named the **Final Score of Change (FSC)**.

The two measures (GSC and FSC) were based upon a 3-point rating scale, -1 (deterioration), 0 (no change) and 1 (improvement). Such scale is based upon the general concept of the Likert Scales where the interval between the scores imply dimensions of change (Freeman and Tyrer, 1989). Nunnaly (1978) "strongly believes that it is permissible to treat most of the measurement methods in psychology and behavioral sciences as interval scales". An arbitrary coding (88) was adopted when an information recorded as present on admission did not have any statement of change reported on discharge.

Appendix K and L give examples of these two 'statements of change'.

Deriving a Final Score of Change (FSC) for each patient

The objective of averaging the scores of individual items to derive a overall mean score is to demonstrate the 'direction' and the 'degree' of change for each patient. Like Curran and Pullen (1990) we describe direction of change with the categories deterioration (negative change), no change (null change), and improvement (positive change) and by degree we mean the numerical value assigned to each patient as a final score.

Since ratings made at discharge for each item express the direction and degree of change related to each item in particular, it seems appropriate to derive an overall mean score measuring direction and degree of change for each patient based upon these individuals items (Snaith, 1981).

Calculation of the Final Score of Change (FSC)

Each item reported on admission obtained one possible rating (-1,0,1) at discharge based on the statement related to this item. Each value for each item were added up and mean scores for each patient were derived for symptoms, behaviour and

social conditions.

However, many items reported as present on admission did not have any statement of change recorded at discharge and were coded 88. These non-response items were considered to be missing at random, as they would have been reported had there been an institutional policy to reassess information elicited at admission, when patients were due for discharge. There was no such policy in operation at the settings studied (Little, 1987: 3).

A procedure to estimate values for this missing information used an approximation to regression analysis. The equation was the simplest approach after considering more sophisticated methods such algorithms and regression analysis using SAS. There was no reason to believe that weighting would be better than simple addition. Furthermore, as we were more interested in testing a method rather than in the results we found justified to use a simple method of estimation.

We did not apply a sensitivity analysis to the missing symptoms. The estimation of the missing values was affected by the score of all patients and items. It was assumed that the predictive value of the non-response symptom is a function of the sum of all patients and items mean scores, and that data would be missing at random.

Appendix J presents a worked example using a real research subject that gives an

overview of the main procedures described so far: (1) the derivation of the SCL score from the symptoms 'reported as present', (2) the concept of information linkage between admission and discharge, (3) the equation used to estimate non-response symptom in order to derive the Final Score of Change.

3.1.5. Data analysis

The Statistical Package for Social Sciences - SPSS - (Norusis, 1988) was used to perform the majority of the analysis of this study.

The results on the availability of the information at admission and at discharge were tested by means of chi-square statistics (X^2). Oneway analysis of variance (Kruskal-Wallis) was used to examine the overall proportion of symptom availability when variance between patients were taken into account.

The differences between settings were examined by presenting their respective demographic and clinical variables and also by showing the results on the outcome measures for each setting. Differences in outcome between 3 different diagnostic groups (DG1, DG2, DG3) were examined. Chi-square statistics, non-parametric oneway analysis of variance (Kruskal Wallis) and median test are used to test significance. When the p value was lower than 0.05 the difference was considered to be statistically significant. For this value we use the convention *, ** for p <

0.01 and *** for p < 0.001 [1].

3.2. The reliability of extracting information from case notes

3.2.1. Inter-rater reliability test

Reliability is the property concerned with the reproducibility and repeatability of

any experimental, test, or measuring procedure. In other words the extent to which

they yield the same results when observed by different examiners - reproducibility

- or at different times - repeatability (Wing et al, 1977; Morley and Snaith,

1989/chapter 8; Regier and Burke, 1989/chapter 5; Carmines and Zeller - chapter

1, 1979).

There are several ways of testing reliability. Test-retest method, split-half

reliability, alternative form method and, internal consistency method are concerned

primarily with measuring the reliability of the contents of an instrument. When the

direct observation of an event is the principal aim, the measuring instrument is the

observer not the instrument itself, then an inter-rater reliability test is another

possibility. Therefore the choice of the method for assessing reliability depends on

the type of information that is used by the instrument, and the objective which one

intents to achieve (Regier and Burke, 1989; Morley and Snaith, 1989).

---

[1]This convention was adopted for all analyses in this thesis.

Having said that, an inter-rater reliability test is described since the interest here is to examine if two observers can reliably extract information from the same case note using the same method. The main concern is to demonstrate reproducibility of the method rather that its repeatability since by collecting information from case notes one is not likely to be affected by subject and occasion variance (Akiskal, 1989: 589).

3.2.2. Procedures

A sample of 15 case notes from the London sample (5 for each setting) were randomly selected by using a random number table (Bland, 1987: 8-13). Another issue examined was if the sample size of 15 cases would be sufficient to meet the requirements for assessing this inter-rater reliability. Cicchetti (1976) presents a formula generated by a Monte Carlo study that a minimal sample size value is given by $N \geq 2k^2$ where k is the number of rating points of the scale. For this extracting procedure we were interested in two rating points ('reported as present', and 'not reported as present') leading to the result of $N = 2 \times 2^2 = 8$, as the minimal sample size value required for this study.

Two meetings were held with the other rater, a research psychiatrist also trained in the P.S.E. He was familiarised with the study coding sheets and a standardisation procedure was established based on the general methodology of

this thesis.

3.2.3. Data analysis

The non-parametric coefficient k - kappa - (Cohen, 1960, 1968) was used to measure agreement between the two raters. It has an undoubted advantage over other coefficients of agreement in that it discounts chance agreement between the raters. Fleiss (1975), examining a dozen indexes of agreement concludes that "only Cohen's k (kappa) and Maxwell and Pilliner's r are defensible both as chance corrected measures and as intraclass correlation coefficients. The latter is commonly used for testing reliability of severity scores (ordinal scales) while the former is applied when reliability is tested for categorical information (nominal scales).

We started with the calculation of a collapsed kappa for symptoms reported as present versus not reported as present. Based on the assumption of Barlow et al (1991) that a physician rating may be influenced by severity of disease, kappa coefficient is also provided for different strata of patients (stratified kappa) assuming different settings with different disease severity. Collapsed and stratified kappas are also produced for behaviour and social items.

Weighted kappa for symptom rating is also calculated by allowing different

degrees of agreement (or disagreement) between the two raters. Cohen (1968) suggested that the degree of agreement (or disagreement) could be treated differently accordingly the relative seriousness of each disagreement. The agreement level, or weight - w - of each cell is determined by the researcher where each weight is an arbitrary value. Hall (1974) suggests the adoption of linear or equal weight intervals, where maximum agreement has a weight of 1 and maximum disagreement a weight of 0.

Although specific software for calculating Kappa has been proposed (Gamsu, 1986) a standard spreadsheet (supercalc-5) was used in this study to compute the coefficient from 2 x 2 matrix. SPSS was employed for crosstabulating symptoms between raters. Chapter 4 (Results) shows the matrices and derivation of kappa coefficients for the extracting process.

Classification of the information collected by the SCL was made by the CATEGO program (Wing and Sturt, 1978), which is a software provided by the MRC Research Unit at the Institute of Psychiatric. The output of the program shows classes and subclasses based on syndrome classification, and ICD-diagnoses. A CATEGO output produced for each rater is shown in the Appendix M. Agreement and kappa coefficients were derived from this data.

3.3. Validation study: Manchester Study

3.3.1. Theoretical background

Validity has been defined as the extent to which a particular measure reflects what it purports to measure (Holland, 1983: 43; Grebb, 1989: 9). One validates, not a test, but an interpretation of data obtained by the test under specified conditions (Carmines and Zeller, 1979: 17, 1979; Morley and Snaith, 1989: 134).

Two types of validity are not normally tested, content (built-in characteristics) and face validity (general acceptance of the test or measure) (Thompson, 1989: 10). Criterion validity, degree of comparability with a more accurate measure, can be tested and the two forms of criterion validity are concurrent (correlation of the two tested measures at the same point in time) and predictive (correlation with some future event) validity [2].

Sensitivity and specificity are two important notions connected with the test of criterion validity. The former is concerned with the capacity of the measure to detect the true affected cases. The latter shows if the measure is capable of identifying the true non-affected cases (Lilienfield and Lilienfield, 1980; Holland, 1983; Kirkwood, 1988; Last, 1988; Regier and Burke, 1989: 313). Another useful

---

[2]Construct validity is also mentioned in the literature but it is said to be more relevant for the social sciences.

78

concept in validity test is the predictive positive value that has been defined as the proportion of apparent cases detected by the measure being tested that are really true affected cases (Regier and Burke, 1989).

3.3.2. Validation procedures

For this study, concurrent validity is chosen as the way to test the accuracy of the concerned measures. These measures are tested against two well established instruments, the Present State Examination (PSE), a diagnostic tool which also yields severity score (Wing, 1974; Thompson, 1989; Freeman and Tyrer, 1989) and the Comprehensive Psychopathological Rating Scale (CPRS), an instrument suitable to measure change (Asberg, 1978; Thompson, 1989; Freeman and Tyrer, 1989).

In this study we were concerned in testing the validity of the following procedures and measures. They are (1) the symptom rating from case notes and the process of (2) derivation of syndromes, (3) the generation of a 'severity score' from case notes, (4) the clinical diagnosis reported in the case notes, (5) the general statement of change and, (6) the Final Score of Change (FSC).

In a study at Manchester University, Creed and colleagues have been carrying out a study to compare the outcome and costs of day hospital and in-patient

treatments. The district units serve a population of 140 000 in one of the most socially deprived inner-city areas in Britain with high psychiatric morbidity. There is lack of psychiatric beds in the District so the Day Hospital was planned to take acute admissions direct from the community. The Day Hospital is situated on the site of a district general hospital and most of the in-patient beds are 7 miles away in a psychiatric hospital. (Creed et al, 1989a, 1990, 1991).

Acute psychiatric patients were randomly assigned to these two services and interviewed at different points in time by means of 2 rating schedules, the PSE and the CPRS. The PSE and the CPRS were applied at admission independently from the admitting doctor who reported his/her findings to the case note, and the CPRS was used again at 2 weeks and at 1m, 2m, 3m, 6m, and 12 months from admission.

Using random numbers 40 patients were selected from this study. For the first stage of the validation study, (above-mentioned procedures (1), (2) and (3)), 6 patients were excluded (4 because their case notes were not found and 2 because their PSE interviews were not comparable with case notes). For the validation of the FSC, we excluded 5 patients (2 because they did not have a CPRS score close to discharge, 2 were self-discharged patients and 1 did not have a FSC). To summarise, 34 patients were used to validate admission information and 29 to validate discharge information. Table 4.2. shows these figures.

Then, from their case notes the same information as was obtained from the

London study was collected (see section 3.1.4.). This procedure thus provided

paired independent observations derived from case notes.

After testing validity we compared London and Manchester samples in terms of

their clinical, demographic and outcome information to demonstrate the

applicability of the method in studying district differences. Chi-square and Mann-

Whitney methods were used to test significance when the two samples were

compared.

3.3.3. Data analysis

3.3.3.1. Admission data

Symptom rating

Ratings made from patients' case notes included in the Manchester study were

compared with their independent PSE interviews (PSE here is used as the criterion

instrument). The question to be answered was whether case notes reflect the real

constellation of symptoms presented by patients. Two analyses are carried out.

Firstly, we cross-tabulated case note and PSE ratings to explore ratings 'inferred

as absent' and 'not reported' made from the case notes (see Appendix O).

Secondly, specificity and sensitivity of case notes and its positive predictive value in reporting symptoms were tested (see Appendix P). These epidemiological concepts, called parameters by Kirkwood (1988), have been adapted in the following way:

- Sensitivity (SS) here measures whether the case notes report as present the symptoms identified by the PSE as present.

- Specificity (SP) measures whether case notes report as absent the symptoms identified by the PSE as absent.

- Positive Predictive Value (PPV) assesses which proportion of symptoms identified as present by the case notes are present as reported by the PSE.

By means of a two-way table, symptoms rated from the 34 case notes were compared with the PSE ratings obtained by the Manchester interviewers. From these tables measures of validity (sensitivity, specificity, and positive predictive value) were calculated. A two-way table to calculate these parameters were constructed, as shown:

PSE RATINGS (criterion)

Case note RATINGS	symptom present	symptom absent
symptom present	a	b
symptom absent	c	d

$$\text{Sensitivity} = \frac{a}{a + c} \qquad \text{False negative} = \frac{c}{a + c}$$

$$\text{Specificity} = \frac{d}{b + d} \qquad \text{False positive} = \frac{b}{b + d}$$

$$\text{Positive Predictive Value} = \frac{a}{a + b}$$

We also present a coefficient based on Manchanda and Hirsch (1986) where all ratings are combined leading to the following coefficient of

$$\text{Validity} = \frac{a + d}{a + b + c + d}$$

Derivation of syndromes

Using the same approach syndrome ratings derived from the 34 patients' case notes were compared with the syndromes derived by the PSE interviews. Overall measures of sensitivity, specificity and positive predictive values are calculated (see Appendix Q).

Validating Syndrome Check List (SCL) score as a measure of disease severity on admission

The SCL is derived from the PSE and can be used to rate presence of each syndrome directly from case notes, interviews or other material concerning the clinical history. Again, in order to compare groups, it is important to know the case-mix of each one. Measuring severity on admission could be a good indicator of group characteristics.

The SCL was not devised to be an instrument to measure severity. However, we decided to explore SCL score as a measure of severity by comparing the SCL scores derived from case notes with the severity score generated by PSE interviews.

Descriptive statistics of the two groups of scores are presented, culminating with the presentation of the Spearman's rank coefficient as an analytical measure of the two scores.

Validating clinical diagnosis

One of the objectives of this research is to compare groups of patients. Diagnosis is one of the most important variables to be considered when groups are

compared. We checked the reported diagnosis in the case notes (ICD) against a diagnosis generated by a standard instrument, the PSE (see Appendix R).

Kappa coefficients are commonly used to test inter-rater reliability where the two sources of data are conceived as being of equal status. However, Cohen (1968) also suggested the use of kappa coefficients for testing validity. In this procedure, different from reliability test, he considers the sources being tested with different status, one being a criterion and the other a predictor. In order to do so he proposes not to assign equal weights to symmetric cells in $k \times k$ tables, as usually made when reliability is tested.

Following Cohen's suggestion, the diagnoses yielded by the PSE interviews (criterion) were cross-tabulated with the diagnoses reported in the case notes (predictor), and weights were assigned to cells according to seriousness of disagreement and a weighted kappa was gene^l *1 ^^his result gives the validity of the diagnosis reported in the case notes.

3.3.3.2. Discharge data

On discharge, two pieces of information, General Statement of Change (GSC) and the Final Score of Change (FSC), have been highlighted as the main sources of outcome data. They also have their validity tested to investigate if they are really

true measures of outcome.

The CPRS interviews yielded 2 scores, one at 'admission' and the other at 'discharge'. The CPRS ratings did not coincide exactly with admission and discharge points, so we took CPRS scores as close as possible from these points. The two scores were compared and percentage change between admission and discharge scores were derived.

The general statement of change - related to a general assessment of patient mental state at the time of discharge - appears to have face validity in that its content makes sense to any clinician reading it. However, we decided to assess further its validity.

The Final Score of Change generated in the Manchester case note review is contrasted with the percentage change of the CPRS scores between admission and discharge and a set of descriptive statistics of the two distributions is presented. Secondly, a graph showing the cumulative frequency of the two scores and the Spearman's rank coefficient are presented.

Categorical ratings were assigned to the FSC, GSC and CPRS scores according their values and paired observations were derived (see Appendix S). It is useful to categorise continuous variables, but there is no generally accepted method of doing so (Altman, 1991). It is thought however, that dividing the continuous

variables into three groups is a reasonable approach (Meyer and Province, 1988 in Altman, 1991).

For this study, we divided the continuous variables into two different categorical scales. First, a 3-point rating scale where negative scores were assigned the category 'deterioration', null scores were categorised as 'no change', and positive scores as 'improvement'. The GSC and the FSC were tested against the CPRS by means of this scale. Second, a 5-point rating scale where the category 'improvement' was divided into 'minimal improvement', 'moderate improvement' and 'marked improvement' on the basis of FSC and CPRS scores (see section S.1. in Appendix S).

Tress et al (1987) points out that the more rating points a scale contains the greater its sensitivity. Therefore we also decided to assess the FSC against the CPRS on the basis of the 5-point scale (see Appendix S).

CHAPTER 4

RESULTS

Introduction

The presentation of the results follows a framework given by the main question of this thesis, '*can case notes be used to measure the outcome of psychiatric treatment?*'. This issue is examined through the study of the availability, reliability and validity of the information recorded in the case notes. The results of outcome measurement are presented according to different settings and diagnoses on the basis of the results of the London case notes sample.

4.1. Subjects and Settings: the London Study

Table 4.1. shows the number of patients excluded and the reasons for exclusion from the study. For assessing availability of items at admission we considered the sample of 225 case notes. When we compared admission information with discharge data 34 patients who discharged themselves were excluded (N=191). For the analysis of the Final Score of Change (FSC) we had N=176.

TABLE 4.1. Number of patients and reasons for exclusion from the study: the London sample

EXCLUSION CRITERIA	S1(N=57)	S2(N=67)	S3 (N=166)	TOTAL (290)†
Organic syndrome Mental retardation	2	2	1	5
To be transferred	0	5	25	30
Chronic conditions	0	0	3	3
No evidence of psychiatric disease	1	0	5	6
Mental state not reported on admission	1	0	1	2
Case note or discharge summary not found	4	0	15	19
SAMPLE 1 ††	49	60	116	225
Self-discharged patients	8	10	16	34
SAMPLE 2 ‡	41	50	100	191
patients without any symptom linkage	4	1	10	15
SAMPLE 3 ‡‡	37	49	90	176

†   Patients included in the study
††   Availability study for admission information
‡   Availability study for discharge information and comparison of variables between the settings
‡‡   Analysis of the FSC

4.2. Subjects and Settings: the Manchester Study

Table 4.2. shows the number of patients excluded and reasons for the exclusion

from the study. For testing validity of data at admission we considered the sample

of 34 patients. When we assessed validity of the discharge measures the sample

reduced to N=29.

TABLE 4.2. Number of patients and reasons for the exclusion from the study: the Manchester study

EXCLUSION CRITERIA	DAY HOSPITAL (N=20)	IN-PATIENT UNIT (N=20)	TOTAL (N=40)†
Case note not found	3	1	4
PSE not comparable with case note	2	0	2
SAMPLE 1 ††	15	19	34
Self-discharged patients	0	2	2
No CPRS close to discharge	1	1	2
Patient without symptom linkage	0	1	1
SAMPLE 2 ‡	14	15	29

† Patients selected by using random numbers (N=40)
†† Sample used to validate admission information (N=34)
‡ Sample used to validate discharge information (N=29)

4.3. Availability of data: London study

This section includes the results on the availability of symptom, behavioural and social information on admission and on discharge. Also, availability of clinical and demographic variables is presented.

4.3.1. On Admission

4.3.1.1. Symptomatology

The total number of symptoms (100) multiplied by the total number of patients (N=225), gives the total ratings made (22500) for all case notes. Availability of symptoms for each rating considering N=225 was: 'reported as present' - RP = 1800(8%), 'reported as absent'- RA = 6070(27%), 'inferred as absent' - IA = 4945(22%), and 'not reported - NR 9686(43%). Table 4.3 shows the proportion of available information (symptoms RP, RA, IA, NR) for each symptom group reported in the mental state examination (MSE) and in the 'history of present complaint' for each setting.

There was no reason to believe that weighting the total ratings made for each symptom group by the number of symptoms contained in the group would be better, since each symptom was considered an independent observation and in practice there was little alteration in the proportion of available information when after weighting for number of symptoms in each group.

Table 4.4. shows the available information for each symptom group considering broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses.

TABLE 4.3. Number of symptoms (percentage) recorded in the case notes for symptom groups by availability ('reported as present'- RP, 'reported as absent' - RA, 'inferred as absent'- IA, and 'not reported'- NR) for each setting

SYMPTOMS GROUPS	S1 (N=49) RP/RA/IA	NR	S2 (N=60) RP/RA/IA	NR	S3 (N=116) RP/RA/IA	NR	Total (N=225) †RP/RA/IA	NR	X^2 p
Appearance/Behaviour (11 symptoms)	469(87%)	70(13%)	481(73%)	179(27%)	919(72%)	357(28%)	1869(76%)	606(24%)	48.7 ***
Speech (7)	292(85)	51(15)	344(82)	76(18%)	617(76)	195(24)	1253(80)	322(20)	11.9 **
Affect (9)	300(68)	141(32)	340(63)	200(27)	595(57)	449(43)	1235(61)	790(39)	16.0 ***
Mood (17)	200(24)	633(76)	235(23)	785(77)	394(20)	1578(80)	829(22)	2996(78)	7.0 *
Non-psychotic thought content (11)	226(42)	313(58)	139(27)	521(73)	408(32)	868(68)	773(31)	1702(69)	61.2 ***
Delusions (27)	662(50)	662(50)	1020(63)	600(37)	2600(83)	532(17)	4282(70)	1794(30)	547.4***
Perceptions (8)	184(47)	208(53)	168(35)	312(65)	483(52)	445(48)	835(46)	965(54)	37.2 ***
Hallucinations (10)	328(67)	162(33)	426(71)	174(29)	963(83)	197(17)	1717(76)	533(24)	62.1 ***
Total (100)	2661(54)	2240(46)	3153(53)	2847(47)	6979(60)	4621(40)	12793(57)	9708(43)	110.3 ***

† Availability of symptoms for each rating (N=225): RP = 1800(8%), RA = 6070(27%), IA =4945(22%), NR 9686(43%)

92

TABLE 4.4 Number of symptoms (percentage) recorded in case notes within symptom groups for each setting considering broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses

SYMPTOM GROUP	DG1 (N=63) RP/RA/IA	NR	DG2 (N=47) RP/RA/IA	NR	DG3 (N=66) RP/RA/IA	NR	Other diagnoses (N=49) RP/RA/IA	NR	Total (N=225) RP/RA/IA	NR	X^2 p
Appearance/ Behaviour (11 symptoms)	481(69%)	212(31%)	386(75)	131(25)	542(75)	184(25)	460(85)	79(15)	1869(76%)	606(24%)	6.1 *
Speech (7)	309(70)	132(30)	252(77)	77(23)	414(90)	48(10)	278(81)	65(19)	1253(80)	322(20)	53.4***
Affect (9)	294(52)	273(48)	233(55)	190(45)	435(73)	159(27)	273(62)	168(38)	1235(61)	790(39)	62.8***
Mood (17)	141(13)	927(87)	203(25)	593(75)	299(27)	820(73)	186(22)	656(78)	829(22)	2996(78)	68.7***
Non-psychotic thought content (11)	66(10)	627(90)	96(19)	421(81)	194(27)	532(73)	417(77)	122(23)	773(31)	1702(69)	64.8***
Delusions (27)	1457(86)	243(14)	777(61)	492(39)	1065(60)	717(40)	983(74)	343(26)	4282(70)	1794(30)	330.2***
Perceptions (8)	149(30)	355(70)	179(48)	197(52)	292(55)	236(45)	215(55)	177(45)	835(46)	965(54)	72.2***
Hallucinations (10)	500(79)	130(21)	390(83)	80(17)	470(71)	190(29)	357(73)	133(27)	1717(76)	533(24)	24.0***
Total (100)	3397(54)	2899(46)	2516(54)	2181(46)	3711(56)	2896(44)	3169(65)	1731(35)	12793(57)	9708(43)	9.5**

93

We dealt in these analyses with a large sample of ratings (for example, 6076 for 'delusions' and 11600 for setting 3) therefore a chi-square statistic was able to detect small variations and showed significant difference between the settings and between the 3 main diagnostic groups (DG1, DG2, DG3)for all symptom groups. However, this result does not take into account variance between patients thus inflating the chi-square value. A oneway analysis of variance (Kruskal-Wallis) showed that there was a significant difference in the availability of symptoms between the settings (X^2 = 11.0, p < 0.01) and that there was no significant difference when the 3 main diagnostic groups were considered (X^2 = 2.2, p = 0.33).

Setting 3 (the intensive care ward) had the highest proportion of available information for psychotic symptoms (delusions and hallucinations). These symptom groups were also highly recorded for the diagnostic group 1 (schizophrenia and paranoid disorder) as shown in Table 4.4. Setting 3 treated more psychotic patients than the other two units (table 4.24B). It appears that case notes tend to have higher availability on psychotic than neurotic symptoms.

4.3.1.2. Behaviour and social items

Behavioural data were collected from nursing notes. 80% (N=48) of the nursing notes of setting 2 were found whereas only 39% (N=45) of them were available

for setting 3. For Setting 1 ward behaviour was not computed because in the Day

Hospital there is no collection of traditional nursing information related to ward

behaviour. Social information was gathered from nursing notes and from the

'current social situation' section. The 'current social situation' section was found

in 83% (N=186) of the case notes and provided social items to be computed.

Tables 4.5A and 4.5B presents the findings related to the availability of behaviour

and social items on admission.

TABLE 4.5A. Number (percentage) of behavioural items recorded in the case notes by availability ('reported as a problem' - RP, 'not reported' - NR) for two ward settings

BEHAVIOUR ITEMS	S2 (N=48) RP	NR	S3 (N=45) RP	NR	Total (N=93) RP	NR	X^2
slowness	0(0%)	48(100%)	0(0%)	45(100%)	0(0%)	93(100%)	-
underactivity	1(2)	47(98)	0(0)	45(100)	1(1)	92(99)	0.82NS
overactivity	12(25)	36(75)	15(33)	30(67)	27(29)	66(71)	0.72NS
conversation	2(4)	46(96)	1(2)	44(98)	3(3)	90(97)	0.16NS
social withdrawal	4(8)	44(92)	5(11)	40(89)	9(10)	84(90)	0.09NS
leisure interests	2(4)	46(96)	0(0)	45(100)	2(2)	91(91)	1.66NS
incomprehensible talk	0(0)	48(100)	1(2)	44(98)	1(1)	92(99)	1.24NS
postures/mannerisms	2(4)	46(96)	0(0)	45(100)	2(2)	91(98)	1.66NS
violent behaviour	6(13)	42(87)	5(11)	40(89)	11(12)	82(88)	0.10NS
remaining in bed	8(17)	40(83)	3(7)	42(93)	11(12)	82(88)	1.66NS
personal appearance	8(17)	40(83)	8(18)	37(82)	16(17)	77(83)	0.02NS
behaviour at mealtime	0(0)	48(100)	1(2)	44(98)	1(1)	92(99)	1.24NS
Total	45(8)	531(92)	39(7)	501(93)	84(8)	1032(92)	0.20NS

The low proportion of behaviour items 'reported as a problem' on admission to

the case notes prevented us from making any firm conclusion on the differences

between the settings.

TABLE 4.5B. Number (percentage) of social items recorded in the case notes by availability ('reported as a problem' - RP, 'not reported' - NR) for the 3 settings

SOCIAL ITEMS	S1 (N=44)		S2 (N=48)		S3 (N=94)		Total (N=186)		
	RP	NR	RP	NR	RP	NR	RP	NR	X^2 p
housing	14(32)	30(68)	14(29)	34(71)	23(25)	71(75)	51(27)	135(73)	0.91 NS
occupation/social role	29(66)	15(34)	32(67)	16(33)	67(71)	27(29)	128(70)	58(30)	0.54 NS
economic situation	9(21)	35(79)	7(15)	41(85)	13(14)	81(86)	29(16)	157(84)	1.04 NS
leisure/social activity	1(2)	43(98)	2(4)	46(94)	1 (1)	93(99)	4(2)	182(98)	1.45 NS
family/social relationship	27(61)	17(39)	32(67)	16(33)	31(33)	63(67)	90(48)	96(52)	18.3 ***
Total	80(37)	140(63)	87(36)	153(64)	135(29)	335(71)	302(32)	628(68)	6.2 *

For social data, we found that there were significant differences between the 3 settings for the item 'family/social relationship' and for the total proportion of items reported. We showed that Setting 3, in contrast with symptom availability, had the lowest proportion of reported items.

4.3.1.3. Demographic and clinical information

Table 4.6. shows the availability of demographic and clinical information.

TABLE 4.6. Number (percentage) of case notes with demographic and clinical variables available for each setting (S)

VARIABLES	S1 (N=49)	S2 (N=60)	S3 (N=116)	Total N=225	X^2	p
DEMOGRAPHIC VARIABLES						
Address	49(100%)	60(100%)	114(98%)	223(99%)	1.89	NS
Marital status	49(100)	59(98)	109(94)	217(96)	4.50	NS
Family type	48(98)	57(95)	106(91)	211(94)	2.7	NS
Country of birth	38(78)	51(85)	87(75)	176(78)	2.33	NS
Educational background	42(86)	48(80)	75(65)	165(73)	9.67	**
Occupational background	45(92)	53(88)	87(75)	185(82)	8.70	**
Employment status	43(88)	55(92)	106(91)	204(91)	0.63	NS
Total	314(92)	383(92)	684(84)	1381(88)	18.4	***
CLINICAL VARIABLES						
Age of illness onset	39(80)	49(82)	91(78)	179(80)	0.61	NS
Mean previous admissions	37(76)	43(72)	69(60)	149(66)	3.21	NS
Duration of present illness	45(92)	55(92)	96(83)	196(87)	4.04	NS
Alcohol consumption	40(82)	40(67)	75(65)	155(69)	4.80	NS
Drug consumption	39(80)	37(62)	69(60)	145(64)	6.35	*
Diagnoses	46(94)	59(98)	110(95)	215(96)	1.55	NS
Total	246(84)	283(79)	510(73)	1039(77)	13.7	**

There was a significant difference in the total proportion of available information (demographic and clinical computed together) between the three settings (X^2 = 29.4, df = 2, p < 0.001), with Setting 3 tending to be different from the other two.

Also there was a significant difference between the settings for the following variables, educational and occupational background and history of drug consumption. Clinical variables appeared to be more under-reported than demographic ones.

4.3.1.4. Summary

The study of data availability on admission (N=225) showed that 57% of symptom ratings were considered available information. We also showed that setting 3 had the highest rate of symptom availability (60%). For behaviour and social data, low rates of available information were found, 8% and 32% respectively. For demographic and clinical data case notes proved to be a good source of information as data were available for 88% and 77% patients respectively.

4.3.2. On Discharge

The outcome assessment proposed in this study is based upon the connecting relationship between the information recorded on admission to the one recorded on discharge. Therefore the characteristic of this linkage and the availability of the discharge data are examined.

Two types of link have been considered. The first one identifies a general statement of change (GSC) in relation to the mental state of patients at the time of discharge. The second one considers a specific statement made on discharge in relation to each item (symptom, social or behaviour) recorded as present on admission.

4.3.2.1. Availability of the GSC

Table 4.7. shows the percentage of the availability of the GSC for each setting.

TABLE 4.7. Number (percentage) of case notes with available and unavailable General Statement of Change (GSC) for each setting

GENERAL STATEMENT OF CHANGE	S1 (N=41)	S2 (N=50)	S3 (N=100)	Total N=191
Absent	10(24%)	12(24%)	30(30%)	52(27%)
Unratable	5(12)	5(10)	18(18)	28(15)
Present	26(64)	33(66)	52(52)	111(58)
Total	41(100)	50(100)	100(100)	191(100)

When present the GSC can generate 3 possible ratings (deterioration, no change, improvement) based on the judgment suggested in the statement. For all settings, 58% of case notes presented a ratable GSC (N=111) whereas 27% had no statement of change reported and 15% had unratable statements. A chi-square test

did not show significant difference between the settings ($X^2 = 3.28$, p = 0.2) in relation to available/unavailable (absent plus unratable) GSC. Appendix K gives examples of statements of change.

4.3.2.2. Linkage between admission and discharge

This section presents the proportion of symptoms 'reported as present' and behaviour and social items 'reported as a problem' on admission (tables 4.3, 4.5A, 4.5B) connected with specific discharge information related to these items. For example, as shown at the bottom of table 4.3. (N=225), 1800 (8%) of the total ratings made (N=22500) were symptoms 'reported as present' (RP) and as shown in tables 4.5A and 4.5B, there were 84 behaviour items and 302 social items 'reported as a problem' on admission

Table 4.8A shows which proportion of symptoms 'reported as present' on admission had related information recorded at discharge. We found that 1406(7.3%) symptoms were 'reported as present' on admission. The excluded patients (N=34) accounted for the difference between 1800 and 1406.

TABLE 4.8A. Number (percentage) of symptoms 'reported as present' on admission with specific related discharge information for each setting

SYMPTOMS	S1 (N=41)	S2 (N=50)	S3 (N=100)	Total N=191
link	118(47%)	257(64%)	393(53%)	768(55%)
no link	133(53%)	137(36%)	368(47%)	638(45%)
Total	251(100%)	394(100%)	761(100%)	1406(100%)

$(x^2 = 26.2, df = 2, p < 0.001)$

† 1 patient was excluded because she did not have any symptom 'reported as present' on admission

This table shows that we were able to link 55% of the symptoms reported as present at admission to related symptom information at discharge when all settings are considered together. There was a significant difference between the settings in terms of symptom linkage for admission and discharge information ($X^2 = 26.2$, df = 2, p < 0.001). The day hospital had the lowest linkage between symptom data recorded on admission and on discharge.

Table 4.8B shows which proportion of symptoms 'reported as present' on admission had related information recorded at discharge for broad diagnostic groups DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses.

TABLE 4.8B. Number (percentage) of symptoms 'reported as present' on admission with specific related discharge information for broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses

SYMPTOMS	DG1 (N=58)	DG2 (N=41)	DG3 (N=52)	Other diagnoses (N=39)	TOTAL (N=190)
link	235(52)	162(51)	192(56)	179(61)	768(55)
no link	217(48)	156(49)	151(44)	114(39)	638(45)
Total	452(100)	318(100)	343(100)	293(100)	1406(100)

$(X^2 = 1.9, df = 2, p = 0.45)$

There was no significant difference between the 3 main diagnostic groups (DG1, DG2, DG3).

Tables 4.5A and 4.5B showed that 84 behaviour (N=93) and 302 social (N=186) items were 'reported as a problem' on admission. Tables 4.9A and 4.9B show which proportion of behavioural (51 items) and social (204 items) information 'reported as a problem' on admission had related discharge information. Self-discharged patients were excluded from the study and accounted for the difference between 84/51 for behaviour and 302/204 for social items as shown in tables 4.9A and 4.9B.

102

TABLE 4.9A. Number (percentage) of behaviour items 'reported as a problem' on admission with specific related discharge information for each setting

BEHAVIOUR ITEMS	S2 (N=41)	S3 (N=35)	Total (N=76)
link	23(77%)	16(76%)	39(76%)
no link	7(23%)	5(24%)	12(24%)
Total	30(100%)	21(100%)	51(100%)

$(X^2 = 0.003, df = 1)$

TABLE 4.9B. Number (percentage) of social items 'reported as a problem' on admission with specific related discharge information for each setting

SOCIAL ITEMS	S1 (N=44)	S2 (N=48)	S3 (N=94)	Total (N=157)
link	26(47%)	16(28%)	18(20%)	6(29%)
no link	29(53%)	42(72%)	73(80%)	144(71%)
Total	55(100%)	58(100%)	91(100%)	204(100%)

$(X^2 = 13.2, df = 2, p < 0.01)$

There was a significant difference between the settings in terms of the linkage of social items. The day hospital had the highest linkage of social items. It is not possible to make any firm comment of behaviour items, due to the small number of patients who had these items recorded on admission. But we could say that when they are reported a linkage is possible for the majority.

4.3.2.3. Summary

The study of data availability on discharge (N=191) demonstrated that 58% of the case notes presented a ratable General Statement of Change (GSC). Also, this section showed that there were 55% of symptoms 'reported as present' on admission that had related information on discharge. It was shown that on one hand, the Day Hospital had the lowest symptom linkage (47%), but on the other hand it had the highest social data linkage (47%). The low proportion of behaviour items reported 'as a problem' at admission prevented us from making any firm conclusion on this type of data.

4.4. Reliability of extracting data: London sample

4.4.1 Study group

The sample of 15 patients drawn for this test (see methods section 3.2.2.) was formed by patients with an age range of 24 to 54 years, 6 were females, 8 were psychotic patients and 7 had the clinical diagnosis of neurosis and associated disorders.

4.4.2. Kappa statistics for symptomatology reporting

Symptoms reported as present versus not reported as present

For the first analysis of symptom data, codings (2) 'reported as absent' - RA, (3) 'inferred as absent' - IA and (4) 'not reported' - NR are aggregate into a new category - 'not reported as present'. We have tested here rater capability of distinguishing the symptoms 'reported as present' - RP - from those not reported as present. This distinction has particular importance since the general methodology of this thesis has proposed to assess change only related to symptoms reported as present. Matrix 4.1. sheds some light on the above-mentioned aggregation.

MATRIX 4.1. Symptoms 'reported as present' versus 'not-reported as present' for the 15 patients

Rater B	Rater A 'Not reported as present'	'Reported as present'	Total
'Not reported as present'	1370	18	1388
'Reported as present'	21	91	112
Total	1391	109	1500

1370 - agreement on symptoms 'not reported as present' (codings 2+3+4)
91 - agreement on symptoms 'reported as present' (coding 1)
21/18 - disagreement on symptoms 'reported as present' (coding 1)

The formula for calculating kappa coefficient is given by

$$k = \frac{P_o - P_c}{1 - P_c}$$

Where Po is the agreement between the observers, Pc the chance agreement, and 1 is the total number of observable values. Leading to the kappa coefficient given by

$$k = \frac{1461 - 1287.1}{1500 - 1287.1} = 0.82$$

This coefficient is known as <u>collapsed</u> kappa since all symptoms for every patient are treated together. On the contrary, when a kappa is calculated for different strata we have a <u>stratified</u> kappa (Barlow et al, 1991). A stratified kappa for each patient and each setting is displayed in Table 4.10.

TABLE 4.10. Results of the calculation of a stratified kappa for each patient and setting

SETTING 1 (Day hospital)	SETTING 2 (In-patient unit)	SETTING 3 (Intensive care unit)
PAT. 1 0.80	PAT. 6 0.73	PAT. 11 0.79
PAT. 2 0.74	PAT. 7 0.84	PAT. 12 0.85
PAT. 3 0.88	PAT. 8 0.66	PAT. 13 0.82
PAT. 4 0.88	PAT. 9 0.87	PAT. 14 0.88
PAT. 5 0.76	PAT. 10 0.78	PAT. 15 0.72
mean kappa 0.81	mean kappa 0.78	mean kappa 0.81

An approximation to the standard error of kappa coefficient (Cohen, 1960) is also

calculated to demonstrate how far from the true value the estimate is likely to be.

This standard error is given by

$$\sigma_k \ = \ \sqrt{\frac{P_o(1 - P_o)}{N (1 - P_c)^2}} \ = \ \sqrt{\frac{1461 (1500 - 1461)}{1500 (1500 - 1287.1)^2}} \ = 0.029,$$

leading to the 95% confidence limits = 0.82 + 1.96 x 0.029. It is estimated then

that the chances are 95% that the population value of k falls between 0.79 and

0.88.

From this statistical tests it can be assumed that a sample size of 15 is adequate

to demonstrate the inter-rater reliability in this study.

Examining ratings separately

In the above presentation of results codings 2,3,4 were aggregated into the

category 'not reported as present'. Here they are presented separately and a

weighted kappa is calculated.

So far kappa coefficient has been calculated considering all disagreements equally.

Now disagreements are weighted based on the relative seriousness of each

disagreement. Matrix 4.2. illustrates this procedure where complete agreement (diagonal cells) is weighted 1 and disagreements are weighted 0, 0.25 and 0.75. Disagreements are considered very serious and weighted 0, when one rater finds the symptom 'reported as present' (coding 1) and the other rates it 'as absent' (coding 2 or 3) or 'not reported' (coding 4). A serious disagreement that is weighted 0.25 occurs when a rater judges a symptom 'reported as absent' (coding 2) and the other one rates it as 'inferred as absent' (coding 3) or 'not reported' (coding 4). A less serious discrepancy is weighted 0.75 when a symptom is rated as 'inferred as absent' (coding 3) and the other rater judges it as 'not reported' (coding 4).

MATRIX 4.2. Weighted-kappa matrix for codings 1,2,3,4

	1	2	3	4
1	88a 1b 8.1c	2 0 40.7	7 0 14.8	12 0 45.6
2	4 0 34.1	403 1 174	27 .25 63.0	32 .25 194.8
3	2 0 8.5	24 .25 43.3	78 1 15.7	12 .75 48.5
4	16 0 59.3	131 .25 302	91 .75 109	571 1 338.2

a observed cell proportion
b agreement weight
c chance-expected cell proportion

Weighted kappa is given by

$$Kw = \frac{\Sigma S - \Sigma Sc}{Sm - ESc}$$

Where S is a x b, Sc is b x c, and Sm is Σa x maximum weight. In this matrix

Kappa is given by

$$Kw = \frac{1270.8 - 805.2}{1500 - 805.2} = \textbf{0.67}$$

The other results for this section are presented in Appendices M and N.

4.4.3. Summary

We showed, for a sample of N=15 drawn from the London Study, that moderate to substantial kappa coefficients (0.67 to 0.85) can be obtained when symptom data are extracted from case notes. For behavioural and social data the kappa coefficients were moderate (0.65 to 0.72).

4.5. Validating the information collected: the Manchester study

4.5.1. Study group and information tested

The sample of 34 patients drawn for this study (see table 4.2.) was formed by patients with a mean age of 34 years old, 15 were female, 16 had the diagnosis of psychosis and 18 were neurotic patients or had the diagnosis of associated disorders.

The first part of the study is the validation of the admission information (see section 3.3.3.1) where the whole sample of 34 patients were tested. Four procedures are tested: (1) the process of symptom rating from the case notes and the associated procedure (2) derivation of syndromes, (3) the generation of a 'severity score' yielded by symptoms collected from case notes, and (4) the clinical diagnosis reported in the case notes.

The second stage of the study focus on discharge information (see section 3.3.3.2) and explores a sample of 29 patients (see table 4.2.) Two measures are tested: (1) the general statement of change and, (2) the Final Score of Change (FSC).

4.5.2. Validating admission data

4.5.2.1. Symptom rating

Exploring codings (3) - 'inferred as absent' - and (4) - 'not reported'

By means of a cross-tabulation analysis symptom ratings made in the P.S.E. interviews have been compared with those made from the case notes (see section 3.3.3.1). Appendix O shows this table for all symptoms combined.

Table 4.11. shows the number (percentage) of symptoms 'inferred as absent' and

'not reported' which coincided with ratings 0 (symptom absent) yielded by the

P.S.E. interviews for each symptom group.

TABLE 4.11. Number (percentage) of case note codings (3) - 'inferred as absent' and (4) 'not reported' that agreed (A) and did not agree (NA) with P.S.E. rating 0 (absent) for each symptom section (N=34)

SYMPTOM GROUP (No of symptoms)	CODING 3		CODING 4	
	A	NA	A	NA
Appearance /Behaviour (11)	255(99%)	3(1%)	26(96%)	1(4%)
Speech (7)	141(99)	1(1)	28(100)	0(0)
Affect (9)	199(100)	0(0)	24(100)	0(0)
Mood (17)	†		259(64)	149(36)
Non-psychotic thought content (11)	168(80)	43(20)	91(97)	3(3)
Delusions (27)	284(93)	22(7)	75(91)	7(9)
Perceptions (8)	34(85)	6(15)	169(95)	9(5)
Hallucinations (10)	98(92)	8(8)	17(94)	1(6)
Total (100)	1179(93)	83(7)	689(80)	170(20)

† The phrases recorded in the mental state examination under this heading did not provide sufficient information for inference.

This table shows that inferring absence of symptoms from case notes reflects in the majority of cases actual absence of symptoms as demonstrated by PSE interviews. A symptom 'not reported' in the case notes usually means actual absence of symptoms as demonstrated by PSE interviews. The exception is the symptom group mood with, 36% of not reported information rated present by the PSE.

Derivation of sensitivity, specificity and positive predictive value measures

The previous chapter (see section 3.3.3.1.) shows how these epidemiological parameters were adapted to this research. Calculations have been carried out to produce results for each symptom section as well as overall measures when all sections are considered together. Appendix P shows the calculation of the overall measures. Table 4.12. presents these results.

TABLE 4.12. Results of the calculation of sensitivity (SS), specificity (SP) and positive predictive value (PPV) measures for each symptom group (N=34)

SYMPTOM GROUPS	SS	SP	PPV
Appearance/Behaviour	.30	.95	.13
Speech	.75	.93	.16
Affect	1	.88	.25
Mood	.43	.89	.77
Non-psychotic thought content	.47	.97	.84
Delusions	.37	.99	.66
Perceptions	.30	.99	.86
Hallucinations	.43	.99	.76
OVERALL MEASURES	.43	.95	.62

From this table we can see that the first 3 sections have low PPV: some symptoms under these headings identified as present by the case notes were not confirmed by the PSE interviews as really present. The SP parameters are high for all symptoms groups indicating that symptoms identified as absent by case note rating are truly absent as shown by the independent PSE interviews. The SS parameters

are lower than the SP ones. This indicates higher proportion of false-negative symptoms. In other words, the PSE interviews identified presence of symptoms which were considered absent by case note rating.

The calculation of a coefficient of validity based on Manchanda and Hirsch (1986) where all ratings are combined (section 3.3.3.1) yielded the value of **0.88** (Appendix P).

4.5.2.2. Syndrome rating

Another analysis has compared SCL ratings yielded by the PSE interviews with SCL ratings produced by the case note review. The same parameters (SS, SP, PPV) were calculated for each syndrome. We aggregate the respective parameter of each syndrome into overall ones and table 4.13 presents them. The cross-tabulation of the syndrome ratings are displayed in the Appendix Q.

TABLE 4.13. Results of the calculation of the overall measures of SS, SP, PPV for syndrome rating (N=34)

OVERALL MEASURES	
Sensitivity (SS)	.56
Specificity (SP)	.91
Positive Predictive Value (PPV)	.69

The overall measures of sensitivity presented by this table are still low but higher than the one presented in table 4.12. This indicates that the aggregation of symptoms into syndromes increase sensitivity decreasing false-negatives.

The calculation of the coefficient of validity based on Manchanda and Hirsch (1986) yielded the value of **0.81** (Appendix Q).

4.5.2.3. Validating Syndrome Check List (SCL) score as a measure of disease severity on admission

The SCL was not devised to be an instrument to measure severity originally. It yields a total syndrome count based on the symptoms rated as present by PSE interviews. We explore this total syndrome count, as a measure of severity, by comparing the SCL scores derived from rating symptoms reported in the case notes with the severity score generated by independent PSE interviews.

Firstly descriptive statistics of the two groups of scores are presented[3] culminating with Spearman's rank coefficient as analytical measure of the two scores. Spearman's coefficient measures to what extent subjects are placed in the same severity rank yielded by their respective scores (SCL and PSE at admission).

---

[3]The frequency distribution and the scatter plot are based on the scores re-scaled to 100.

114

Descriptive statistics

Table 4.14. displays the frequency distribution of each score for the Manchester study (N=34) and table 4.15. shows the measures of central tendency and variance for each score distribution. Figure 4.1 shows the two scores in a scatterplot.

TABLE 4.14. Frequency distribution of the SCL and PSE scores for the Manchester study (N=34)

SCORES	SCL SCORE			PSE SCORE		
	Freq.	Cum. freq.	%	Freq.	Cum. freq.	%
0 - 19	1	1	2.9	3	3	8.8
20 - 39	3	4	11.8	11	14	41.2
40 - 59	9	13	38.2	15	29	85.3
60 - 79	16	29	85.3	3	32	94.1
80 - 100	5	34	100.0	2	34	100.0
Total	34	34	100.0	34	34	100.0

TABLE 4.15. Measures of central tendency and variance of the PSE and SCL score distributions (N=34)

MEASURES	SCL SCORES	†	PSE SCORES	†
Mean	17.4	(62.2)	17.2	(45.4)
Median	18.0	(64.2)	18.0	(47.3)
S.D.	5.3	(19.0)	7.7	(20.3)
Interquartile range	5 -- 28	(18-100)	2 -- 38	(5-100)
5% percentile	6.5	(23.2)	3.5	(9.1)
95% percentile	27.0	(97.3)	33.5	(88.1)

† Values after re-scaling them to 100.

115

Figure 4.1
Scatterplot of SCL and PSE scores
(scaled to 100) for 34 patients

Syndrome Check List (SCL)

Present State Examination (PSE)
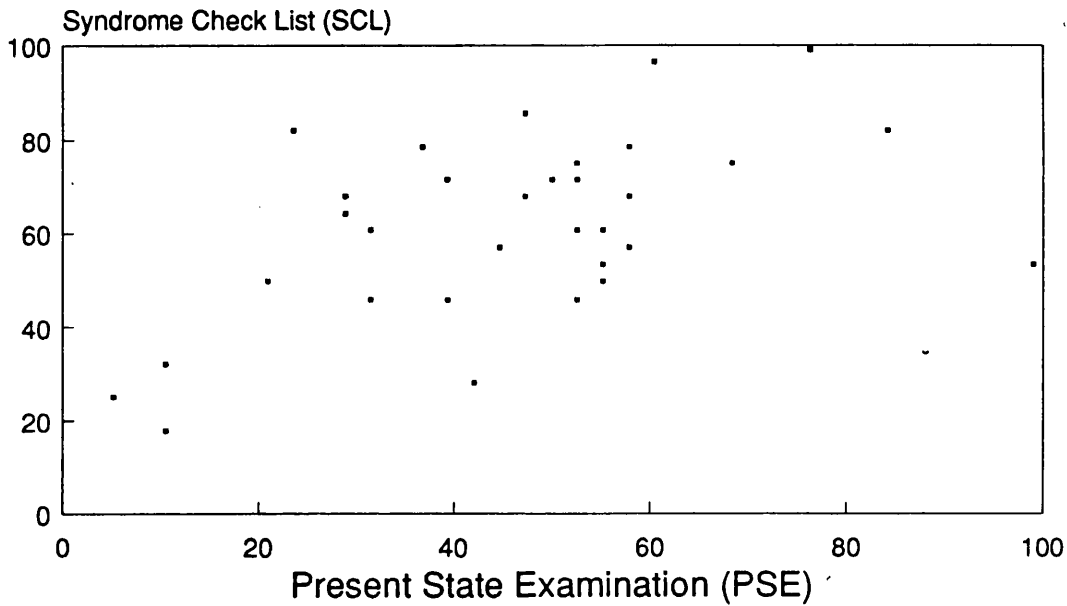
Spearman's (ρ) rank correlation gave ρ = 0.42. The significance was assessed

using a test as described by Kirkwood (1988) given by

$$t = \rho\sqrt{\frac{(n-2)}{(1-\rho^2)}} \quad \text{d.f.}= n - 2 \ = \ 0.42\sqrt{\frac{32}{1-0.42^2}} = 2.61$$

this is significant at the 1% level for 34 pairs confirming a positive association

between the two scores. Despite the weakness of the coefficient, the positive and

significant association suggests that the SCL also mesures severity.

4.5.2.4. Validating clinical diagnosis

The diagnoses yielded by the PSE interviews (criterion) were cross-tabulated with the diagnoses reported in the case notes (predictor) (N=23), weights assigned to cells related to the seriousness of disagreement and a weighted kappa of **0.73** was generated (table and calculation in the Appendix R).

4.5.3. Validating discharge information

On discharge two pieces of information, the General Statement of Change (GSC) and the Final Score of Change (FSC), have been highlighted as the main source of outcome data. They also have their validity tested to investigate if they can be used as measures of outcome.

4.5.3.1. The General Statement of Change

Available general statements (N=18) (62%)[4] - rated -1 (deterioration), 0 (no change), and 1 (improvement) - collected from Manchester case notes were compared to the percentage change between the CPRS scores at admission and at discharge (see section 3.3.3.2).

---

[4]38% of case notes did not present a general statement of change at discharge.

Using paired observation we compared the scores yielded by the measures (GSC and CPRS) on the basis of the outcome category given by the 3-point rating scale (see section S.1. in Appendix S). We found that 94% of the general statement categories matched with the percentage change categories of the CPRS scores, but 6% did not coincide (see section S.2. in Appendix S).

4.5.3.2. The Final Score of Change (FSC)

In this section the Final Score of Change generated in the Manchester case note review is contrasted with percentage change of the CPRS scores between admission and discharge.

Descriptive statistics

Table 4.16. displays the frequency distribution of the FSC and of the CPRS percentage change. Table 4.17. shows measures of central tendency and variance. Figure 4.2 plots the the FSC with the CPRS percentage change as cumulative frequencies.

TABLE 4.16. Frequency distribution of FSC and CPRS (percentage change) for the Manchester sample (N=29)

SCORES	FSC			CPRS		
	Freq.	Cum. Freq.	%	Freq.	Cum. Fr.	%
-1.00 -- -0.70	0	0	0	0	0	0
-0.69 -- -0.40	0	0	0	0	0	0
-0.39 -- -0.10	1	1	3.4	2	2	6.8
-0.09 -- 0.20	2	3	10.2	6	8	20.4
0.21 -- 0.50	1	4	13.8	6	14	48.2
0.51 -- 0.80	8	12	41.3	10	24	82.7
0.81 -- 1.00	17	29	100.0	5	29	100.0
Total	29	29	100.0	29	29	100.0

The FSC presents a skewed distribution as the majority of the values are above 0.66. This is a direct consequence of the scale used to produce the FSC. While for the CPRS the score improves only when a symptom is absent, for the FSC a score of 1 (improvement) can be given to a symptom still present. This fact makes the FSC distribution to have a skewed distribution to the left.

TABLE 4.17. Measures of central tendency and variance of the FSC and the CPRS (percentage change) distributions (N=29)

MEASURES	FSC	CPRS
Median	.85	.53
Mean	.72	.45
S.D.	.31	.34
Interquartile range	-.22 -- 1	-.37 -- .95
5%. percentile	-.11	-.35
95% percentile	1.0	.93

Figure 4.2
Cumulative frequencies of the FSC and CPRS (percentage change)
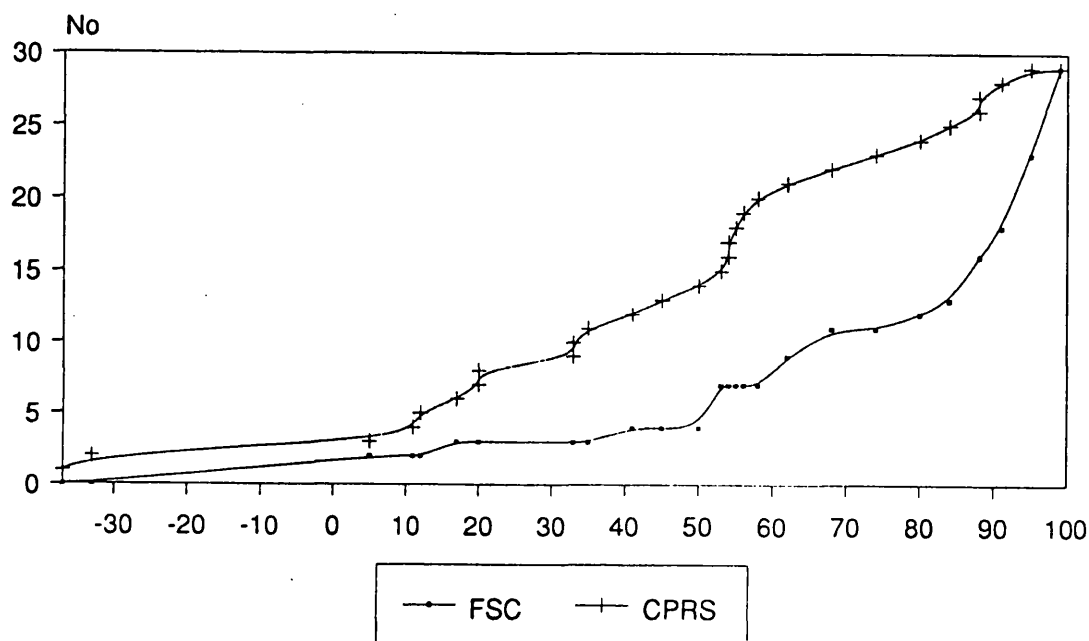
Spearman's (ρ) rank coefficient gave ρ = 0.26. The significance was assessed

using a test as described by Kirkwood (1988) given by

$$t = \rho \sqrt{\frac{(n - 2)}{(1 - \rho^2)}} \quad \text{d.f.} = n - 2 \quad = \quad 0.26 \sqrt{\frac{27}{1 - 0.26^2}} = 1.39$$

This is not significant indicating no association between the two scores.

Analysis of the categorical ratings

The above results showed the analysis of the two outcome measures (FSC and

CPRS) as continuous variables. We also assigned outcome categories to each

measure taking into consideration the values of each score (section S.1. in

Appendix S). Using paired observation we compared the scores yielded by the

measures (FSC and CPRS) on the basis of the outcome category given by the 3-

point and 5-point rating scales (see section S.3. in Appendix S). We found 96%

of matched pairs between the outcome measures when a 3-point rating scale was

considered, and 45% when a 5-point scale was the yardstick (section S.3. in

Appendix S).

4.5.4. Summary

The validation study used the Manchester sample for assessing admission data (N=34). We showed that 93% of the symptoms 'inferred as absent' and 80% of the symptoms 'not reported' in the case notes were rated absent by independent PSE interviews. We obtained substantial specificity coefficients (0.95 and 0.91) for symptom and syndrome rating, respectively and low sensitivity (0.43 and 0.57) and moderate positive predictive value (0.61 and 0.69) measures. From these parameters we deduced that there were 57% of false negative symptoms and only 5% of false positive. A validity coefficient based on Manchanda and Hirsch (1986) for symptom and syndrome rating yielded the values of 0.89 and 0.81, respectively.

SCL scores derived at admission from case notes were compared with independent PSE scores and gave $\rho = 0.42$ ($p < 0.01$). There was considerable agreement (0.73) between CATEGO-diagnoses and clinical diagnoses reported in the case notes.

To test discharge information we used 29 case notes from the Manchester sample. We showed that the GSC (N=18) collected from case notes matched with 94% of the percentage change categories of the CPRS scores. However, ratings derived from the GSC were not very sensitive. For example, CPRS yielded percentage change of 5% and 95% while the GSC only demonstrated one level of

improvement. The correlation of the FSC with CPRS (percentage change) between admission and discharge yielded Spearman rank coefficient of 0.26. A paired observation of the categorical ratings of the FSC and the CPRS percentage change gave 96% of matched pairs between the outcome measures when a 3-point scale was used and 45% for a 5-point rating scale. However, the sample was small and had only 3 cases that deteriorated or had no change.

4.6. London and Manchester samples compared

We compared London with Manchester sample in terms of their clinical, demographic, availability of symptom information at admission and outcome information. Firstly, we compared the samples combined, London - setting 1 and setting 2, N=91 - and Manchester -day hospital and in-patient unit, N=34. Secondly, setting 1 (N=41) is compared with Manchester day hospital (N=15) and setting 2 (N=50) is compared with Manchester in-patient unit (N=19).

4.6.1. Demographic and clinical variables comparison

Tables 4.18A and 4.18B present the comparison of the two samples for Manchester (N=34) and London (N=91) (see Appendix I for details on the demographical variables).

TABLE 4.18A. Comparison of demographic and administrative variables for Manchester (N=34) and London (N=91) samples

DEMOGRAPHIC VARIABLES	Manchester (N=34)	London (N=91)
Sex		
M	19(55)	43(48)
F	15(45)	48(52)
Mean age (years)	37.7	34.6
median length of stay (days)**	30.0	51.0
Patients living		
in the district	31(91)	74(88)
outside the district	4(9)	17(12)
Marital status		
widowed/divorced/single	22(65)	70(77)
married	12(35)	20(22)
information not available	0	1(1)
Family type		
proportion of patients living alone	14(42)	45(49)
other conditions	19(56)	42(46)
information not available	1(2)	4(4)
Country of birth *		
British nationality	27(79)	46(51)
non British	7(21)	27(30)
information not available	0	18(19)
Social class		
1/2	5(15)	21(23)
3/4/5	21(61)	52(57)
not classified	3(9)	11(12)
information not available	5(15)	7(8)
Employment status		
unemployed	23(68)	46(51)
employed	7(21)	12(13)
other status	3(8)	23(25)
information not available	1(3)	10(11)

length of stay (X^2 = 6.54, df = 4, p < 0.01)
country of birth (X^2 = 10.5, p = 0.03)

There were significant differences between the samples for the variables 'length of stay' and 'country of birth' and SCL score.

TABLE 4.18B Comparison of clinical variables for Manchester (N=34) and London (N=91) samples

CLINICAL VARIABLES	Manchester (N=34)	London (N=91)
Age at onset of first psychiatric illness (years)	26.0	24.0
Mean number of previous admissions	2.7	3.3
Duration of present illness	19(56)	41(45)
-3 months	14(41)	38(41)
	1(3)	12(13)
+3 months		
information not available		
Diagnoses		
schizophrenia	10(29)	18(20)
affective disorder	2(6)	17(19)
paranoid disorder	3(9)	3(3)
neurotic depression	7(20)	18(20)
personality disorder	2(6)	12(13)
addiction	1(3)	4(4)
other neuroses	1(3)	3(3)
other diagnoses†	6(18)	14(15)
information not available	2(6)	2(2)
SCL score at admission (median)***	18.0	11.0
% of 'not reported symptoms' at admission***	25%	43%

† ICDs 309 (adjustment reaction), 307.5 (disorders of eating), 308 (acute reaction to stress), 298 (other non-organic psychosis).

We also compared the units of London and Manchester separately to examine which modality of treatment is contributing most to the difference in SCL score. As they also have differences in the rate of symptoms reporting, this kind of analysis is relevant according to the aims of this thesis. Tables 4.19A and 4.19B presents the comparison of the day hospitals and tables 4.20A and 4.20B shows the in-patient units.

TABLE 4.19A. Comparison of demographic and administrative variables for Manchester (N=15) and London (N=41) day hospitals

DEMOGRAPHIC VARIABLES	Manchester (N=15)	London (N=41)
Sex		
M	8(53)	17(42)
F	7(47)	24(48)
Mean age (years)	39.7	35.2
median length of stay (days)**	35.0	101.5
Patients living		
in the district	14(93)	37(90)
outside the district	1(7)	4(10)
Marital status		
widowed/divorced/single	7(47)	33(80)
married	8(53)	8(20)
information not available	0	
Family type *		
proportion of patients living alone	4(27)	24(59)
other conditions	11(73)	16(39)
information not available	0	1(2)
Country of birth		
British nationality	10(67)	22(54)
non British	5(33)	9(22)
information not available		10(24)
Social class		
1/2	2(13)	8(20)
3/4/5	9(60)	28(68)
not classified	2(13)	5(12)
information not available	2(13)	
Employment status		
unemployed	9(60)	23(56)
employed	3(20)	5(12)
other status	2(13)	7(17)
information not available	1(7)	6(15)

length of stay ($X^2 = 7.3$, p < 0.01)
family type ($X^2 = 10.3$, df = 4, p < 0.05)

TABLE 4.19B Comparison of clinical variables for Manchester (N=15) and London (N=41) day hospitals

CLINICAL VARIABLES	Manchester (N=15)	London (N=41)
Age at onset of first psychiatric illness (years)	29.7	23.2
Mean number of previous admissions **	1.9	2.9
Duration of present illness	8(53)	12(29)
-3 months	6(40)	22(54)
+3 months	1(7)	7(17)
information not available		
Diagnoses		
schizophrenia	2(13)	5(12)
affective disorder	1(7)	8(20)
paranoid disorder	3(20)	1(2)
neurotic depression	4(27)	12(29)
personality disorder	0(0)	6(15)
addiction	1(7)	2(5)
other neuroses	0(0)	2(5)
other diagnoses†	3(20)	3(7)
information not available	1(7)	2(5)
SCL score at admission (median) ***	19.0	10.0
% of 'not reported symptoms' at admission ***	27%	44%

† ICDs 309 (adjustment reaction), 307.5 (disorders of eating), 308 (acute reaction to stress), 298 (other non-organic psychosis).

We showed that there were significant differences between the two day hospitals for 'length of stay', 'family type', mean number of previous admissions and SCL score.

TABLE 4.20A. Comparison of demographic and administrative variables for Manchester (N=19) and London (N=50) in-patient units

DEMOGRAPHIC VARIABLES	Manchester (N=19)	London (N=50)
Sex		
M	11(58)	26(52)
F	8(42)	24(48)
Mean age (years)	37.7	34.3
median length of stay (days)	30.0	35.0
Patients living		
in the district	17(89)	37(74)
outside the district	2(11)	13(26)
Marital status		
widowed/divorced/single	4(21)	37(74)
married	15(79)	12(24)
information not available		1(2)
Family type		
proportion of patients living alone	10(53)	21(42)
other conditions	8(42)	26(52)
information not available	1(5)	3(6)
Country of birth **		
British nationality	17(90)	24(48)
non British	2(10)	18(36)
information not available		8(16)
Social class 1/2	3(16)	13(26)
3/4/5	12(63)	24(48)
not classified	1(5)	6(12)
information not available	3(16)	7(14)
Employment status		
unemployed	14(74)	23(46)
employed	4(21)	7(14)
other status	1(5)	16(32)
information not available	0(0)	4(8)

country of birth (X^2 = 8.9, df = 4, p < 0.01)

TABLE 4.20B. Comparison of clinical variables for Manchester (N=19) and London (N=50) in-patient units

CLINICAL VARIABLES	Manchester (N=19)	London (N=50)
Age at onset of first psychiatric illness (years)	24.3	24.7
Mean number of previous admissions	3.3	3.6
Duration of present illness	11(57)	29(58)
-3 months	8(43)	15(32)
+3 months		5(10)
information not available		
Diagnoses		
schizophrenia	8(42)	13(26)
affective disorder	1(5)	9(18)
paranoid disorder	0(0)	2(4)
neurotic depression	3(16)	6(12)
personality disorder	2(11)	6(12)
addiction	0(0)	2(4)
other neuroses	1(5)	1(2)
other diagnoses†	3(16)	11(22)
information not available	1(5)	0(0)
SCL score at admission (median)	18.0	12.0
% of 'not reported symptoms' at admission	25%	43%

† ICDs 309 (adjustment reaction), 307.5 (disorders of eating), 308 (acute reaction to stress), 298 (other non-organic psychosis).

For the in-patient units we found that there were significant differences for the variable country of birth' and SCL score at admission.

4.6.2. Outcome information compared

In this topic the percentage of overall change (table 4.21) and the FSC for each

sample are examined (table 4.22). Firstly some descriptive statistics are presented

and later tests of significance are described.

TABLE 4.21. Comparison of the percentage of overall change for the study samples

OUTCOME CATEGORY	MANCHESTER N=29	LONDON N=86
Deterioration	1(3)	4(5)
No change	1(3)	7(8)
Minimal improvement	1(3)	12(14)
Improved	7(25)	14(16)
Much improved	19(66)	49(57)
Total	29(100)	86(100)

There was no significant difference between the samples in terms of their outcome

categories according to a chi-square statistics ($X^2 = 4.2$, df = 4, p = 0.37).

TABLE 4.22. Measures of central tendency and variance of the FSC (score) for the two samples

STATISTICS	MANCHESTER N=29	LONDON N=86
Mean	.73	.60
Median	.86	.77
S.D.	.31	.42
Skewness	-1.6	-1.3
5% percentile	-.11	0
95% percentile	1.0	1.0

These figures suggest that both distributions are non-symmetrical and negatively skewed. A Mann-Whitney test did not show any significant difference between the means of the samples (p = 0.17). And also a median test also did not show a significant difference between the samples (p = 0.24). These results would suggest that the London and Manchester sample do not differ much from each other in terms of outcome.

4.6.3. Summary

We demonstrated that there were significant differences between the units for the variables 'length of stay', 'country of birth', 'family type and mean number of previous admissions. It was shown that Manchester and London units had different SCL score at admission and different proportion of 'not reported' symptoms on admission. Also, when we compared the modalities of treatment separately the differences in SCL score at admission and 'not reported' symptoms were maintained. It will be discussed later, that local characteristics can also explain the differences in SCL score, but it may be that the differences in symptom reporting can account for this difference. However, in terms of outcome data, there was no significant difference between the samples.

4.7. Examining differences between settings and between diagnostic groups: the London study

As mentioned at the beginning of this chapter, the London study provides the results to explore differences between settings and between diagnostic groups. In this topic we compare units by their main demographic and clinical characteristics. Differences between diagnostic groups are also examined.

4.7.1. Demographic and clinical characteristics

First of all, the 3 settings are compared (see Appendix I for details on the demographic and clinical variables). Table 4.23. shows the referral pattern of the 3 settings and tables 4.24A and 4.24B show demographic and clinical variables.

TABLE 4.23. Number (percentage) of patients referred to the 3 settings by different referral agencies

REFERRAL AGENCIES	S1(N=49)	S2(N=60)	S3(N=116)	N=225
GPs	5(10%)	2(3%)	1(1%)	8(4%)
Social workers	2(4)	2(3)	1(1)	5(2)
Emergency services	1(2)	38(63)	50(43)	89(39)
General Hospitals	1(2)	4(7)	3(2)	8(4)
CPNS	3(6)	0(0)	0(0)	3(1)
Police	0(0)	0(0)	24(20)	24(11)
Self-referrals	6(12)	0(0)	1(1)	7(3)
In-patient units	12(25)	1(2)	14(12)	27(12)
Out-patient services	11(23)	7(12)	2(2)	20(8)
Others	8(16)	6(10)	20(17)	34(16)
Total	49(100)	60(100)	116(100)	225(100)

$(X^2 = 116.0, df = 18, p < 0.001)$

There were significant differences in the referral patterns of the 3 settings indicating considerable heterogeneity between them. For Setting 1 (day hospital) referrers are mainly in-patient units, out-patient services and GPs. For Setting 2 (the general ward) the great majority of patients are admitted from the hospital emergency services. Finally, patients admitted to Setting 3 (the high dependency ward) are referred from emergency services, police, prisons and other in-patient settings.

133

TABLE 4.24A. Comparison of demographic and administrative variables for the 3 London settings

DEMOGRAPHIC VARIABLES	S1 N=41	S2 N=50	S3 N=100	N=191
Sex				
M	17(42)	26(52)	59(59)	102(54)
F	24(48)	24(48)	41(41)	89(46)
Mean age (years)	35.2	34.3	32.5	33.5
median length of stay (days)***	†101.5	35.0	10.0	21.0
Patients living **				
in the district	37(90)	37(74)	54(54)	128(67)
living outside the district	4(10)	13(26)	44(44)	61(32)
	0(0)	0(0)	2(2)	2(1)
Marital status				
widowed/divorced/single	33(80)	37(74)	74(74)	144(75)
married	8(20)	12(24)	21(21)	41(22)
information not available		1(2)	5(5)	6(3)
Family type				
proportion of patients living alone	24(59)	21(42)	57(57)	102(53)
other conditions	16(39)	26(52)	33(33)	75(32)
information not available	1(2)	3(6)	10(10)	14(15)
Country of birth				
British nationality	22(54)	24(48)	39(39)	85(45)
non British	9(22)	18(36)	35(35)	62(33)
information not available	10(24)	8(16)	26(26)	44(22)
Social class				
1/2	8(20)	13(26)	19(19)	40(21)
3/4/5	28(68)	24(48)	52(52)	104(54)
not classified	5(12)	6(12)	5(5)	16(8)
information not available		7(14)	24(24)	31(17)
Employment status	23(56)	23(46)	55(55)	101(53)
unemployed	5(12)	7(14)	18(18)	30(16)
employed	7(17)	16(32)	20(20)	43(22)
other status	6(15)	4 (8)	7(7)	17(9)
information not available				

† Day patients did not attend every day, but we considered their attendance as 5 days per week because they are under day hospital staff responsibility during this period (information gathered from a member of staff).

median length of stay (X^2 = 59.3, p < 0.001)
address (X^2 = 21.0, df = 8, p < 0.01)

TABLE 4.24B Comparison of clinical variables for the 3 London settings

CLINICAL VARIABLES	S1 N=41	S2 N=50	S3 N=100	N=191
Age at onset of first psychiatric illness (years)	23.2	24.7	22.4	22.9
Mean number of previous admissions	2.9	3.6	3.7	3.5
Duration of present illness***				
-3 months	12(29)	29(58)	71(71)	112(59)
+3 months	22(54)	16(32)	12(12)	50(26)
information not available	7(17)	5(10)	17(17)	29(15)
Diagnoses **				
schizophrenia	5(12)	13(26)	36(36)	54(28)
affective disorder	8(20)	9(18)	25(25)	42(22)
paranoid disorder	1(2)	2(4)	1(1)	4(2)
neurotic depression	12(29)	6(12)	6(6)	24(13)
personality disorder	6(15)	6(12)	12(12)	24(13)
addiction	2(5)	2(4)	7(7)	11(6)
other neuroses	2(5)	1(2)	1(1)	4(2)
other diagnoses†	3(7)	11(22)	8(8)	22(11)
information not available	2(5)		4(4)	6(3)
SCL score at admission (median)	10.0	12.0	12.0	12.0

† ICDs 309 (adjustment reaction), 307.5 (disorders of eating), 308 (acute reaction to stress), 298 (other non-organic psychosis).

duration of present illness (X^2 = 41.9, df = 8, p < 0.001)
diagnoses (X^2 = 38.4, df = 18, p < 0.01)

There were significant differences between the settings in terms of the variables 'length of stay', 'address', 'duration of the present illness' and diagnoses. In fact, S3 presented the highest proportion of psychotic patients. Although these characteristics bring about case-mix issues, there were no significant differences between the settings when the medians of the SCL score were tested (X^2 = 3.2, p = 0.2).

Table 4.25. presents the SCL score at admission for broad diagnostic groups, DG1

(schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses.

TABLE 4.25. Comparison of the SCL score at admission for broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses

DIAGNOSTIC GROUPS	SCLSCORE	MINIMUM SCORE	MAXIMUM SCORE
DG1 (N=58)	13	2	32
DG2 (N=41)	12	0	23
DG3 (N=52)	10	2	23
Other diagnoses (N=40)	12	3	29
Total (N=191)	12	0	32

There was significant differences ($X^2 = 6.5$, p = 0.04) between the three main diagnostic groups (DG1, DG2, DG3) when their respective SCL scores at admission are examined by a non-parametric median test. It may be considered that there were differences in the SCL score between the 3 main diagnostic groups with non-psychotic group with lower score. However, as shown in the availability study, it appeared that case notes had higher availability on psychotic than on neurotic symptoms. This may be counted for the significant difference between the 3 main diagnostic groups as SCL score is affected by the rate of reporting at admission.

4.7.2. Outcome information

Two variables are considered, the General Statement of Change (GSC) and the Final Score of Change (FSC).

4.7.2.1. General Statement of Change

For this outcome measure only the patients with a ratable GSC were considered (see table 4.7.). Table 4.26. shows the overall mean and number (percentage) of patients by outcome category for each setting based upon the GSC (see section 3.1.4.3).

TABLE 4.26. Number (percentage) of patients by outcome category based upon the General Statement of Change and the overall mean for each setting

OUTCOME CATEGORY	S1(N=26)	S2(N=33)	S3(N=52)	Total (N=111)
GSC overall mean	0.62	0.67	0.67	0.66
Deterioration	4(15)	0 (0)	2(4)	6(5)
No change	2(8)	11(33)	13(25)	26(24)
Improvement	20(77)	22(67)	37(71)	79(71)
Total	26(100)	33(100)	52(100)	111(100)

A oneway analysis of variance (Krushkal-Wallis) showed that there was no difference between the settings in terms of their mean value of GSC ($X^2 = 0.2$, p=

0.9). The GSC indicates greater likelihood of some deterioration for day hospital patients, but similar levels of improvement for the three settings.

4.7.2.2. Final score of change (FSC)

The presentation of the results on this score includes the following steps:

- Table 4.27. Frequency distributions of the scores for each setting.

- Table 4.28. Categories of overall change for the study groups based upon the Final Score of Change.

- Table 4.29. Measures of central tendency and variance of the score for each setting.

- Table 4.30. Measures of central tendency and variance of the Final Score of Change and categories of overall change considering broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses.

TABLE 4.27. Frequency distribution of the FSC for each setting

Scores	S1(N=37) Freq. %		S2(N=49) Freq. %		S3(N=90) Freq. %		Total(N=176) Freq. %	
-1.00 -- - 0.70	1	2.7	0	0.0	1	1.1	2	1.1
-0.69 -- - 0.40	1	5.4	1	2.0	0	1.1	2	2.2
-0.39 -- - 0.10	1	8.1	0	2.0	0	1.1	1	2.8
-0.09 -- 0.20	6	24.3	9	20.4	13	15.5	28	18.8
0.21 -- 0.50	4	35.1	6	32.6	16	33.3	26	33.5
0.51 -- 0.80	9	59.4	9	51.0	15	50.0	33	52.2
0.81 -- 1.00	15	100.0	24	100.0	45	100.0	84	100.0
Total	37		49		90		176	

TABLE 4.28. Number (percentage) of patients by categories of overall change based on the Final Score of Change (FSC) for the study groups

OUTCOME CATEGORIES	S1(N=37)	S2(N=49)	S3(N=90)	Total (N=176)
Deterioration	3(8)	1(2)	1(1)	5(3)
No change	4(11)	3(6)	6(7)	13(8)
Minimal improvement	4(11)	8(17)	15(17)	27(15)
Moderate improvement	6(16)	7(14)	14(15)	27(15)
Marked improvement	20(54)	30(61)	54(60)	104(59)
Total	37(100)	49(100)	90(100)	176(100)

There was no significant difference between the settings in terms of their outcome categories (x^2 = 6.3, df = 8, p = 0.61). The categories of overall change derived from the FSC suggest greater likelihood of symptom deterioration and no change for the day hospital, but the level of improvement is similar between the settings.

TABLE 4.29. Measures of central tendency and variance of the FSC distribution for the 3 settings

MEASURES	SETTING 1	SETTING 2	SETTING 3
Median	.74	.78	.81
Mean	.55	.62	.65
S.D.	.47	.39	.37
Skewness	-1.2	-1.0	-1.2
Interquartile range	-.92 -- 1	-.59 -- 1	-.94 -- 1
5% percentile	-.54	-.50	.0
95% percentile	1.0	1.0	1.0

These differences between medians and means suggest that none of the settings presented distributions close to the normal curve. In fact, the three distributions have negative skewness coefficient indicating that they are asymmetrical.

The results of a oneway analysis of variance (Kruskal-Wallis procedure) performed to compare score mean of each setting did not show any significant difference in outcome between them ($X^2 = 0.89$, $p = 0.63$).

TABLE 4.30. Measures of central tendency and variance and number (percentage) of patients by categories of overall change based upon the Final Score of Change for broad diagnostic groups, DG1 (schizophrenia and paranoid disorder), DG2 (affective disorder), DG3 (neuroses and personality disorder), and other diagnoses

CATEGORIES OF OVERALL CHANGE	DG1 (N=53)	DG2 (N=37)	DG3 (N=50)	Other diagnoses (N=36)	Total (N=176)
MEAN	0.66	0.59	0.63	0.60	0.63
MEDIAN	0.77	0.85	0.74	0.76	0.78
S.D.	0.31	0.52	0.34	0.43	0.39
Deterioration	0(0)	4(11)	1(2)	0(0)	5(3)
No change	1(2)	3(8)	4(8)	5(14)	13(8)
Minimal improvement	12(22)	4(11)	7(14)	4(11)	27(15)
Moderate improvement	11(21)	2(5)	9(18)	5(14)	27(15)
Marked improvement	29(55)	24(65)	29(58)	22(61)	104(59)
Total	53(100)	37(100)	50(100)	36(100)	176(100)

The results of a oneway analysis of variance (Kruskal-Wallis procedure) performed to compare score mean of the 3 main diagnostic groups (DG1,DG2, DG3) did not show any significant difference in outcome between them ($X^2 = 0.04$, p = 0.1).

4.7.3. Summary

We demonstrated that the 3 settings had some differences between them. Setting 1 had the highest length of stay, and the greatest proportion of patients living in

the district. It also admitted patients with longer duration of present illness. Setting 3 had the highest proportion of psychotic patients whereas Setting 1 presented more patients with the diagnosis of neurotic depression. These characteristics bring about the issue of case-mix and show that we were not dealing in this research with homogeneous groups, an aspect to be aware of when collecting information from case notes. However, there were no significant differences between the settings as measured by the outcome measures GSC and the FSC.

4.8. Final summary

At admission

We found that 57% of the symptom information were available (table 4.3.), but we showed in the parallel validation study that 80% of the 'not reported' symptoms (43%) were really absent as demonstrated by PSE interviews (table 4.11). In other words, the 57% rate may comprise nearly all symptoms actually present and reported 'as absent' at the time of admission. We showed that there was no significant difference in the symptom availability between 3 main diagnostic groups (table 4.4). Moreover, two observers agreed moderate to substantially when they rated symptoms from the same set of case notes (matrices 4.1 and 4.2.). The overall measures of specificity were high (tables 4.12. and 4.13) but sensitivity and positive predictive value parameters were low. From these

parameters we deduced that there were 57% of false negative symptoms and only 5% of false positive.

From symptoms 'reported as present' on admission a SCL score was derived for each patient. The parallel validation study compared SCL score derived from the case notes with independent PSE scores for the same patients. Despite the weakness of the obtained coefficient, the levels of significance indicate that the association is unlikely to have arisen by chance. We can not say that the results obtained in this study completely validate the SCL score as a measure of disease severity at admission. However, the significant association with the PSE suggests that SCL score indicates severity. Nevertheless, the SCL score derived from case notes is prone to be affected by under-reporting as suggested by the results shown in tables 4.18B, 4.19B and 4.20B and by the high rate of false negative symptoms recorded at admission in the case notes.

At discharge

A score related to a general assessment of the patient's mental state at discharge (GSC) was made, but no differences between the settings were found (table 4.26.). However, ratings derived from the GSC were not very sensitive. The number of case notes that recorded a GSC was very small (N=18) and showed that all patients but one had improved. Moreover, the CPRS yielded percentage change of

5% and 95% while the GSC only demonstrated one level of improvement.

A correlation of the FSC with CPRS (percentage change) between admission and discharge yielded Spearman rank coefficient of 0.26. This demonstrated that the two variables yielded different degrees of change. Neverthless a paired observation of the categorical ratings of the FSC and the CPRS percentage change gave 96% of matched pairs between the outcome measures when a 3-point scale was used and 45% for a 5-point rating scale. This association suggests that the FSC indicates symptom change between admission and discharge. However, the sample was small and had only 3 cases that deteriorated or had no change.

Information about 55% of the symptoms reported as present on admission was also recorded on discharge (table 4.8.) and could be assessed for change. There was no significant difference in the proportion of link between symptoms reported as present on admission and related information recorded at discharge when 3 diagnostic groups were considered. Dummy values were estimated for the 45% missing by means of an approximation to regression analysis methods. However, our results should be interpreted cautiously due to the high percentage of missing data. Ratings for each symptom were made and total scores were derived for each patient (the Final Score of Change - FSC). There were no differences between settings and between 3 diagnostic groups using this score (tables 4.29 and 4.30).

Our results on proportion of overall change for each outcome category presented

in the tables 4.26 and 4.28 are proposed as reference levels of symptom change

for each modality of care considered in the present study, providing a link between

our work and medical audit.

CHAPTER 5

DISCUSSION

Introduction

This chapter follows the framework given by the main attributes of the outcome information tested in this study, availability, reliability, validity where the results produced for each attribute is examined. The results of outcome measurement are discussed according to different settings and diagnoses. After that, we present the implications and limitations of this research.

5.1. Availability

In the field of health services research, availability of information plays an important role. Obstacles to undertake an appropriate collection of data vary from the difficulties in accessing the sources (Buck, 1987; Pounder, 1992; Clements, 1992) to the completeness of the information available (Liptzin, 1974; Kiernan, 1976; Black and Creed, 1988; Barrie and Marsh, 1992). Even trivial criticisms to the readability of the sources are reported (Edward et all, 1977; Cunningham, 1991).

In our research, access to the medical case notes may be considered satisfactory since in the London study 93% were found (table 4.1), contrasting with 44% unavailable notes in Dunnill and Gould's audit experiment (1992). However, we did not find 47% of the nursing sections of the case notes. Only 1% of the patients did not have the mental state examination reported at admission, contrasting with 5% in Cunningham's study (1991).

We can consider the percentage of data available in the case notes at admission and at discharge as if it were the 'response rate' of the information concerned (symptoms, behaviour and social items, clinical and demographic variables). By and large, investigators consider a response rate of between 70 to 80 per cent a desirable one (Bland, 1987).

5.1.1. On admission

We found an overall 'response rate' of 83% for demographic and clinical variables, of 57% for symptoms, 8% for behaviour items, and 32% for social items (tables 4.3, 4.5A, 4.5B, 4.6). The 'response rate' for behaviour and social items was low partly because these items were not systematically reported in either the medical or nursing case notes, and partly because some of the nursing notes where we expected behaviour items to be recorded were not found.

In the case of demographic and clinical variables we can say that, in general, psychiatric case notes are a suitable source of information, but there is less information on the variables 'previous admissions', 'history of alcohol' and 'drug consumption'. Setting 3 had more missing information than the other two settings (table 4.6). The fact that this setting admitted more patients into different sections of the Mental Health Act (1983) who were unable at admission to give information might explain this variation. Nevertheless, the information should have been sought, and recorded subsequently.

The 'response rate' for symptoms is lower than the desirable rate of 70% to 80%. However, in the validation procedure, the comparison of 'not reported' symptoms with separate PSE ratings for the same patients revealed that 80% of 'not reported' symptoms in the case notes were really absent (table 4.11). In other words, the 57% 'response rate (codings 1,2,3) comprise nearly all symptoms actually present and reported absent at the time of admission. Statistically significant differences in reporting of symptom groups were found between settings, and by diagnostic groups, but no clear pattern emerged.

5.1.2. On discharge

A General Statement of Change based on the overall assessment of the patient's mental state at discharge, could be identified in 58% of the sample case notes (excluding patients who discharged themselves). No previous studies appear to

have investigated outcome recording in hospital psychiatric case notes, but the GSC proportion is similar to Curran and Pullen (1990) finding of mental health outcome in GP notes. Although we could not use this as a measure of outcome for all case notes (table 4.7), it did give a useful assessment of the overall change when recorded.

Interesting findings arise from the analysis of the linkage between the information recorded on admission and on discharge (tables 4.8A, 4.8B, 4.9A, 4.9B).

Firstly, we were able to link 55% of the symptoms reported at admission to related symptom information at discharge (table 4.8A). The equation proposed in the chapter 3 provided dummy values for symptoms not reported at discharge allowing us to reduce the amount of missing information. Every symptom reported as present at admission was taken into account for outcome consideration. The literature on missing data (Little, 1987) does not suggest at what level of missing data estimation procedures are permissible. However, our results should be interpreted cautiously due to the high percentage of missing data.

Secondly, for behaviour items, due to the small number of patients who had these items recorded on admission, it is not possible to make any firm comment. When they were reported a solid linkage can be established.

Thirdly, the day hospital showed the lowest linkage of the 3 settings for symptom

information and the highest for social data. If we consider the amount of symptom recorded between admission and discharge as one criterion for assessing the quality of notekeeping, this result may indicate lower quality of the day hospital notes.

Finally, for the social reporting we can not suggest lower quality of the notes. We observed during collection of the data that day hospital statements at discharge tended to focus more on social areas than on assessing the mental state of patients. Possibly, the longer 'length of stay' of day hospital patients (Table 4.24A) gives the staff more time for social interventions, thus more interest in social information. The following comment from a patient admitted to the day hospital involved in this study published in The Bloomsbury Mental Health News (March, 1992: 7) illustrates our point:

> "Before you leave here they make sure you've got a place to go, you're on a college course, doing voluntary work, somewhere to stay, or something like that. They keep you in until that's sorted out..."

5.1.3. Concluding remarks

We demonstrated that symptom information was almost complete at admission and that dummy values could be supplied when lacking at discharge. Behavioural and

social data were not present in sufficient amounts to qualify them to be used as measures of outcome.

Outcome is a multidimensional attribute and behaviour and social functioning are important components. It is important to develop means of systematically reporting them in the case notes improving symptom reporting.

What can be done to improve case note reporting? On one hand, the fact that medical audit is dependent on data recorded in the case notes will increase the pressure to improve notes (Ellis et al, 1991). On the other hand, it is open to question whether better record keeping bears any association with good management and higher quality care (McGuire, 1985).

5.2. Reliability

5.2.1. Acceptable level

An acceptable level of reliability depends on the purpose of the method (Remington et al, 1979). It has been suggested that reliability estimates in the range of 0.70 to 0.80 are high enough for most purposes in basic research (Grant and Kaplan, 1989). Everrit (1989) presents a descriptive table of the full range of kappa values and the respective strengths of agreement, and suggests values of

0.60 as moderate, and 0.61-0.80 as substantial. Inter-rater reliability for this study attained moderate to substantial levels. It is worth pointing out, that the two raters were highly trained and it is not likely that an audit activity would use the same type of professionals. This pose some constrains on the generalisability of the results.

5.2.2. Comparison with published studies

Since there appear to be no similar studies in the literature, it is difficult to compare results. However, a comparison can be established with previous studies that reported inter-rater reliability tests for the P.S.E. or for social and behavioural schedules since they all have one common objective: to determine whether or not the measuring instrument produces an acceptable level of reproducibility.

These studies present the following characteristics. Firstly, they were based upon structured or semi-structured interviews applied by either medical or non-medical interviewers. Secondly, the investigation concentrated on intraclass correlation coefficients for severity scores as the principal measure of reliability. Thirdly kappa coefficients were reported only when an item-by-item analysis for each rater combination was undertaken. Some papers presented a simple kappa, others a weighted one. Finally, it is worth noting that the characteristics of the samples were different in each investigation.

Bearing in mind the differences between these studies and the present one, inter-

rater coefficients generated by these studies are presented in tables 5.1 and 5.2. so

that the reader can compare their scores with ours.

TABLE 5.1. Comparison of different studies of inter-rater reliability of the P.S.E

SOURCE	SAMPLE ASPECTS	DATA ANALYZED	COEFFICIENT
Wing (1967)	Out-patients, day-patients and in-patients	diagnosis	84% (not corrected for chance)
Kendell (1968)	In-patients	item-by- item	0.73 (mean kappa) 0.76 (mean weighted kappa)
WHO (1973)	young patients with functional psychosis	item-by-item	0.77 (median Intra Class Coefficient - I.C.C)
Wing (1974)	young patients with functional psychosis	diagnosis	84% (not corrected for change)
Wing (1977)	community	Syndrome agreement	0.52 (mean kappa)
Cooper (1977)	community	item-by-item	0.74 (mean kappa) 0.67 (mean weighted kappa)
Luria (1979)	In-patients	20 symptoms	0.86 (I.C.C)
		Syndrome agreement	0.96 (generalised kappa)
This study	acute day and in- patients	Item-by-item .	0.82 (collapsed kappa) 0.67 (collapsed weighted kappa)
		Syndrome agreement	0.85 (mean kappa)

Table 5.2. Comparison of different studies of inter-rater reliability for reported social and behaviour items

SOURCE	SAMPLE ASPECTS	DATA ANALYZED	COEFFICIENT
Clare (1976)	Patients attending general practitioner	social adjustment	0.55 - 0.94 (weighted kappa)
Platt (1980)	in-patients	social performance and behaviour	48% (weighted kappa 0.90) 68% (weighted kappa 0.80) 83% (weighted kappa 0.70)
Wikes (1986)	variety of placements	behaviour	0.72 (mean weighted kappa) †
WHO/DAS (1988)	Patients with severe psychiatric disorder	social role	0.24 - 1.00 (kappa coefficient)
This study	acute day and in-patients	social problems	0.67 (mean kappa)
		ward behaviour	0.72 (collapsed kappa)

† The authors did not calculate mean kappa.

5.3. Validity

5.3.1. Literature findings

Ferguson and Tyrer (1989: 148) point out that "unfortunately, although reliability can be recorded satisfactorily in most psychiatric research, there are few instances when a scale can be regarded as having proven validity". The study of validity is more complex than the assessment of reliability, so less easily quantifiable (Thompson, 1989). To be valid an instrument should be reliable but reliability does not endorse validity. However, in psychiatry, measures of agreement are

presented as proxy for validity. For example, the comparison of the Diagnostic Interview Schedule (DIS) with an alternative clinical diagnostic procedure has yielded kappa coefficients as measure of its validity (Robins et al, 1982; Thompson, 1989).

Where a new instrument is tested against a well established instrument ('criterion validity') the results are usually presented in terms of the sensitivity and the specificity of the test or in terms of correlation between scores (Burguess, 1991; Barsky, 1992; Bech et al, 1992; Hyler et al, 1992; Kim et al, 1992). Even when the tested instrument does not achieve a high coefficient of validity, authors do not discard it; instead they point to its advantages and suggest ways of improving the new instrument in the future (e.g. Hyler et al, 1992). We have not found any study in the literaure where symptom data recorded in case notes are tested against well established instruments.

5.3.2. Validating the rating process using case notes

This study compared the ability of a research to abstract symptom data at admission from case notes with symptoms recorded separately using a standard instrument, the Present State Examination (PSE). The process of inferring symptoms as absent from case notes reflects in the majority of cases actual absence of symptoms as demonstrated by PSE interviews. However, 20% of

symptoms 'not reported' in the case notes (N=170) were found as present by the

PSE ratings decreasing sensitivity parameters, and 38% of symptoms 'reported as

present' in the case notes (N=133) were found 'as absent' by the PSE ratings

increasing the amount of false-positive symptoms, thereby decreasing the positive

predictive value (see Appendix O). This issue deserves discussion

A clinical episode is a cross-sectional picture of patient symptomatology whereas

the PSE includes symptoms exhibited during the month prior to the interview in

its ratings. These might increase the rate of false-negative symptoms reported in

the case notes, and thereby decrease its sensitivity. By way of a contrast, the PSE

interviews may be carried out after patients had been admitted, when observed

symptoms, such as agitation, for example, might have disappeared due to

medication. So symptoms like these are rated as absent by the PSE, but present

in case notes. This fact might increase the rate of false-positive symptoms, thereby

decreasing the predictive positive value.

Consider the following quote from Regier and Burke, 1989:

> " Higher values of sensitivity and specificity are always desirable.
> For a given instrument, there are trade-offs between these two
> values. The only way to improve both without a trade-off is to
> improve the instrument itself."

We think that the only way to improve case notes so that they reflect the care

provided is to create motivation and incentives for this improvement by

demonstrating the suitability of clinical records for outcome research. However,

as discharge assessment is based on the symptom 'reported as present' on admission, to increase sensitivity of case note reporting (ability to report at admission true present symptoms) appears to be the most sensible approach to be adopted.

5.3.3. Validating SCL score as a severity score at admission

Adding together the scores assigned to different symptoms in order to obtain an overall severity score has been criticised by many authors (Snaith, 1981 for example). However, Hamilton (1968) and Thompson (1989) defend this approach on both mathematical and theoretical grounds since the item-scores being added represent the same concept that is intended to be shown by the overall score.

There are no studies that have tried to derive a severity score at admission by rating symptoms recorded in case notes. The present study has attempted to do so by using the SCL to rate symptoms from the MSE at admission.

The SCL was not devised to be an instrument to measure severity. However, we decided to explore SCL score as a measure of disease severity at admission by comparing the SCL scores derived from case notes with the severity score generated by PSE interviews of the same patient. This approach has particular importance since the available measurement tools in the psychiatric field are

designed for research rather than routine clinical practice (Glover, 1990).

In order to test the validity of the SCL score as a severity measure we compared scores attained at admission for patients in the Manchester study with the PSE ratings for the same patients. The results given in table 4.15 suggest similarities for the two distributions.

We can not say that the results obtained in this study completely validate the SCL score as a measure of disease severity at admission as indicated by PSE scores applied in previous studies (Hoult, 1986; Tyrer et al, 1987; Creed et al, 1990). The positive and significant association with the PSE (see 4.5.2.3) suggests that SCL score can also be used to indicate severity.

The SCL is a simple instrument and can support researchers in collecting baseline outcome information from the routine clinical practice since the majority of the existent measurement tools are less straightforward to apply. However, it should be said that the SCL score derived from case notes is prone to be affected by under-reporting as suggested by the results shown in tables 4.18B, 4.19B and 4.20B. Moreover, the amount of false negative symptoms indicated by this study is also a contributing factor to reduce SCL score.

5.3.4. Validating clinical diagnosis

The kappa coefficient of 0.73 obtained for this validation procedure indicates that it is valid to use diagnoses reported in the case notes in order to classify and compare groups for outcome purposes. A previous study by Schmid et al (1982) comparing independent PSE diagnoses with final diagnosis from the patients' case notes found a considerable lower coefficient of 0.51. This is to be expected as additional sources of information are available when a final diagnosis is made (Schimid, 1982).

5.3.5. Validating discharge information

Previous outcome studies assessed change in Mental State through the comparison of scores measured at baseline with those measured at follow-up points. The same scale was used to rate patient severity at the various measurement points.

In this present study a different approach was adopted because the information reported in the case notes at discharge was not made in terms of absence or presence of symptoms and thus did not meet the requirements for the completion of the SCL schedule. Instead, we measured "direction of change" (Curran and Pullen, 1990) - deterioration, no change and improvement between admission and discharge as measured by the General Statement of Change (GSC) and the Final

Score of Change (FSC). Also, we assume that score values (from -1 to +1) are measuring degree of change along a continuum between these two intervals.

The GSC must be interpreted with caution, because ratings derived from the GSC were not very sensitive when compared with the CPRS. For example, CPRS yielded percentage change between admission and discharge of 5% and 95% while the GSC only demonstrated improvement in mental state without showing different ranges of improvement. Furthermore, the number of case notes that recorded a GSC was very small (N=18) and showed that all patients but one had improved.

Validation against CPRS measurements suggest that the FSC indicates symptom change between admission and discharge. However, the FSC and the CPRS yielded different degrees of change. For example, the FSC presented 14% of its values below 0.46, and 48% between 0.86 - 1.0 whereas the CPRS had 45% below 0.46 and only 14% between 0.86 - 1.0. This is a direct consequence of the scale used to produce the FSC. While for the CPRS the score improves only when a symptom is absent, for the FSC a score of 1 (improvement) can be given to a symptom still present. In addition, we should be cautious in interpreting FSC scores because of the symptoms (45%) that had their values estimated at discharge.

The process of estimating scores for the symptoms without information at

discharge made the FSC not very robust. Therefore, for future research we suggest

a combination of collecting information from case notes with a prospective

assessment at discharge of the symptoms reported as present on admission.

5.3.6. The London and Manchester samples compared

Ideally, a representative sample should have been drawn from the London study

for the validation procedure. However, due to time and financial limitations we

decided to use the PSE and CPRS interviews already carried out in Manchester for

our validation.

A comparison of the two samples revealed that they did not differ significantly in

terms of outcome information (table 4.21), although they presented differences in

the variables 'length of stay', 'country of birth' and 'family type' (tables 4.18A,

4.19a, 4.20A).

The difference in the SCL score between the two day hospital samples (table

4.19B) suggests that the Manchester day unit is admitting more severe patients.

In fact, Creed et al (1987), commenting on Central Manchester psychiatric service,

pointed out that "the day hospital has been used for seriously and acutely ill

patients as a primary treatment facility". In fact, it was designed to take acute

patients direct from the community due to the shortage of psychiatric beds in a

District with high psychiatric morbidity (Creed et al, 1989a; Creed et, 1990).

When comparing London in-patient unit with the Manchester one (tables 4.20A and 4.20B), there was also a significant difference in the SCL score at admission between them. We showed in table 4.11 that 80% of the symptoms 'not reported' at admission were really absent as shown by the PSE interviews. However, the London sample had higher proportion of symptoms 'not reported' at admission (tables 4.20B) which may account for the difference in severity between the in-patient units.

Most clinical research use samples drawn from populations which are more restricted than those about we wish to draw conclusions. This does not mean that studies based on local groups of patients have no value (Bland, 1987). Even when random selection is carried out differences between the samples are shown. Two randomised controlled studies were carried out in Manchester and Blackburn to assess the effectiveness of day care and in-patient treatment but they did not produce comparable groups of patients at each centre (Creed et al, 1991).

5.4. Examining differences between settings and between diagnostic groups

This study provides the first report of the outcome of routine psychiatric care assessed through case note review. The results were considered according to

different settings and diagnoses in order to make more sense to clinicians and to

bring out case-mix between different modalities of treatment. The results obtained

through the comparison of settings and diagnostic groups do not relate directly to

the hypothesis testing process but provide the operational means to show the use

of case notes as a source of outcome data. The selection of the study sample was

constrained by aspects imposed by using case notes. For example, 64 patients were

excluded because the consultants in charge at admission an at discharge were

different not ensuring the same recording standards. As a result we could include

in our study only 49% of all patients admitted to the 3 settings in 1990. This

contrasts with 22% in Herz et al (1971), 15% in Washburn et al (1975), 19% in

Fenton et all (1979), 100% in Hoult (1986) and 94% in Creed et al (1990).

However, our study group sizes (from 37 to 90) are comparable with the range of

study samples presented in the literature (38-60). Contrasting with other studies

(e.g. Dick et al, 1985 who dealt with only neurotic and personality disorders and

adjustment reaction) we considered more diagnostic groups. However, the

distribution of diagnostic groups between settings is unequal and for some settings

it was very small and this impaired the comparison of diagnostic group outcome

by setting. For example, DG1 (schizophrenia and paranoid disorder) was

represented by 5 patients in the day hospital (S1), 19 in the general ward (S2), and

33 in the intensive care ward (S3).

5.4.1. Referral patterns and demographic and clinical variables of the 3 settings

The referral patterns for each setting presented in table 4.23 differ considerably. This is confirmed by the literature in the field that characterizes referral behaviour within the mental health system as a random procedure sometimes dominated by non-clinical factors (Bowman et al, 1983; Fahy et al, 1987; Gotlieb and Wolfson, 1987; Roland and Morris, 1988; Wilkinson, 1989).

In a era where the split between provider and purchaser is a reality the study of the referral behaviour of the system is essential. This study just gives a hint on the subject but we think that further research is necessary since the organisation of the referral process will certainly occupy a high rank in the rationing agenda.

In terms of demographic and clinical variables (table 4.24) the day hospital had the highest median length of stay (3-4 months), more patients with the duration of present illness with more than 3 months (65%), less psychotic patients (35%) and more neurotic patients (35.6%). From table 4.23 we gather that the day hospital had the lowest proportion of patients admitted from emergency service (2.0%).

These characteristics bring about the issue of case-mix and show that we were not dealing in this research with homogeneous groups. But if we had adopted a more strict criteria of inclusion the selection would have became more severe reducing

the size of the groups.

The data obtained make us think about the discrepancy between plans and reality. An internal document written by the day hospital staff states that "the unit mean to focus on the more acute end of the spectrum of psychiatric disorder and we expect the majority of patients to be improved and discharged within 3-4 months of admission".

Just as a matter of comparison, the day hospital in Creed's study (N=41) had 39% patients with the duration of present illness with more than 3 months, 54% of psychotic patients and 27% of neurotic patients, and 49% of patients admitted from emergency services. It should be said, however, that this hospital was originally planned to take acute admissions from the community due to the lack of inpatient beds "in one the most deprived inner areas in Britain with a high psychiatric morbidity" (Creed, 1989a; Creed, 1990)

The day hospital of our study may not be focusing on the more acute end of the spectrum of psychiatric disorder but it is accomplishing its aim of discharging the majority of its patients as improved (81%, table 4.28) within 3-4 months of admission.

The kind of analysis carried out in this section and in the previous one (5.3.6), where psychiatric units in different districts were compared, shows one potential

application of the method, hinting at geographical or service variations.

5.4.2. Outcome measures

The first outcome measure, the GSC, showed no significant differences between the settings (table 4.26). It is worth pointing out that the GSC must be interpreted with caution, because ratings derived from the GSC were not very sensitive as shown in the validation study.

Whereas the GSC expresses a general view of the patient's mental state at the time of discharge, the FSC gives a more sensitive measure at discharge because it considers every symptom that was reported as present on admission.

Comparing the mean FSC for each setting (table 4.29) and by 3 main diagnostic groups (table 4.30) there were no significant differences between the units in terms of symptom change between admission and discharge. These findings are compatible with previous outcome studies (Washburn et al, 1976; Fink et al, 1978; Dick et al, 1985; Creed et al, 1990) that found no long term difference in symptomatology when they compared patients treated by day or in-patient care. Hospital admission represents only one stage of the whole process of treatment. However, the compatibility of this study's findings with previous studies that measured long term treatment outcome might indicate that measuring changes

within the course of hospitalisation might represent a proxy outcome for the whole process of treatment.

5.5. Implications and limitations

Various criticisms are likely to be raised about the use of case notes as a source of evaluative data, the most cogent being the unsystematic and unstandardised way in which the information is recorded. However, this kind of criticism does not only apply to clinical case notes, it applies equally to another written source commonly employed in epidemiological research, the death certificate. It has been commented that "death certificates are not primarily intended for epidemiological research but researchers often rely on them" (Royal College of Physicians of London, 1978). The Medical Services Study Group of the Royal College of Physicians of London (1978) found discrepancies in 49% of the death certificates when they were compared with hospital case notes, consultants' opinions and the necropsy findings.

Others also report inaccuracies in death certification practice. Diehl and Gau (1982) found that doctors who qualified before 1955 tended to certify death by stomach cancer more often than their younger colleagues and Percy et al (1981) found that only 65% of death certificates for cancer were accurate. Kircher et al (1985) compared death certificates with autopsy reports and found a 29% of

disagreement and Alderson and Meade (1967) compared the coding of the 'Principal Condition Treated in the Hospital' with the 'Underlying Cause of Death' for 1216 patients dying in hospital and found discrepancies in 39% of cases. Finally, Dillner (1991) cites a report on *The Autopsy and Audit* suggesting that mortality statistics not backed by necropsies may be inaccurate.

Despite such high levels of inaccuracy, epidemiologists have not given up using information from death certificates. But rather, they demand improvements in the way in which death certificates are completed. By the same token, we suggest that case notes could offer a potential source of data for outcome measurement if more attention was given to increase recording levels of outcome assesment at discharge to related information collect at admission.

In an era where computerised information is increasingly significant, it may seem an eccentric enterprise to propose changes in the way professionals report clinical information in case notes. However, only by showing the usefulness of clinical case notes, as epidemiologists have done with death certificates, can we prompt change in case notes reporting.

It is important to consider the main limitations of assessing outcome of psychiatric care. Sartorius and Harding (1983) argue that "outcome is a hypothetical point at which disease ends rather than the point at which individual is discharged from a facility with a label, say 'improved'. Since our knowledge of the true end point

of mental disease is insufficient, measurement of outcome is impossible, except in the sense of measuring the condition at one point in time". Since mental illness affects different spheres of life in order to measure outcomes of treatment various indicators (psychopathology, behaviour, social role) should be taken into account.

Outcome is the ultimate measure of the quality of care. Quality is a complex concept where different dimensions such as effectiveness, efficiency, equity, accessibility, acceptability and adequacy play their part (Holland, 1983; Donabedian, 1989; Vuori, 1989, Black, 1990). In addition each dimension has its own advocates and is important to explicit from whose perspective we are talking about (Crozier and Spiby, 1989; Vuori, 1989).

In order to conclude, we should make explicit what kind of outcome this present study has measured. It measured the short term effectiveness of psychiatric treatment taking into account one indicator of change, symptom, from the provider point of view. Behaviour and social data were also examined but were not considered in the assessment of change due to lack of systematic reporting.

CHAPTER 6

CONCLUSION AND
FURTHER PROPOSALS

6.1. Introduction

The results showed that case notes do not contain complete information to assess
the outcome of routine psychiatric treatment but the study indicates the potential
for assessing outcome from hospital case notes if recording were improved.

The study of data availability on admission showed that when symptom
information is well reported in the case notes on admission it can be compared
with related data on discharge. For the demographic and clinical variables,
psychiatric case notes were an acceptable source of information, although was less
information on clinical variables. Behaviour and social data had a poor recording
rate and we did not use this information in the measurement of change.

It was demonstrated that data can be extracted from case notes reliably, but the
highly trained raters employed in this study pose obstacles to the generalisability
of the results.

We demonstrated that the process of rating symptoms 'as absent' and 'inferred as absent' from case notes reflected in the majority of cases actual absence of symptoms as demonstrated by PSE interviews. However, one fifth of symptoms 'not reported' in the case notes were found 'as present' by the PSE ratings decreasing sensitivity parameters and SCL scores at admission.

A General Statement of Change based on the overall assessment of the patient's mental state at discharge, could be identified in 58% of the sample case notes. While symptoms 'reported as present' on admission could be linked to information on discharge for 55% of symptoms. But the process of estimating symptoms without information at discharge made the FSC a not very robust measurement.

Case notes may offer a cheap and potential source of data to measure the outcome of routine psychiatric treatment in terms of collecting baseline symptom information. But more attention is needed to improving the ability to report at admission true present symptoms and to recording outcome assessment at discharge.

It should be stated that we are not trying to convince anybody that assessing symptom change between admission and discharge is a panacea for the ailments of outcome measurement in psychiatry. However, by presenting this proposal we are suggesting that using the routine clinical contact as a source of outcome data could open up new possibilities for audit research. This leads to the discussion on

the replicability of the method. In other words, how can this method be applied in public health terms?

6.2. Applicability of the method

Shaw and Costain (1989) states that the process of audit should be repeatable and objective. To follow this advice in practice, instruments should be adopted to facilitate the assessment of the information relevant to the audit process. To replicate this method using only case notes can make it tedious and tardy lowering morale of the professionals in charge of data collection. Therefore, for future research we suggest a combination of collecting information at admission with a prospective assessment at discharge of symptoms reported as present at admission.

In order to start such a process, some conditions should be met. First of all, researchers and administrators have to decide to invest time and financial resources into the field of outcomes of psychiatric care.

Secondly, health professionals need incentives to adopt a different reporting strategy, to combine the conventional reporting on information needed for diagnosis with information relevant to outcome assessment. In particular, the mental state examination needs to be recorded at discharge to allow comparison with the MSE at admission to provide information on symptom change.

Thirdly, there could be encouragement to publish research on outcome using information generated by the clinical process, in other words, giving an incentive to make maximum use of existing clinical information, avoiding expensive primary data collection.

Fourthly, there needs to be development of research in related subjects that could support the process of collecting information from the routine clinical interaction to assess treatment outcomes. For example, more research is needed on linking quality of the notekeeping process with management of the patient. It is still not clear if there is any association between quality of the notekeeping, management of the patient and outcome of treatment. If we are able to demonstrate to professionals the clinical utility of case notes we may create incentives for their improvement. Otherwise, clinicians are apt to continue seeing case notes only as mandatory accountability documents.

Finally, psychiatric treatment involves a comprehensive set of long-term interventions, and positive outcomes are often not seen in a short period. A hospital admission is likely to represent only one segment of the whole process of treatment. Although, in the psychiatric literature, there are authors (e.g. Zusman, 1988; Fauman, 1989) who comment that changes within the course of a hospitalisation might represent a proxy outcome measure for the whole process of the treatment, the issue is still contentious. Therefore, if we think that a result achieved at the time of the discharge is to be considered a good one only if it can

be sustained over time, research on comparing discharge outcome for acute patients with long term results is still needed.

6.3. The method and audit in psychiatry

To set criteria and develop standards are two pre-requisites for an audit process. Criteria should be understood as explicit statements on what constitutes appropriate care and standards are quantitative specifications of what is acceptable practice (Fauman, 1989; Vuori, 1989; Shaw, 1990; Black, 1990; Russel and Wilson, 1992). Shaw (1990) proposes outcome of care as one of the headings to be subjected to the process of criterion definition in an audit process. The question is 'is there any evidence that the initial objectives of the intervention have been met?'.

However, before answering this question another one should be raised: 'what are the objectives of a psychiatric treatment?'. This is a grey area. For example, we do not know what proportion of patients should demonstrate some improvement or, for that matter, what proportion should not be expected to improve at all. In other words, we do not know yet what is appropriate (criteria) in terms of symptom change for the different diagnostic categories.

Our results of overall change for each outcome category presented in tables 4.26

and 4.28 are proposed as reference levels of symptom change (standard) for each modality of care considered in the present study. They could be used as a baseline standard to assess the performance of psychiatric settings and to make comparisons across districts and different modalities of treatment as long as they produce the same kind of information. The replication of such assessment might speed up the development of criteria in terms of symptom change for the different modalities of psychiatric care and different diagnostic categories.

The idea of pooling data for audit assessment - "global audit" - from different units to extend comparison is not new in the surgical field (Gruer et al, 1986; Black, 1989c), but we do not know of such an experience in the psychiatric area. We see this kind of approach of particular interest for measuring outcomes of psychiatric treatment.

In this area, as Zusman (1989) points out "differing outcome may reflect the nature of the patient and the illness as much as they do the care provided". Moreover, it is said that research assessing different modalities of psychiatric care is not a matter of comparing an effective treatment with a relatively ineffective one. But an exercise of comparing two potent treatments and how they are interrelated and searching for specific effects which would discriminate between the treatments (Guy et al, 1969; Michaux et al, 1973; Tansella et al, 1989). It is, though, suggested that the research question should be formulated in a different way: "For whom and under what conditions is one treatment superior to the other"

(Guy, 1969).

Therefore, because comparison between few psychiatrists or few units may generate meaningless results, we will need studies with large samples with group of patients with similar attributes (e.g. diagnosis, occupational and educational background, family type, age, sex to cite some) to be assessed when treated by different treatments.

The existence of large data sets offers the potential of "natural experiment of tremendous size" and a continuing accumulation of data (Epstein, 1990). This can also expand the research focus to a broader set of outcomes. Furthermore, large data set can enhance the assessment of met needs and geographical equity of health service use (Black, 1990).

6.4. Final remarks

The current state of case note reporting does not permit a valid measurement of outcome of routine psychiatric treatment. The considerable availability of symptom information found at admission was not matched by the same amount of discharge statements. However, we were able to link 55% of symptoms reported as present at admission with related information at discharge and this could be raised if more forethought given (or required of them) by clinicians.

The estimation of scores for missing discharge data was not very robust. Therefore, future research should combine collecting information at admission from case notes with a prospective assessment at discharge of the symptoms reported as present at admission.

Differences between setting and diagnoses were examined. Little differences were found in the availability of symptoms and also little differences in outcome were shown as literature would suggest.

We presented the percentage of overall change based on the FSC as bench marks against which future research or audit results can be measured.

In order to conclude we summarise the main implications of this research.

Firstly, reviewing information produced by case notes may involve clinicians in evaluation exercises creating incentives to improve records.

Secondly, assessing quality of the notekeeping process and outcome of the care associated with it can help to answer the question whether bad reporting is associated with bad management of patients.

Thirdly, working with routine clinical data reported in case notes can enhance research through associating the process of immediate care with long term

outcomes.

Fourthly, as psychiatric case notes are structured in a similar way between different districts, collecting information from them can help to raise questions on geographical variations in the care patterns.

REFERENCES

ACCESS TO HEALTH RECORDS ACT (1990). London: Her Majesty Stationery Office.

AKISKAL, H.S. (1989). The classification of mental disorders. In: Kaplan, H.I. & Sadock. B.J. (eds) *Comprehensive Textbook of Psychiatry*, Vol. 1, 5th ed. Baltimore: Williams and Wilkins.

ALDERSON, M.R. & MEADE, T.W. (1967). Accuracy of diagnosis on death certificates compared with that in hospital records. *British Journal of Preventive and Social Medicine*, 21: 22-29.

ALDERSON, M. (1983). Routine sources of morbidity and other statistics. In: Alderson, M. *Introduction to Epidemiology*. London: MacMillan.

ALTMAM, D.G. (1991). Categorising continuous variables. *British Journal of Cancer*, 64: 975.

ASBERG, M., MONTGOMERY, S.A., PERRIS, C., SCHALLING, D., & SEDVALL, G. (1978). A Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica* (Supplement 271).

THE AUTOPSY AND AUDIT. Report of the Joint Working Party of the Royal College of Pathologists, Royal College of Physicians of London, Royal College of Surgeons of England. (cited by Dillner, 1991).

AVISON W.R. & SPEECHLEY K.N. (1987). The discharged psychiatric patient: a review of social, social psychological, and psychiatric correlates of outcome. *American Journal of Psychiatry*, 144: 10-18.

BACHRACH, L. (1976). *Deinstitutionalisation: an analytical review and sociological perspective*. Rockville, Maryland: US Department of Health,

Education and Welfare, NIMH. (cited by Thornicroft and Bebbington, 1989)

BALLARD (1990). Psychiatric in-patient audit - the patients' perspective. *Psychiatric Bulletin*, **14**: 674-675.

BARSKY, A.J., CLEARY, P.D., WYSHAK, G., SPITZER, R.L., WILLIAMS, J.B., & KLERMAN, G.L. (1992). A structured diagnostic interview for hypochondriasis. A proposed criterion standard. *Journal of Nervous and Mental Disease*, **180**: 20-7.

BARLOW, W., LAI, M.Y., & AZEN, S.P. (1991). A comparison of methods for calculating a stratified kappa. *Statistics in Medicine,* **10**: 1465-1472.

BARRIE, J.L. & MARSH, D.R. (1992). Quality of data in the Manchester orthopaedic database. *British Medical Journal*, **304**: 159-162.

BECH, P., ALLERUP, P., MAIER, W., ALBUS, M., LAVORI, P, & AYUSO, J.L. (1992). The Hamilton scales and the Hopkins Symptom Check List (SCL-90). A cross-national validity study in patients with panic disorders. *British Journal of Psychiatry*, **160**: 206-211.

BLACK, D. & CREED, F. (1988). Assessment of self-poisoning patients by psychiatrists and junior medical staff. *Journal of the Royal Society of Medicine*, **81**: 97-99.

BLACK, N., CHAPPLE, J., DALZIEL, M., & SPIBY, J. (1989a). Step by step to audit. *The Health Service Journal*, **99**: 140-141.

BLACK, N. (1989b). A regional computerised surgical audit project. *Quality Assurance in Health Care*, **2**: 263-270.

BLACK, N. (1989c). Beating the offside trap: developing a regional approach to clinical audit. *Journal of Management in Medicine*, **4**: 38-42.

BLACK, N. (1990). Quality assurance of medical care. *Journal of Public Health Medicine*, **12**(2): 97-104.

BLACK, N. (1992). Research, audit and education. *British Medical Journal*, **304**: 698-700.

BLAND, M. (1987). *An Introduction to Medical Statistics*. Oxford: Oxford University Press.

BLOOMSBURY MENTAL HEALTH NEWS (1992). User views on community care, March. London: Bloomsbury Community Health Council.

BLOOMSBURY AND ISLINGTON HEALTH AUTHORITY (1990). *Mental Health Services (Southside): Business Plan 1991/92*. Draft 4, October 1990.

BLOOMSBURY AND ISLINGTON HEALTH AUTHORITY (1991). *Public Health Report*.

BOWDEN, D. & WALSHE, K. (1991). When medical audit starts to count. *British Medical Journal*, **303**: 101-103.

BOWLING, A (1991). *Measuring Health*. Milton Keynes: Open University Press.

BOWMAN, E.P., SHELLEY, R.K., SKEFFINGTON, A.S.& SINANAN, K. (1983). Day patient versus in-patient: factors determining selection of acutely ill patients for hospital treatment. *British Journal of Psychiatry*, **142**: 584-587.

BRANCATI, F.L. (1992). Readers of the lost chart: an archaeologic approach to the medical record. *Journal of the American Medical Association*, **267**: 1860-1861.

BRAUN P., KOCHANSKY, G., SHAPIRO, R., GREEMBERG, S., GUDEMAN, J.E., JOHNSON, S., & SHORE, M.F. (1981). Overview: Deinstitutionalisation of psychiatric patients, a critical review of outcome studies. *American Journal of Psychiatry*, **138**: 736-749.

BROWN, G.W., BONE, M., DALISON, B., et al (1966). *Schizophrenia and Social Care: A Comparative Study of 339 Schizophrenic Patients*. London: Oxford University Press (cited by Braun et al, 1981).

181

BUCK, N., DEVLIN, H.B., & LUNN, J.N. (1987). *The Report of a Confidential Enquiry into Perioperative Deaths*. London: Nuffield Provincial Hospital Trust.

BURGESS, J.W. (1991).The Personality Inventory Scale: a self-rating clinical instrument for diagnosis of personality disorder. *Psychological Report*, **69**: 1235-1246.

BURTON, M. (1986). What do we mean by evaluation? *The Health Service Journal*, 17 July, 954-955.

CAHN, C. & RICHMAN, A. (1985). Quality assurance in psychiatry. *Canadian Journal of Psychiatry*, **30**: 148-152.

CARMINES, E.G. & ZELLER, R.A. (1979). Reliability and Validity Assessment. In: Sullivan J.L. (ed.) *Quantitative Applications in the Social Sciences*. Beverly Hills: Sage Publications Inc.

CASPER, E.S.(1987). A management system to maximise compliance with standards for medical records. *Hospital and Community Psychiatry*, **38**: 1191-1194.

CHALMERS, A.F. (1982). *What is this thing called Science?* Milton Keynes: The Open University Press.

CICCHETTI, D.V. (1976). Assessing inter-rater reliability for rating scales: resolving some basic issues, *British Journal of Psychiatry*, **129**: 452-456.

CLARE, A.W. & CAIRNS, V.E. (1978). Design, development and use of a standardized interview to assess social maladjustment and dysfunction in community studies. *Psychological Medicine*, **8**: 589-604.

CLEMENTS, D. (1992). Audit: the crucial role of medical records. *British Medical Journal*, **304**: 643.

COCHRANE, A.L. (1971). *Effectiveness and Efficiency*. London: Nuffield Provincial Hospital trust.

COHEN, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**: 37-46.

COHEN. J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, **70**: 213-220.

COOPER, J.E., COPELAND J.R.M., BROWN, G.J., HARRIS, T., & GOURLAY, A.J. (1977). Further studies on interviewer training and inter-rater reliability of the Present State Examination (PSE). *Psychological Medicine*, **7**: 517-523.

CRADDOCK, N. & CRADDOCK, B. (1990) Audit of psychiatric discharge summaries. *Psychiatric Bulletin*, **14**: 618-620.

CRAFT, M. (1958). An evaluation of treatment of depressive illness in a day hospital. *The Lancet*, **ii**: 149-151.

CREED F., OSBORNE M., THOMAS P., BLACK D., & ANTHONY P. (1987). Psychiatric day-hospitals. *The Lancet*, **2**(8563): 1465.

CREED F., ANTHONY P., GODBERT K. & HUXLEY P. (1989a). Treatment of severe psychiatric illness in a day hospital. *British Journal of Psychiatry*, **154**: 341-347.

CREED F., BLACK D. & ANTHONY P. (1989b). Day-hospital and community treatment for acute psychiatric illness: a critical appraisal. *British Journal of Psychiatry*, **154**: 300-310.

CREED F., BLACK D., ANTHONY P., OSBORNE M., THOMAS P. & TOMENSON B. (1990). Randomised controlled trial of day patient versus inpatient psychiatric treatment. *British Medical Journal*, **300**: 1033-1037.

CREED F., BLACK D., ANTHONY P., OSBORNE M., THOMAS P., FRANKS, D., POLLEY, R., LANCASHIRE, S., SALEEN, P., & TOMENSON B. (1991). Randomised controlled trial of day and in-patient treatment. 2: comparison of two hospitals. *British Journal of Psychiatry*, **158**: 183-189.

183

CROZIER, A.M. & SPIBY, J. (1989). *Audit in Mental Health*. North West Thames Regional Health Authority.

CSERNANSKY, J.G., YESAVAGE, J.A., MALONEY, W., & KAPLAN, J. (1983). The Treatment Response Scale: a retrospective method of assessing response to neuroleptics. *American Journal of Psychiatry*, **140**: 1210-1213.

CUNNINGHAM, S. (1991). Are your casenotes perfect? *Psychiatric Bulletin*, **15**: 672-674.

CURRAN, S.M. & PULLEN, I. (1990). Audit of a psychiatric liaison service - the value of general practice casenotes. *Psychiatric Bulletin*, **14**: 727-729.

DEPARTMENT OF HEALTH (1989). *Working for Patients*. London: Her Majesty Stationery Office

DICK P., CAMERON L., COHEN D., BARLOW M. & INCE A. (1985). Day and full time psychiatric treatment: a controlled comparison. *British Journal of Psychiatry*, **147**: 246-250.

DIEHL, A.K. & GAU, D.W. (1982). Death certification by British doctors: a demographic analysis. *Journal of Epidemiology and Community Health*, **36**: 146-149.

DILLNER, L. (1991). Necropsies and audit. *British Medical Journal*, **303**: 807-808.

DONABEDIAN, A (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly*, **44**: 166-206.

DONABEDIAN, A (1977). Evaluating the quality of medical care. In: Ertel, P.Y. & Aldridge, M.G. (eds) *Medical Peer Review*. Saint Louis: The C. V. Mosby Company.

DUNNILL , M.G.S. & GOULD, S.R. (1992). Audit of gastrointestinal bleeding in a district general hospital. *British Medical Journal*, **304**: 383-384.

184

EDINCOTT, J. & SPITZER, R.L. (1978). A diagnostic interview: the schedule for affective disorders and schizophrenia. *Archives of General Psychiatry*, **35**: 837-844.

EDWARDS, D.W., YARVIS, R.M., & MUELLER, D.P. (1979). Evidence for efficacy of partial hospitalization: Data from two studies. *Hospital and Community Psychiatry*, **30**: 97 (cited by Greene and De La Cruz, 1981).

EDWARDS, J.G., NUMM C.M.H., & RICKETTS, B.S.(1987). Three years of medical audit in a psychiatric unit. *Bulletin of the Royal College of Psychiatrists*, **11**: 154-155.

ELLIS, P.M., MELLSOP, G.W., PEACE, K.A., & WILSON, J.M. (1987). Peer review as an aid to improving the completeness of psychiatric case notes. *Medical education*, **21**: 493-497.

ELLIS, P.M., BLACKSHAW, G., PURDIE, G.L., & MELLSOP, G.W. (1991). Clinical information in psychiatric practice: what do doctors know, what do they think is known and what do they record? *Medical Education*, **25**: 438-443.

ELLIS, B.W. & SENSKY, T. (1991). A clinician's guide to setting up audit. *British Medical Journal*, **302**: 704-707.

EPSTEIN. A.M. (1990). The outcome movement - Will it get us where we want to go? *The New england Journal of Medicine*, **323**: 266-270.

EVERITT, B.S. (1989). *Statistical Methods For Medical Investigations*. New York: Oxford University Press.

FAHY, T.A., BERMINGHAM, D., & DUNN, J. (1987). Police admissions to psychiatric hospitals: a challenge to community psychiatry. *Medicine Science and Law*, **27**: 263-268.

FAUMAN, M.A. (1989). Quality assurance monitoring in psychiatry. *American Journal of Psychiatry*, **146**: 1121-1130.

185

FEINSTEIN, A.R., PRITCHETT, J.A., & SCHIMPFF, C.R. (1969). The Epidemiology of Cancer Therapy. *Archives of Internal Medicine*, **123** (April - special communication).

FENTON, F.R. TESSIER, L., STRUENING, S.L. (1979). A comparative trial of home and hospital psychiatric care. *Archives of General Psychiatry*, **36**: 1073-1079.

FENTON, G.W., BALLINGER, B.R., & NAYLOR, G.J. (1990). Medical audit in a scottish psychiatric service. *Psychiatric Bulletin*, **14**: 136-139.

FERGUSON, B. & TYRER, P. (1989). Rating instruments in psychiatric research. In: Freeman, C. & Tyrer, P. (eds) *Research Methods in Psychiatry*, London: Royal College of Psychiatrists.

FINK, E.B., LONGABAUGH, R. & STOUT, R. (1978). The paradoxical underutilisation of partial hospitalization. *American Journal of Psychiatry*, **135**: 713-716.

FISH, D.B. (1989). Medical audit: outcome measures needed. *British Medical Journal*, **299**: 1361.

FITZ-GIBSON, C.T. & MORRIS, L.L. (1987). *How to Design a Program Evaluation*. Newsbury Park: Sage Publication.

FLEISS, J.L. (1975). Measuring agreement between two judges on the presence or absence of a trait. *Biometrics*, **31**: 651-659.

FONTANA, A.F. & DOWDS, B. N. (1975). Assessing treatment outcome. *The Journal of Nervous and Mental Disease*, **161**: 221-230.

FRATER, A. & COSTAIN, D. (1992). Any better? Outcome measures in medical audit. *British Medical Journal*, **304**: 519-520.

GABBAY, J. & LAYTON, A.J. (1992). Evaluation of audit of medical inpatient records in a district general hospital. *Quality in Health Care*, **1**: 43-47.

FREEMAN, C. & TYRER, P. (1989) (eds) *Research Methods in Psychiatry*, London: Royal College of Psychiatrists.

GAMSU, C.V. (1986). Calculating reliability measures for ordinal data. *British Journal of Clinical Psychology*, **25**: 307-308.

GARDEN, G. (1989). Audit in Psychiatry. *Psychiatric Bulletin*, **13**: 278-281.

GATH, A (1989). Audit. *Psychiatric Bulletin*, **15**: 23-25.

GELDER, M., GATH, D., MAYOU, R. (1989). *Oxford Textbook of Psychiatry*. Oxford: Oxford University Press.

GLICK, I.D., HARGREAVES, W.A. & GOLDFIELD, W.D. (1974). Short vs long hospitalization: a prospective controlled study. I. Preliminary results of one year follow-up of schizophrenics. *Archives of General Psychiatry*, **30**: 363-369.

GLICK, I.D., HARGREAVES, W.A., DRUES, J., & SHOWSTACK, J.A. (1976a). Short vs long hospitalization: a prospective controlled study. IV. One year follow-up results for schizophrenics patients. *American Journal of Psychiatry*, **133**: 509-514.

GLICK, I.D., HARGREAVES, W.A., DRUES, J., & SHOWSTACK, J.A. (1976b). Short vs long hospitalization: a prospective controlled study. V. One year follow-up results for non-schizophrenics patients. *American Journal of Psychiatry*, **133**: 515-517.

GLICK, I.D., HARGREAVES, W.A., DRUES, J., SHOWSTACK, J.A., & KATZOW, J.J. (1977). Short vs long hospitalization: a prospective controlled study. VII. Two year follow-up results for non-schizophrenics patients. *Archives of General Psychiatry*, **34**: 314-317.

GLOVER, G. (1990). The audit of mental health services. *Quality assurance in Health Care*, **2**: 181-188.

GOFFMAN, E. (1961). *Asylums: essays on the social situation of mental patients and other inmates*. Anchor Books, Doubleday & Co.

GOLDBERG, D.P., COOPER, B., EASTWOOD, M.R., KEDWARD, H.B., & SHEPHERD, M. (1970). A standardized psychiatric interview for use in community surveys. *British Journal of Preventive and Social Medicine*, **24**: 18-23.

GOLDESTEIN J.M. & HORGAN C.M. (1988). Inpatient and outpatient psychiatric services: substitutes or complements?. *Hospital and Community Psychiatry*, **39**: 632-636.

GOTTLIEB, J.F. & OLFSON, M. (1987) Current referral practices of mental health care providers. *Hospital and Community Psychiatry*, **38**: 1171-1181.

GRAD, J., & SAINSBURY, P. (1966). Evaluating the community psychiatric service in Chichester: results. **In**: E.M. Gruenberg (ed) *Evaluating the Effectiveness of Mental Health Service*, New York: Milbank Memorial Fund.

GRANT, R.L. (1981). The capacity of the psychiatric record to meet changing needs. **In**: Siegel, C. & Fischer, S. (eds) *Psychiatric Records in Mental Health Care*. New York: Brunner/Mazel.

GRANT, I. & KAPLAN, R.M. (1989). Statistics and experimental design. **In**: Kaplan, H.I. & Sadock. B.J. (eds) *Comprehensive Textbook of Psychiatry*, Vol. 1, 5th ed. Baltimore: Williams and Wilkins.

GREBB, J. (1989). Psychiatric Rating Scales. **In**: Kaplan, H.I. & Sadock. B.J. (eds) *Comprehensive Textbook of Psychiatry*, Vol. 1, 5th ed. Baltimore: Williams and Wilkins.

GREENE, L.R. & DE LA CRUZ. A. (1981). Psychiatric day treatment as alternative to and transition from full time hospitalization. *Community Mental Health Journal*, **17**: 191-202.

GRUER, R., GORDON, D.S., GUNN, A.A., & RUCKLEY, C.V. (1986). Audit of surgical audit. *The Lancet*, i: 23-26.

GUILLETE, W., CROWLEY, B., SAVITZ, S.A., & GOLDBERG, F.D. (1978). Day hospitalization as a cost effective alternative to inpatient care: a pilot study. *Hospital and Community Psychiatry*, **29**: 525-527.

GULLIFORD, M.C., PETRUCKEVITCH, A., & BURNEY, P.G.J. (1991). Hospital case notes and medical audit: evaluation of non-response. *British Medical Journal*, **302**: 1128-1129.

GUY, W., GROSS, M., HOGARTY, G.E., & DENNIS, H. (1969). A controlled evaluation of day hospital effectiveness. *Archives of General Psychiatry*, **20**: 329-338.

HAFNER, H. & KLUG, J. (1982). The impact of an expanding community mental health service on patterns of bed usage: evaluation of a four-year period of implementation. *Psychological Medicine*, **12**: 177-190.

HALL, J.N. (1974). Inter-rater reliability of ward rating scales. *British Journal of Psychiatry*, **125**: 248-255.

HAMILTON, M. (1968). Some notes on rating scales. *The Statistician*, **18**: 11-17.

HARGREAVES, W.A., GLICK, I.D., DRUES, J., SHOWSTACK, J.A., & FEIGENBAUN, E. (1977). Short vs long hospitalization: a prospective controlled study. VI. Two year follow-up results for schizophrenics patients. *Archives of General Psychiatry*, **34**: 305-311.

HEATH, D.A. (1990). Random review of hospital patient records. *British Medical Journal*, **300**: 651-652.

HERZ M.I., ENDICOTT J., SPITZER R.L. & MESNIKOFF A. (1971). Day versus inpatient hospitalisation: a controlled study. *American Journal of Psychiatry*, **127**: 1371-1382.

HERZ M.I., ENDICOTT J., SPITZER R.L. (1975). Brief hospitalization of patients with families: initial results. *American Journal of Psychiatry*, **132**: 413-418.

HERZ M.I., ENDICOTT J., SPITZER R.L. (1977). Brief hospitalization: a two year follow-up. *American Journal of Psychiatry*, **134**: 502-504.

HILL, D., BIRLEY, J.L.T., CAWLEY, R.H., KENDELL, R.E., LISHMAN, W.A., POST, F., RUTTER, M.L., WING, J.K., & WOLFF, H.H.(1973). *Notes on Eliciting and Recording Clinical Information*. Oxford: Oxford University Press.

HIRSCH, S.R., PLATT, S., KNIGHTS, A., & WEYMAN, A. (1979). Shortening hospital stay for psychiatric care: effect on patients and their families. *British Medical Journal*, **1**: 442-446.

HOGARTY G.E., DENNIS H., GUY W. & GROSS G.M. (1968). "Who goes there?" - A critical evaluation of admissions to a psychiatric day hospital. *American Journal of Psychiatry*, **124**: 934-943.

HOLLAND, W., ed (1983). *Evaluation of Health Care*. Oxford: Oxford University Press.

HOLMAN, C. (1989). Medical audit in psychiatry. *Psychiatric Bulletin*, **13**: 281-284.

HOPKINS, A. (1991). Approaches to Medical Audit. *Journal of Epidemiology and Community Health*, **45**: 1-3.

HOULT, J. (1986). Community care of the acutely mentally ill. *British Journal of Psychiatry*, **149**: 137-144.

HYLER, S.E., SKODOL, A.E., OLDHAM, J.M., KELLMAN, H.D. & DOIDGE, N.(1992). Validity of the Personality Diagnostic Questionnaire-Revised: a replication in an out-patient sample. *Comprehensive Psychiatry*, **33**: 73-77

JACOBSON, A., KOEHLER, J.E., JONES-BROWN, C. (1987). The failure of routine assessment to detect histories of assault as experienced by psychiatric patients. *Hospital and Community Psychiatry*, **38**: 386-389.

JENKINS, R. (1990). Towards a system of outcome indicators for mental health *British Journal of Psychiatry*, **157**: 500-514.

JOHN, K, GAMMON, G.D., & WEISSMAN, M.M. (1987). Assessment of psychosocial status: measures of subjective wellbeing, social adjustment and psychiatric symptoms. In: Abelin, T., Brzezinski, Z.J., & Carstairs, V.D.L (eds) *Measurement in Health Promotion and Protection*. Copenhague: World Health Organization.

JOHNSON J.C., KERSE, N.M., GOTTLIEB, G., WANICH, C. SULLIVAN, E., & CHEN, K. (1992). Prospective versus retrospective methods of identifying patients with delirium. *Journal of the American Geriatric Society*, **40**: 316-319.

JOHNSTONE E.C., OWENS, D.G.C., GOLD, A., CROW, T.J., & MACMILLAN, J.F. (1981). Institutionalization and the defects of schizophrenia. *British Journal of Psychiatry*, **139**: 195-203.

JONES, E. (1991). Audit in psychiatric: "failed discharges". *Psychiatric Bulletin*, **15**: 26-27.

KAPLAN, H.I. & SADOCK, B.J. (1989). Typical signs and symptoms of psychiatric illness. In: Kaplan, H.I. & Sadock, B.J. (eds) *Comprehensive Textbook of Psychiatry*, Vol. 1, 5th ed. Baltimore: Willians & Wilkins.

KECMANOVIC, D. (1985). Post release adjustment of day and inpatients. *International Journal of Social Psychiatry*, **31**: 74-79

KEENAN, B. (1975). Designing mental health studies: the pragmatics of non-experimental design. *Hospital and Community Psychiatry*, **26**: 739-740.

KENDEL, R.E., EVERRIT, B., COOPER, J.E., SARTORIUS, N. and DAVID, M.E. (1968). The reliability of the 'Present State Examination'. *Social Psychiatry*, **3**: 123-129.

KIERNAN, W.E.S., MCCREADIE, R.G., & FLANAGAN, W.L. (1976). Trainees' competence in psychiatric case writing. *British Journal of Psychiatry*, **129**: 167-172.

KIM, S.W., DYSKEN, M.W., & KUSKOVSKI, M. (1992). The Symptom Check List-90: Obsessive-Compulsive Subscale: a reliability and validity study. *Psychiatry Research*, **41**: 37-44.

KIRCHER, T., NELSON, J., & BURDO, H. (1985). The autopsy as a measure of accuracy of the death certificate. *The New England Journal of Medicine*, **313**: 1263-1269.

KIRKWOOD, B.R. (1988). *Medical Statistics*. London: Blackwell Scientific Publications.

KNIGHTS, A., HIRSCH, S.R, & PLATT, S.D.(1980). Clinical change as a function of brief admission to hospital in a controlled study using the Present State Examination. *British Journal of Psychiatry*, **137**: 170-180.

KOLAKOWSKA, T., WILLIAMS, A.O., ARDERN, M., REVELEY, M.A., JAMBOR, K., GELDER, M.G., & MANDELBROTE, M. (1985). Schizophrenia with good and poor outcome. *British Journal of Psychiatry*, **146**: 229-246.

KREITMAN, N. (1975). The use of clinical records in retrospective research. In: Sainsbury P. & Kreitman N. (eds) *Psychiatric Research*. Oxford: Oxford Medical Publications.

KRIS, E.B. (1965). Day hospitals. *Current Therapeutic Research*, **7**: 320-323 (cited by Greene and De La Cruz, 1981).

KUTCHER, S.P., WHITEHOUSE, A.M., FREEMAN, C.P.L. (1985). 'Hidden' eating disorders in Scottish inpatients. *American Journal of Psychiatry*, **142**: 1475-1478

LAKATOS, I. (1974). Falsification and the methodology of scientific research programmes. In: Lakatos, I & Musgrave, A. (eds) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press (cited by Chalmers, 1982).

LAPATRA, J.W. (1975). *Health Care Delivery Systems*. Springfield: Charles C. Thomas

LAST, J.M. (1988). *A Dictionary of Epidemiology*, 2nd ed. New York: Oxford University Press.

LAWRENCE, R.E., COPAS, J.B., & COOPER, P.W. (1991). Community care: Does it reduce the need for psychiatric bed? *British Journal of Psychiatry*, **159**: 334-340.

LEFF, J. (1992). Problems of Transformation. *International Journal of Social Psychiatry*, **38**: 16-23.

LELLIOTT, P. (1992). The College Research Unit and its clinical audit programme. *Psychiatric Bulletin*, **16**: 313-315.

LEON, R.L., BOWDER, C.L., & FABER, R.A. (1989). The psychiatric interview, history, and mental status examination. In: Kaplan, H.I. & Sadock, B.J. (eds) *Comprehensive Textbook of Psychiatry*, Vol. 1, 5th ed. Baltimore: Willians & Wilkins.

LILIENFIELD, A. M. & LILIENFIELD, D. E. (1980). *Foundations of Epidemiology*, 2nd ed. New York: Oxford University Press.

LIPTZIN, B. (1974). Quality assurance and psychiatric practice: a review. *American Journal of Psychiatry*, **131**: 1374-1377.

LITTLE, R.J.A.(1987). *Statistical Analyses with Missing Data*. New York: John Wiley & Sons.

LURIA, R.E. and BERRY, R. (1979). Reliability and validity of PSE syndromes. *Archives of General Psychiatry*, **36**: 1187-1195.

LYONS, T.F. & PAYNE, B.C. (1974). The relationship of physicians' medical recording performance to their medical care performance. *Medical Care*, **12**: 463-469.

LYSTAD, M.H. (1958). Day hospital care and changing family attitudes toward the mentally ill. *Journal of Nervous and Mental Disease*, **127**: 145-152.

MCDONALD, R. & MARKS, I.M. (1988). Quality assurance of outcome in mental health care: a model for routine use in clinical settings. *Health Trends*, **20**: 111-114.

MCGRATH, G. & TANTAM, D. (1987). Long-stay patients in a psychiatric day hospital: a case note review. *British Journal of Psychiatry*, **150**: 836-840.

MCGUFFIN, P., KATZ, R., & ALDRICH, J. (1986). Past and Present State Examination: the assessment of 'lifetime ever' psychopathology. *Psychological Medicine*, **16**: 461-465.

MCGUIRE, C.H. (1985). Medical problem solving: a critique of the literature. *Journal of Medical Education*, **60**: 587-595.

MCKEE, C.M., LAUGLO M. & LESSOF, L. (1989). Medical audit: a review. *Journal of the Royal Society of Medicine*, **82**: 474-478.

MCSHANE, R., ROWE, D., & JULIER, D. (1992). Will the information recorded in psychiatric notes change when patients have right to read them? *Psychiatric Bulletin*, **16**: 404-405.

MANCHANDA, R. & HIRSCH, S.R. (1986). Rating scales for clinical studies in schizophrenia. **In:** Bradley, P.B. & Hirsch, S.R. (eds) *The Psychopharmacology and Treatment of Schizophrenia*. Oxford: Oxford University Press.

MARK, M., RABINOWITZ, J., KINDLER, S., RABINOWITZ, S., MUNITZ, H., & BLEICH, A. (1991). A system for improving psychiatric record keeping. *Hospital and community Psychiatry*, **42**: 1163-1166.

MARKANTONAKIS, A. & WEIR, I.K. (1991). Medical audit quality of note-keeping. *Psychiatric Bulletin*, **15**: 283-284.

MECHANIC D. & AIKEN L.H. (1987). Improving the care of patients with chronic mental illness. *The New England Journal of Medicine*, **317**: 1634-1638.

MENTAL HEALTH ACT (1983). London: Her Majesty Stationery Office.

MENTAL HEALTH FOUNDATION (1989). *Mental Illness: The Fundamental Facts*. London: Mental Health Foundation.

MEYER, J.S. & PROVINCE, M. (1988). Proliferative index of breast carcinoma by thymidine labelling: prognostic power independent of stage, oestrogen and progesterone receptors. *Breast Cancer Research and Treatment*, **12**: 191 (cited by Altman, 1991).

MICHAUX M.H., CHELST, M.R., FOSTER, S.A, & PRIUM , R.J. (1972). Day and full time psychiatric: a controlled comparison. *Current Therapeutic Research*, **14**: 279-292 (cited by Michaux et al, 1973).

MICHAUX M.H., CHELST M.R., FOSTER S.A., PRIUM, R.J., & DASINGER, E.M. (1973). Postrelease adjustment of day and full-time psychiatric patients. *Archives of General Psychiatry*, **29**: 647-651.

MITCHELL, M.W. & FOWKES, F.G.R. (1985). Audit reviewed: does feedback on performance change clinical behaviour? *Journal of the Royal College of Physicians of London*. **19**: 251-254.

MONTGOMERY, S., ASBERG, M., JORNESTEDT, L., THOREN, P., TRASKMAN, L., MCAULEY, R., MONTGOMERY, D., & SHAW, P. (1978). Reliability of the CPRS between the disciplines of psychiatry, general practice, nursing and psychology in depressed patients. *Acta Psychiatrica Scandinavica*, Supplement, 271.

MORLEY,S. & SNAITH, P. (1989). Principles of psychological assessment. In: Freeman, C. & Tyrer, P. (eds) *Research Methods in Psychiatry*. London: Royal College of Psychiatrists.

MOSHER, L.R., MENN, A. & MATHEWS, S.M. (1975). Soteria: Evaluation of a home-based treatment for schizophrenia. *American Journal of Orthopsychiatry*, **45**: 455-467.

MOSHER, L.R. (1983). Alternatives to psychiatric hospitalization. *The New England Journal of Medicine*, **309**: 1579-1580.

195

NORRIS, V. (1959). *Mental Illness in London*. London: Oxford University Press

NORUSIS, M.J. (1988). *SPSS/PC + V2.0 Base Manual for the IBM PC/XT/AT and PS/2*. Chicago: S.P.S.S. Inc.

NUNNALY, J.C. (1978). *Psychometric Theory*. New York: McGraw-Hill, Inc.

OVERALL, J.E. & GORHAN, D.R. (1962). The Brief Psychiatric Rating scale. *Psychological Report*, **10**: 799-812.

PANG, J. (1985). Partial Hospitalization. *Psychiatric Clinics of North America*, **8**: 587-595.

PASAMANICK, B., SCARPITTI, F.R. & DINITZ, S. (1967). *Schizophrenics in the community: an experimental study in the prevention of hospitalization*. New York: Appleton-Century-Crofts.

PENK, W.E., CHARLES, H.L. & VAN HOOSE, T.A.(1978). Comparative effectiveness of day hospital and inpatient psychiatric treatment. *Journal of Consulting and Clinical Psychology*, **46**: 94-101.

PETERS, B.S., BECK, E.J., COLEMAN, D.G., WADSWORTH, M.J.H., MCGUINESS, O., HARRIS, J.R.W., & PINCHING, A..J (1991). Changing disease patterns in patients with AIDS in a referral centre in the United Kingdom: the changing face of AIDS. *British Medical Journal*, **302**: 203-207.

PERCY, C., STANEK, E., & GLOECKLER, L. (1981). Accuracy of cancer death certificates and its effects on cancer mortality statistics. *American Journal of Public Health*, **71**: 242-250.

PHILIPS, E.M. & PUGH, D.S. (1987). *How to Get a PhD*. Milton Keynes: Open University Press.

PLATT, S., WEYMAN, A., HIRSCH, S. & HEWETT, S. (1980). The Social Behaviour Assessment Schedule (SBAS): rationale, contents, scoring and reliability of a new interview schedule. *Social Psychiatry*, **15**: 43-55.

POLACK, P.R. & KIRBY, M.W. (1976). A model to replace psychiatric hospitals. *The Journal of Nervous and Mental Disease*, **162**: 13-22.

PORGOURIDES C., PRASHER, V.P., & OYEBODE, F. (1992). Use of section 5(2) in clinical practice. *Psychiatric Bulletin*, **16**: 14-16.

POTOMIANOS, G., NORTH, W.R.S., MEADE, T.W., TOWNSED, J. & PETERS, T.J. (1986). Randomised trial of community-based centre versus conventional hospital management in treatment of alcoholism. *The Lancet*, **ii**: 797-799.

POUNDER, R.E. (1991). Audit in practice: some reports from those who have tried. *Journal of the Royal College of Physicians of London*, **25**: 339-340.

REGIER, D.A. & BURKE, J.D. (1989). Epidemiology. In: Kaplan, H.I. & Sadock, B.J. (eds) *Comprehensive Textbook of Psychiatry* Vol. 1, 5th ed: Baltimore: Williams and Wilkins.

REMINGTON, M., TYRER, P., NEWSON-SMITH, J., & CICCHETI, D.V. (1979). Comparative reliability of categorical and analogue rating scales in the assessment of psychiatric symptomatology. *Psychological Medicine*, **9**: 765-770.

RICE, C.E., BERGER, D.G., SEWALL, L.G. & LEMKAU, P.V. (1961). Measuring social restoration performance of public psychiatric hospitals. *Public Health Reports*, **76**: 437-446.

ROBINS, L.N., HELZER, J.E., RATCLIFF, K.S., & SEYFRIED, W. (1982). Validity of the Diagnostic Interview Schedule, version II: DSM-III diagnoses. *Psychological Medicine*, **12**: 855-870.

ROBINSON, M. (1991). Medical audit: basic principles and current methods. *Psychiatric Bulletin*, **15**: 21-23.

ROLAND, M. & MORRIS, R. (1988). Are referral by general practitioners influenced by the availability of consultants? *British Medical Journal*, **297**: 599-600.

ROSIE J.S. (1987). Partial hospitalization: a review of recent literature. *Hospital and Community Psychiatry*, **38**: 1291-1298.

ROSSI, P.H. & FREEMAN, H.E. (1989). Evaluation: a systematic approach, 4th ed. Newbury Park: Sage Publications

ROY, D. (1991). Setting up district audit meetings in psychiatry. *Psychiatric Bulletin*, **15**: 417-418.

ROYAL COLLEGE OF PHYSICIANS OF LONDON (1978). Death certification and epidemiological research. *British Medical Journal*, **2**: 1063-1065.

ROYAL COLLEGE OF PSYCHIATRISTS (1989). Preliminary report on medical audit. *Psychiatric Bulletin*, **13**: 577-580.

RUBENSTEIN, L.V., KAHN, K.L., HARRISON, E.R., SHERWOOD, M.J., ROGERS, W.H., & BROOK, R. (1991). *Structured Implicit Review of the Medical Record*. Santa Monica: RAND publications.

RUSSEL, I.T. & WILSON, B.J. (1992). Audit: the third clinical science? *Quality in Health Care*, **1**: 51-55.

SAPPINGTON, A.A. & MICHAUX, M.H. (1975). Prognostic patterns in self-report, relative report, and professional evaluation measures for hospitalized and day-care patients. *Journal of Consulting and Clinical Psychology*, **43**: 904-910.

SARTORIUS, N. & HARDING, T.W. (1983). Issues in the evaluation of mental health care. In: Holland, W. ed, *Evaluation of Health Care*. Oxford: Oxford University Press.

SCHENE, A.H. & GERSONS, B.P.R. (1986). Effectiveness and application of partial hospitalization. *Acta Psychiatrica Scandinavica*, **74**: 335-340.

SCHENE, A.H., HENDERSON, J.H., KNUDSEN, H.C., RIJKSCHROEFF, R., & THORNICROFT, G. (1992). The evaluation of mental health care transformation in the cities of Europe. *The International Journal of Social Psychiatry*, **38**: 40-49.

SCHMID, W., BRONISH, T., & ZERSSEN, D.V. (1982). A comparative study of PSE/CATEGO and DiaSiKa: two psychiatric computer diagnostic systems. *British Journal of Psychiatry*, **141**: 292-295.

SECRETARY OF STATE FOR HEALTH (1992). *The Health of the Nation: a strategy for health in England.* London: HMSO.

SHAW, C.D. & COSTAIN, D.W. (1989). Guidelines for medical audit: seven principles. *British Medical Journal*, **299**: 498-499.

SHAW, C.D. (1989). Clinical outcome indicators. *Health Trends*, **21**: 37-40.

SHAW, C.D. (1990). Criterion based audit. *British Medical Journal*, **300**: 649-651.

SHEPHERD, M. (1957). *A Study of the Major Psychoses in an English County.* London: Oxford University Press.

SIEGEL, C. & FISCHER, S.K. (eds) (1981). *Psychiatric Records in Mental Health Care.* New York: Brunner/Mazel.

SMITH, S. & CROSS, E. (1957). Review of 1000 patients treated at a psychiatric day hospital. *International Journal of Social Psychiatry*, **2**: 292-298 (cited by Herz et al, 1971).

SMITH, H.E., RUSSELL, G.I., FREW, A.J., & DAWES, P.T. (1992). Medical audit: the differing perspectives of managers and clinicians. *Journal of the Royal College of Physicians*, **26**: 177-80.

SNAITH, R.P. (1981). Rating Scales. *British Journal of Psychiatry*, **138**: 512-514

SNAITH, R.P. (1991). Measurement in psychiatry. *British Journal of Psychiatry*, **159**: 78-82.

SOREFF, S., GULKIN, T., & PIKE, J.G. (1990). The evolving clinical chart. *Psychiatric Clinics of North America*, **13**: 127-133.

SPITZER R.L., ENDICOTT J. & FLEISS J.L. (1967). Instruments and recording forms for evaluating psychiatric status and history: rationale, method of development and description. *Comprehensive Psychiatry*, **8**: 321-343.

STEIN, L.I., TEST, M.A. & MARX, A.J. (1975). Alternative to the hospital: a controlled study. *American Journal of Psychiatry*, **132**: 517-522.

STEIN, L.I. & TEST, M.A. (1980). Alternative to mental hospital treatment. *Archives of General Psychiatry*, **37**: 392-397.

STERN, G., COTTREL, D., & HOLMES, J. (1990). Patterns of attendance of child psychiatric out-patients with special reference to Asian families. *British Journal of Psychiatry*, **156**: 384-387.

TAINTOR, Z. (1981). The use of psychiatric records in training psychiatrists. In: Siegel, C. & Fischer, S. (eds) *Psychiatric Records in Mental Health Care*. New York: Brunner/Mazel.

TAIT, I. (1981). History of our records. *British Medical Journal*, **282**: 702-704.

TALBOT J.A. (1985). Community care for the chronically mentally ill. *Psychiatric Clinics of North America*, **8**: 437-448.

TANSELLA, M. (1986). Community Psychiatry without mental hospitals - the italian experience: a review. *Journal of the Royal Society of Medicine*, **79**: 664-669.

TANSELLA, M. (1989). Evaluating community psychiatric services. In: Willians P., Wilkinson G., & Rawnsley, K. (eds) *The Scope of Epidemiology Psychiatry*. London: Routledge.

TANTAM, D. (1985). Alternatives to psychiatric hospitalisation. *British Journal of Psychiatry*, **146**: 1-4.

TEST, M.A. & STEIN, L.I. (1978). Community treatment of the chronic patient: research overview. *Schizophrenia Bulletin*, **4**: 350-364.

THOMPSON, G.E. & HANDELMAN, I. (1978). *Health Data and Information Management*. Boston: Butterworths.

THOMPSON, C. (1989). *The Instruments of Psychiatric Research*. Chichester: John Wiley & Sons.

THORNICROFT, G & BEBINGTON, P. (1989). Deinstitutionalisation: from hospital closure to service development: a review of the literature. *British Journal of Psychiatry*, **155**: 739-755.

THORNICROFT, G. & STRATHDEE, G. (1991). Mental Health. *British Medical Journal*, **303**: 410-412.

TRESS, K.H., BELLENIS, C., BROWNLOW, J.M., LIVINGSTON, G. and LEFF, J.P. (1987). The Present State Examination Change Rating Scale, *British Journal of Psychiatry*, **150**, 201-207.

TYRER P.J. & REMINGTON M. (1979). Controlled comparison of day-hospital and out patient treatment for neurotic disorders. *The Lancet*, **ii**: 1014-1016.

TYRER P., REMINGTON M., & ALEXANDER J. (1987). The outcome of neurotic disorders after out-patient and day hospital care. *British Journal of Psychiatry*, **151**: 57-62.

VAUGHAN P.J. (1983). The disordered development of day care in psychiatry. *Health Trends*, **15**: 91-94.

VORT, W.V. & MATTSON, M.R. (1989). A strategy for enhancing the clinical utility of the psychiatric record. *Hospital and Community Psychiatry*, **40**: 407-409.

VUORI, H. (1989). Research needs in quality assurance. *Quality Assurance in Health Care*, **1**: 147-159.

VUORI, H. & ROGER, F. (1989). Issues in quality assurance - the European scene. *Quality assurance in Health Care*, **1**: 125-135.

WALLACE C.H. & HASS B.J. (1983). Assessing outcome in chronic populations treated in day-treatment programs. In: Lambert, M.J. & Christensen E.R., Dejulio S.S. (eds) *The Assessment of Psychotherapy Outcome*, New York: Wiley.

WARREN, M.D. (1979). Evaluation as a tool in health planning and management. In: Hobson, W. ed, *The Theory and Practice of Public Health*. Oxford: Oxford University Press.

WASHBURN, S., VANNICELLI, M., LONGABAUGH, R. & SCHEFF, B.J. (1976). A controlled comparison of psychiatric day treatment and inpatient hospitalization. *Journal of Consulting and Clinical Psychology*, **44**: 665-675.

WEISSMAN, M.M. (1975). The assessment of social adjustment:a review of techniques. *Archives of General Psychiatry*, **32**: 357-365.

WILDER J.F., LEVIN G. & ZWERLING I. (1966). A two-year follow-up evaluation of acute psychotic patients treated in a day hospital. *American Journal of Psychiatry*, **122**: 1095-1101.

WILKINSON, G. (1984). Day care for patients with psychiatric disorders. *British Medical Journal*, **288**: 1710-1712.

WILKINSON, G. (1989). Referrals from general practitioners to psychiatrists and paramedical mental health professionals. *British Journal of Psychiatry*, **154**: 72-76.

WILLIAMS, J.G., KINGHAM, M.J., MORGAN, J.M., & DAVIES, A.B. (1990). Retrospective review of hospital patient records. *British Medical Journal*, **300**: 991-993.

WING, J.K. (1960). Pilot experiment in the rehabilitation of long-hospitalized male schizophrenic patients. *British Journal of Preventive and Social Medicine*, **14**: 173-180.

WING, J.K., BIRLEY, J.L.T., COOPER, J.E., GRAHAM, P. & ISAACS, A. (1967). Reliability of a procedure for measuring and classifying 'Present Psychiatric State'. *British Journal of Psychiatry*, **113**: 499-515.

WING, J.K. (1972). Principles of evaluation. **In:** Wing, J.K. & HAILEY, A.M. (eds) *Evaluating a Community Psychiatric Service:* The Camberwell Register. London: Oxford University Press.

WING, J.K., COOPER, J.E. & SARTORIUS, N. (1974). *The Measurement and Classification of Psychiatric Symptoms,* London: Cambridge University Press.

WING, J.K., NIXON, J.M., MANN, S.A., & LEFF, J.P. (1977). Reliability of the PSE (ninth edition) used in a population study. *Psychological Medicine,* **7:** 505-516.

WING, J.K. (1978). *Reasoning about Madness.* London: Oxford university Press.

WING, J.K. & STURT, E. (1978). *The PSE-ID-CATEGO System Supplementary Manual.* London: MRC-Psychiatric Unit

WING, J.K. (ed) (1982). Long-term community care: experience in a London Borough. *Psychological Medicine,* Monograph Supplement 2.

WING, J.K. (1983). Use and misuse of the PSE. *British Journal of Psychiatry,* **143:** 111-117.

WING, J.K. (1989). The measurement of 'social disablement'. *Social Psychiatry and Psychiatric Epidemiology,* **24:** 173-178.

WORLD HEALTH ORGANIZATION (1973). *The International Pilot Study of Schizophrenia.* Geneva: WHO.

WORLD HEALTH ORGANIZATION (1988). *The Psychiatric Disability Assessment Schedule.* Geneva: WHO

WYKES, T. and 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,* **148:** 1-11.

ZUSMAN, J. (1988). Quality assurance in mental health care. *Hospital and Community Psychiatry*, **39**: 1287-1290.

APPENDICES

APPENDIX A

EHICAL PERMISSION AND CONSULTANT APPROVAL

UNIVERSITY COLLEGE &
MIDDLESEX SCHOOL OF MEDICINE
UNIVERSITY COLLEGE LONDON

UCMSM
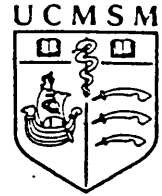
Dean of the School of Medicine and of
the Faculty of Clinical Sciences
SIR WILLIAM SLACK, K.C.V.O., M.A., M.CH., F.R.C.S.

Dean, Faculty of Medical Sciences
PROFESSOR A.P. MATHIAS, B.Sc., PH.D.

Secretary
DIANA F. SANDERS, M.A., PH.D.

Please reply to: School of Medicine
Administrative Offices
Riding House Street
London W1P 7PN
Telephone: 01-636 8333

(From 6 May 1990: 071-636 8333)
Ext.
Fax: 01-436 0184
(From 6 May 1990: 071-436 0184)

Ext 9383

24 April 1990

Dr M McCarthy
Senior Lecturer
Dept of Community Medicine
UCL

Dear Dr McCarthy,

Joint UCH/UCL Committee on the Ethics of Clinical Investigations

Submission No : 2227

Title : Outcomes of mental health care in Bloomsbury.

On the information received the Chairman of the Joint
UCH/UCL Committee on the Ethics of Clinical Investigations has
given Chairman's approval for the above study.

Yours sincerely

DCutler.

Debbie Cutler
Administrative Officer

THE UNIVERSITY COLLEGE AND MIDDLESEX
SCHOOL OF MEDICINE

Dr David P. Frost
Medical Knowledge-Based Systems Unit

Tel: 071-380 9472

Wolfson Building
The Middlesex Hospital
Riding House Street
London W1N 8AA

Medical Record Officer
UCH and Middlesex Hospitals

12 February 1991

Dear Sir,
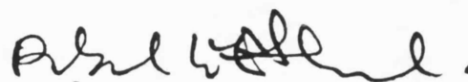
Could you please allow Dr Marco Akerman to have access to notes of patients who
have been admitted to Laffan Ward. He is conducting some research with Dr Mark
McCarthy on outcome measures for psychiatric patients. The research project
has been given Ethical Committee approval.

I have discussed the research with Dr Akerman and I am happy for him to have
access to notes from Laffan Ward. Any queries can be addressed to myself.

Yours sincerely

Dr David P. Frost
Senior Lecturer
Honorary Consultant Psychiatrist, Laffan Ward

R.D. LITTLEWOOD
READER
HON. CONSULTANT PSYCHIATRIST, LAFFAN
WARD)

2/9/91

ST. LUKE'S WOODSIDE HOSPITAL
WOODSIDE AVENUE
MUSWELL HILL
LONDON N10 3HU

25th March 1991

Dr. M. Akerman,
UCMSM,
Department of
Community Medicine,
66-72, Gower Street,
London WC1E 6EA.

Dear Dr. Akerman,

Further to your letter of 13th March 1991 to Dr. Gurling,
and our telephone conversation, I am able to write that
Dr. Gurling gives his permission for you to collect
data from his patients on Noel Harris Ward for your
research project.

Yours sincerely,

RButler

Dr. R.E. Butler,
SHO to Dr. H. Gurling,
Consultant Psychiatrist: Noel Harris Ward.

BLOOMSBURY HEALTH AUTHORITY

ST. PANCRAS HOSPITAL

4 ST. PANCRAS WAY,

LONDON,

NW1 0PE

Telephone: 071-387 4411   EXT

MG.40

11 April 1991

Medical Records

TO WHOM IT MAY CONCERN

I, Dr Maurice Greenberg, Consultant Psychiatrist and Clinical
Tutor in Psychiatry, give Mr Marco Ackerman, Department of
Community Medicine, University College Hospital, permission of
access to patients' medical notes who attended the Jules Thorn
Day Hospital in 1988 - 1989.

Yours faithfully

PP *Hlarty*

Dr Maurice Greenberg
Consultant Psychiatrist/Clinical Tutor in Psychiatry
JULES THORN DAY HOSPITAL

210

Bloomsbury Health Authority

ST. LUKE'S WOODSIDE HOSPITAL
WOODSIDE AVENUE
MUSWELL HILL
LONDON N10 3HU

10th May, 1991.

Dr. Marco Akerman
Dept.of Community Medicine
University College Hospital
66-72, Gower Street
London, WC1E 6EA.

Dear Dr. Akerman,

With reference to your request, I confirm that you may have
access to my patients' relevant information on Noel Harris Ward,
St. Luke's-Woodside Hospital, for your research.

Yours sincerely,

Dr. O.W. Hill, FRCP, FRCPsych.
Consultant   Psychiatrist

BLOOMSBURY HEALTH AUTHORITY

UNIVERSITY COLLEGE &
MIDDLESEX SCHOOL OF MEDICINE
DEPARTMENT OF PSYCHIATRY

Head of Department
Professor R. M. Rosser

Wolfson Building
Middlesex Hospital
Riding House Street
London W1N 8AA
Tel: 071-380 9468
Fax: 071-323 1459

28 May 1991.

Mr. Marco Akerman,

Department of Community Medicine,

Fourth Floor,

University College London,

66-72 Gower Street,

London WC1 ECA.

Dear Dr. Akerman,

This is to confirm that I have no objection to your scrutinising the files of my patients for your research project. The Psychiatry Advisory Committee also supports your work.

Yours sincerely,

Dr. Stuart Turner,

Senior Lecturer in Psychiatry.

APPENDIX B

Pilot Study

B.1. Objectives

The following objectives were set to be achieved by the Pilot Study:

- to start contact with medical records offices within the target area (South Camden area - Bloomsbury and Islington Health Authority, London).

- to assess characteristics of case notes from different modalities of care (out-patient, in-patient and day-hospital).

- to test the coding sheet for collecting demographic and clinical variables.

- to identify areas within the case notes for collecting symptom information and social/behaviour items on admission and to test the coding sheet for assembling these data.

- to examine the linkage existing between information recorded on admission and on discharge.

B.2. Sample selection

Case notes (N=37) were selected at random (table of random numbers) from the filing cabinets of the respective medical records offices, day-hospital (N=6), 2 in-

patient units (N=21) and out-patient service (N=10). Patients admitted in 1988 and 1989 were considered. After applying exclusion criteria (described in Chapter 3, topic 3.1.3.) we were left with the study sample of 25 case notes.

B.3. Main conclusions

As the result of the Pilot Study the main conclusions were:

- not to include in the main study out-patient case notes because little information was reported on each consultation. Defining the points of admission and discharge for out-patient treatment were found to be difficult.

- to drop from the collection the variables 'ethnic group', 'type of prior treatment' and, 'months of previous hospitalisations' due to low reporting rate.

- to include 3 new variables - 'duration of present illness', 'current employment status' and, 'alcohol and drug problems' - with reasonable reporting rate.

- to change 2 categorical variables ('onset of illness' and 'prior hospitalisation') into numerical ones ('age of first psychiatric illness' and 'number of previous admissions' respectively) to increase statistical power of the variable.

- to drop from the symptom coding sheet the heading of the MSE 'memory and

cognition' and the symptoms related to these two attributes, 100 (dissociative states), 101 (conversion), 102 (clouding and stupor), 103 (organic impairment of memory). It was found very difficult to match these symptoms with reporting under this heading.

- to drop from the coding sheet the symptoms 94 (evasiveness), 139 (misleading answers), 140 (interview doubtfully adequate) - there is no reporting in the case notes concerned with these 'symptoms'.

- to drop symptom 39 (pre-menstrual exacerbation). The fact that this symptom do not apply to males would create two different denominators for males and females (total numbers of symptoms considered) when calculation of proportions is carried out.

- to join symptoms 35 (delayed sleep) and 37 (early waking) into a new symptom 35/37 called sleep disturbance since case note reporting occasionally discriminate the individuals symptoms.

- to consider for rating symptom 34 not only loss of weight but also the general reporting of appetite disturbance.

_ to drop from the collection PSE symptoms number 5,6,7,8,10,11,14,15,16,17 (general symptoms and autonomic anxiety) due to the high percentage of non-

reported information (99%).

- to collect symptom information on admission from the following 'areas' or 'sections' within the case notes - 'history of present complaint', 'mental state examination' (MSE) at admission and 'nursing notes' on admission.

- to collect behaviour items from 'nursing notes' on admission.

- to collect social data from the section 'current social situation' and 'nursing notes'.

- to collect discharge information from items recorded at admission from 'clinical notes', 'nursing reports' and from the section 'treatment and progress'. It was observed that in these sections there is general information about patients mental state at discharge and specific information for relevant symptoms, social and behaviour items reported as present on admission.

APPENDIX C

INFORMATION COLLECTED ON ALL PATIENTS ADMITTED IN 1990 TO THE THREE SETTINGS

DATA FROM ALL PATIENTS ADMITTED IN 1990

SETTING   ---------

PATIENT NAME   ------------------

PATIENT CODE NUMBER   ----------

DATE OF ADMISSION   ----/----/-----

DATE OF DISCHARGE   ----/----/----

CONSULTANT NAME   ---------------

REFERRAL AGENCY   ------------------

SEX   ---------

AGE   -----(D.O.B.---------

ADDRESS   ------------------

DIAGNOSIS   --------------------

DISCHARGE CONDITIONS   ----------------------

OTHER NOTES:

APPENDIX D

THE SYNDROME CHECK LIST FORM

SYNDROME CHECK LIST (April, 1973)

(9th Edition of PSE, April, 1973)

The syndromes formed out of the 140 symptoms rated in the 9th edition of the PSE are specified below, together with their symptom - composition. The presence of each syndrome may be directly rated from case notes, interviews or other material concerning the clinical history. Syndromes may be rated on the basis of the whole clinical history, each separate episode of illness, or the present episode. The period involved should always be specified. For the purposes of the CATEGO computer program, it is only necessary to rate syndromes 1 - 25. The others may, of course, be needed for other purposes.

When rating from case-records, always rate conservatively, Do not rate a syndrome present unless there is good evidence to this effect. For example, "patient feels influenced" would not be sufficient evidence for 1. (NS). Write down an example whenever a positive rating is made. Use X (Not known) freely. (Y = Not applicable).

Syndrome No.	Syndrome Name	Symptoms	
1. (NS)	Nuclear Syndrome	55. Thought intrusion 56. Thought broadcast 57. Thought commentary 58. Thought withdrawal	62. Voices about patient 71. Delusions of control 81. Delusions of alien penetration
		0 No symptoms 1 NS? = Partial delusions only 2 NS+ = 1 symptom 3 NS++ = 2+ symptoms	
2. (CS)	Catatonic Syndrome	116. Mannerisms and posturing 119. Catatonic movements	
		0 No symptoms 2 CS+ = 1 symptom 3 CS++ = 2 symptoms	

221

Syndrome No.	Syndrome Name	Symptoms
3. (IS)	Incoherent Speech	135. Neologisms 136. Incoherence of speech 0 No symptoms 2. IS+ = 1 symptom 3 IS++ = 2 symptoms
4. (RS)	Residual Syndrome	60. (2) Hears muttering, whispering 118. Behaves as if hallucinated 132. Non-social speech 0 No symptoms 2 RS+ = 1 symptom 3 RS++ = 2+ symptoms
5. (DD)	Depressive Delusions and Hallucinations	61. (2) Depressive hallucinations 88. Delusions of guilt 91. Hypochondriacal delusions (bowels blocked up) 92. Delusions of catastrophe 0 No symptoms 2 DD+ = 1 symptom 3 DD++ = 2+ symptoms
6. (SD)	Simple Depression	19. Inefficient thinking 23. Depressed mood 24. Hopelessness 25. Suicidal plans or acts 121. Depression on examination 0 No symptoms 1 SD? = If no symptom 23 or 121 2 SD+ = 1 symptom (23 or 121) 3 SD++ = 2+ symptoms (including 23 or 121)

222

Syndrome No.	Syndrome Name	Symptom
7. (ON)	Obsessional Neurosis	44. Checking and repeating
		45. Cleanliness and rituals
		46. Obsessional ideas and rumination
		0 No symptoms
		2 ON+ = 1 symptom
		3 ON++ = 2+ symptoms
8. (GA)	General Anxiety	11. Anxiety
		14. Panic attacks
		120. Anxiety on examination
		0 No symptoms
		2 GA+ = 1 symptom
		3 GA++ = 2+ symptoms
9. (SA)	Situational Anxiety	15. Situational anxiety
		17. Specific phobias
		18. Anxiety avoidance
		0 No symptoms
		2 SA+ = 1 symptom
		3 SA++ = 2+ symptoms
10. (HT)	Hysteria	64. (2) Dissociative hallucinations (Not sub-cultural)
		100. Dissociative states
		101. Conversion symptoms
		122. Histrionic
		0 No symptoms
		2 HT+ = 1 symptom
		3 HT++ = 2+ symptoms
11. (AF)	Flattening	128. Blunted affect
		0 No symptoms
		2 AF+ = 1 symptom

Syndrome No.	Syndrome Name	Symptoms
12. (HM)	Hypomania	41. Subjective euphoria
		42. Ideomotor pressure
		43. Grandiose ideas and actions
		123. Hypomanic affect
		137. Hypomanic content of speech
		0 No symptoms
		1 HM? = If only symptoms 42, 43 or 137
		2 HM+ = 1 symptom (41 or 123)
		3 HM++ = 2+ symptoms (including 41 or 123)
13. (AH)	Auditory Hallucinations	63. Voices to patient (not depressive)
		0 No symptoms
		2 AH+ = 1 symptom
14. (PE)	Delusions of Persecution	74. Delusions of persecution
		0 No symptoms
		1 PE? = Partial delusions only
		2 PE+ = 1 symptom
15. (RE)	Delusions of Reference	72. Delusions of reference
		73. Delusions of mis-interpretation
		0 No symptoms
		1 RE? = Partial delusions only
		2 RE+ = 1 symptom
		3 RE++ = 2 symptoms
16. (GR)	Grandiose and Religious Delusions	76. Delusions of grandiose ability
		77. Delusions of grandiose identity
		78. Religious delusions
		0 No symptoms
		1 GR? = Partial delusions only
		2 GR+ = 1 symptom
		3 GR++ = 2+ symptoms

Syndrome No.	Syndrome Name	Symptoms

17. (SF)	Sexual and Fantastic Delusions	59. Thoughts read .
		 70.(2) Delusional elaboration of hallucinations
		 75. Delusions of assistance
		 79. Delusional explanation (hypnotism, etc.)
		 80. Delusional explanation (rays, etc.)
		 84. Morbid jealousy
		 85. Delusions of pregnancy
		 86. Sexual delusions
		 87. Fantastic delusions
		 89. Delusions concerning appearance
		 90. Delusions concerning lack of organs
		 82. Primary delusions
		    · 0 No symptoms
		
		     1 SF? = Partial delusions only
		
		     2 SF+ = 1 symptom
		
		     3 SF++ = 2+ symptoms

18. (VH)	Visual Hallucinations	66. (2) Visual hallucinations 0 No symptoms 2 VH+ = 1 symptom
19. (OH)	Olfactory Hallucinations	68. Olfactory hallucinations 69. Delusion that patient smells 0 No symptoms 2 OH+ = 1 symptom 3 OH++ = 2 symptoms
20. (OV)	Overactivity	112. Gross excitement 113. Irreverent behaviour 115. Embarrassing behaviour 0 No symptoms 2 OV+ = 1 symptom 3 OV++ = 2+ symptoms

Syndrome No.	Syndrome Name	Symptoms
21. (SL)	Slowness	110. Slowness and underactivity
		130. Slow speech
		133. Muteness
		134. Restriction of quantity of speech
		0 No symptoms
		2 SL+ = Any other symptoms but muteness
		3 SL++ = Muteness (133)

22. (NP)	Non-specific Psychosis	49. Unfamiliarity and delusional mood
		50. Heightened perception
		52. Changed perception
		53. Changed perception of time
		60.(1) Hears music, tapping, etc.
		61.(1) Hears voice calling name
		66.(1) Minor visual hallucinations
		70. (1) Other minor hallucinations
		94. Evasiveness concerning delusions
		102. Clouding or stupor
		109. Bizarre appearance
		117. Stereotypies
		125. Suspicion
		126. Perplexity
		129. Incongruous affect
		0 No symptoms
		1 NP? = 1 symptom
		2 NP+ = 2 symptoms
		3 NP++ = 3+ symptoms

23. (DE)	Depersonalisation	47. Derealisation
		48. Depersonalisation
		0 No symptoms
		2 DE+ = 1 symptom
		3 DE++ = 2 symptoms

226

Syndrome No.	Syndrome Name	Symptoms

24. (ED)	Special Features of Depression	29. Self-depreciation
		32. Guilty ideas of reference
		33. Guilt
		51. Dulled perception
		54. Lost affect
		
		    0 No symptoms
		
		    2 ED+ = 1 symptom
		
		    3 ED++ = 2+ symptoms

25. (AG)	Agitation	111. Agitation on examination
		
		    0 No symptoms
		
		    2 AG+ = 1 symptom

26. (NG)	Self-neglect	108. Self neglect
		
		    0 No symptoms
		
		    2 NG+ = 1 symptom

27. (IR)	Ideas of Reference	31. Ideas of reference
		
		    0 No symptoms
		
		    2 IR+ = 1 symptom

28. (TE)	Tension	5. Tension pains
		7. Muscular tension
		8. Restlessness
		
		    0 No symptoms
		
		    2 TE+ = 1 symptom
		
		    3 TE++ = 2+ symptoms

29. (LE)	Lack of energy	36. Subjective anergia
		
		    0 No symptoms
		
		    2 LE+ = 1 symptom

Syndrome No.	Syndrome Name	Symptoms	
30. (WO)	Worrying, etc.	4. Worrying	
		6. Tiredness	
		10. Nervous tension	
		21. Neglect through brooding	
		35. Delayed sleep	
		0 No symptoms	
		2 WO+ = 1 symptom	
		3 WO++ = 2+ symptoms	☐
31. (IT)	Irritability	40. Irritability	
		124. Hostile irritability	
		0 No symptoms	
		2 IT+ = Symptom 40.	
		3 IT++ = Symptom 124.	☐
32. (SU)	Social Unease	16. Anxiety on meeting people	
		28. Social withdrawal	
		30. Lack of self confidence	
		0 No symptoms	
		2 SU+ = 1 symptom	
		3 SU++ = 2+ symptoms	☐
33. (IC)	Loss of interest and concentration	20. Poor concentration	
		22. Loss of interest	
		0 No symptoms	
		2 IC+ = 1 symptom	
		3 IC++ = 2 symptoms	☐
34. (HY)	Hypochondriasis	9. Hypochondriasis	
		0 No symptoms	
		2 HY+ = 1 symptom	☐

Syndrome No.	Syndrome Name	Symptoms	
35. (OD)	Other Symptoms of Depression	27. Morning depression 34. Loss of appetite· 37. Early waking 38. Loss of libido 39. Premenstrual exacerbation 0 No symptoms 2 SO+ = 1 symptom 3 SO++ = 2+ symptoms	☐
·36. (OR)	Organic Impairment	67. Delirious visual hallucinations 103. Organic impairment of memory 0 No symptoms 2 OR+ ≤ 1+ symptom	☐
37. (SC)	"Subcultural" Delusions or Hallucinations	64.(1) "Subcultural" hallucinations 83. "Subcultural" delusions 0 No symptoms 2 SC+ = 1+ symptom	☐
38. (DI)	Doubtful Interview	139. Misleading answers 140. Interview doubtfully adequate 0 No item 2 DI+ = 1 item	☐

Symptoms not used in SCL

1, 2, 3, 12, 13, 61. (3), 65, 93, 95-99, 104-107, 114, 127, 131, 138.

APPENDIX E

SYMPTOM CODING FORM AND
A SAMPLE OF THE ACTUAL DATA COLLECTED

E.1. Symptom coding form

COLLECTION OF PSYCHIATRIC SYMPTOMS FROM CASE-NOTES

Setting code number: _____ 1   Inc.Crit:_____

Patient name: _____ 2-7

Patient code number: _____

SYMPTOMS	R P 1	R A 2	I A 3	N R 4
A) APPEARANCE AND GENERAL BEHAVIOUR	■	■	■	■
108 self neglect 8				
109 bizarre appearance 9				
110 slowness and underactivity 10				
111 agitation 11				
112 gross excitement and violence 12				
113 irreverent behaviour 13				
115 embarrassing behaviour 15				
116 mannerisms and posturing 16				
117 stereotypes and tics 17				
118 behaves as if hallucinated by voices 18				
119 catatonic movements 19				
B) SPEECH	■	■	■	■
130 slowness of speech 20				
132 non-social speech 22				
133 muteness 23				
134 restricted quantity of speech 24				
135 neologisms and bizarre use of words and phrases 25				
136 incoherence of speech 26				
137 flight of ideas (hypomanic content of speech) 27				
C) AFFECT	■	■	■	■
120 observed anxiety 47				
121 observed depression 48				

231

	R P 1	R A 2	N R 4	
122 histrionic behaviour 49				
123 hypomanic affect 50				
124 hostile irritability 51				
125 suspicion 52				
126 perplexity 53				
128 blunted affect 55				
129 incongruity of affect 56				

SYMPTOMS	R P 1	R A 2	N R 4
D) MOOD	■	■	■
19 inefficient thinking 29			
20 poor concentration 30			
21 neglect due to brooding 31			
22 recent loss of interest 32			
23 depressed mood 33			
24 hopelessness 34			
25 suicidal plans or attacks 35			
27 morning depression 36			
28 social withdrawal 37			
34 loss of weight/appetite disturbance 38			
35/37 sleep disturbance 39			
36 subjective anergia and retardation 40			
38 loss of libido 41			
40 irritability 43			
41 expansive mood 44			
42 subjective ideomotor pressure 45			
43 grandiose ideas and actions 46			

232

F) NON-PSYCHOTIC THOUGHT CONTENT	R P 1	R A 2	I A 3	N R 4
4 worrying 69				
9 hypochondriasis 70				
17 specific phobias 71				
29 self depreciation 72				
30 lack of self confidence with people 73				
31 simple ideas of reference 74				
32 guilty ideas of reference 75				
33 pathological guilt 76				
44 obsessional checking and repeating 77				
45 obsessional cleanliness and rituals 78				
46 obsessional ideas and rumination 79				
G) DELUSIONS				
55 thought insertion 80				
56 thought broadcast or thought sharing 81				
57 thought echo or commentary 82				
58 thought block or withdrawal 83				
59 delusions of thought being read 84				
71 delusions of control 85				
72 delusions of reference 86				
73 delusional misinterpretation and misidentification 87				
74 delusions of persecution 88				
75 delusions of assistance 89				
76 delusions of grandiose ability 90				
77 delusions of grandiose identity 91				
78 religious delusions 92				
79 delusional explanation (paranormal and occult) 93				

233

80 delusional explanations (physical) 94				
81 delusion of alien forces penetrating or controlling mind or body 95				
82 primary delusions 96				
83 subculturally influenced delusions 97				
84 morbid jealousy 98				
85 delusions of pregnancy 99				
86 sexual delusions and hallucinations 100				
87 delusional, memories, confabulations, and fantastic delusions 101				
88 delusions of guilt 102				
89 simple delusions concerning appearance 103				
90 delusions of depersonalisation 104				
91 hypochondriacal delusions 105				
92 delusions of catastrophes 106				
H) PERCEPTIONS	■	■	■	■
47 derealisation 107				
48 depersonalisation 108				
49 delusional mood 109				
50 heightened perception 110				
51 dulled perception 111				
52 changed perception 112				
53 changed perception of time 113				
54 lost affect 114				
I) HALLUCINATIONS				
60 non-verbal hallucinations 115				
61 affective or non-specific verbal hallucinations 116				
62 non-affective verbal hallucinations (about the subject) 117				

63 non-affective verbal hallucinations (spoken to the subject) 118				
64 dissociative hallucinations 119				
66 visual hallucinations (clear consciousness) 120				
67 visual hallucinations (clouded consciousness) 121				
68 olfactory hallucinations and delusions 122				
69 delusion that subject smells 123				
70 other hallucinations and delusional elaborations 124				

E.2. Sample of the actual data collected at admission:

10 patients for each setting and 1 symptom for each heading were chosen to demonstrate the actual data collected:

Setting 1: Jules Thorn Day Hospital, St Pancras Hospital
Setting 2: Laffan Ward, U.C.H.
Setting 3: Noel Harris Ward, St Luke's Hospital

Coding 1: symptom 'reported as present'
Coding 2: symptom 'reported as absent'
Coding 3: symptom 'inferred as absent'
Coding 4: symptom 'not reported'

PATCODE	SETTING	S108	S130	S120	S19	S4	S55	S47	S60
FR8917	1	2	3	1	4	4	2	3	2
FX9303	1	3	1	4	4	4	2	2	3
FV7683	1	4	2	4	4	4	2	4	4
FY1891	1	3	3	3	4	4	3	4	2
FX2137	1	2	2	4	4	4	3	2	2
FS7351	1	2	4	4	4	4	4	4	4
EH0001	1	1	2	4	4	4	2	4	3
FW7191	1	2	2	1	4	1	2	2	2
EE7939	1	1	2	4	4	3	4	4	4
BK5545	1	2	2	4	4	2	2	2	2
EB7258	2	1	1	4	4	4	4	4	4
438996	2	1	2	4	4	4	3	4	4
337029	2	1	1	3	4	4	4	4	4
FV5545	2	4	2	4	4	4	3	4	3
BJ2076	2	4	2	4	4	4	4	4	4
447948	2	1	4	3	4	4	2	4	4
440526	2	2	3	4	4	2	2	2	2
FX8556	2	2	2	3	4	4	2	2	2
G57685	2	2	2	1	4	4	3	2	2
423871	2	1	3	4	4	4	4	4	2
030001	3	4	4	4	4	2	3	2	2
442355	3	1	2	4	4	4	4	4	4
436445	3	2	4	1	4	4	3	4	2
353510	3	1	4	1	4	2	2	4	4
438990	3	1	2	4	4	2	2	2	2
190735	3	1	3	1	4	4	3	4	3
439005	3	1	2	4	1	2	2	4	3
442319	3	4	4	4	4	4	3	4	3
438998	3	2	2	4	4	3	2	2	2
433468	3	2	2	1	4	2	2	2	2

APPENDIX F

BEHAVIOUR AND SOCIAL ITEMS CODING FORM
AND A SAMPLE OF THE ACTUAL DATA COLLECTED

F.1. Behaviour and social items coding form

COLLECTION OF WARD BEHAVIOUR AND SOCIAL ITEMS FROM CASE-NOTES

Setting code number: _____ (1)

Patient code number: _____ (2-7)

Patient name: _____ (8-13)

WARD BEHAVIOUR ITEMS	REPORTED AS A PROBLEM (1)	NOT REPORTED (2)
slowness of movement 14		
underactivity 15		
overactivity 16		
conversation 17		
social withdrawal 18		
leisure interests 19		
irrelevant or incomprehensible talk 20		
posturing and mannerisms 21		
threatening or violent behaviour 22		
tendency to remain in or return to bed 23		
personal appearance 24		
behaviour at mealtime 25		
SOCIAL ITEMS	REPORTED AS A PROBLEM (1)	NOT REPORTED (2)
housing conditions 26		
occupation/social role 27		
economic situation 28		
leisure/social activities 29		
family and social relationships 30		

F.2. A sample of the actual data collected on behaviour and symptom information:

10 patients for each setting, 3 behaviour and 2 social items were chosen to demonstrate the actual data collected:

Setting 1: Jules Thorn Day Hospital, St Pancras Hospital
Setting 2: Laffan Ward, U.C.H.
Setting 3: Noel Harris Ward, St Luke's Hospital

Coding 1: item 'reported as a problem'
Coding 2: item 'not reported'

PATCODE	SETTING	SLOWNESS	UNDERAC	OVERAC	HOUSING	OCCUSOCR
FV7683	1	2	2	2	1	1
FY1891	1	2	2	2	2	1
FX2137	1	2	2	2	2	2
FS7351	1	2	2	2	2	1
EH0001	1	2	2	2	1	2
EE7939	1	2	2	2	1	2
FW8565	1	2	2	2	1	1
FX3193	1	2	2	2	1	1
FY3872	1	2	2	2	2	1
EB7258	2	2	2	2	1	1
438996	2	2	2	2	1	1
337029	2	2	2	1	2	2
FV5545	2	2	2	2	2	2
BJ2076	2	2	2	1	2	1
447948	2	2	2	1	2	2
440526	2	2	2	2	2	2
FX8556	2	2	2	2	2	2
G57685	2	2	2	2	2	2
423871	2	2	2	1	2	2
030001	3	2	2	2	2	1
363636	3	2	2	1	2	2
436445	3	2	2	2	2	2
353510	3	2	2	2	2	1
438990	3	2	2	2	2	2
190735	3	2	2	2	2	1
439005	3	2	2	1	2	2
442319	3	2	2	2	2	2
438998	3	2	2	2	2	2
020300	3	2	2	2	1	1
FR8917	1	2	2	2	2	1

APPENDIX G

EXAMPLES OF SYMPTOM CODING

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
self neglect	Malodorous - Unkempt - Dirty feet with plastered toes - scruffy - Poor personal hygiene - Dishevelled	Clean and tidily dressed - Well presented - Reasonably kempt
bizarre appearance	Dressed in paper overall - Odd clothes - Draw arches and crosses over his body	Neatly dressed
slowness and underactivity	Retarded - Motor retardation - Little spontaneity of movement	No retardation
agitation	Agitated - Motor overactivity - Pacing up and down - Unable to stand still	No agitation
gross excitement and violence	Threatening to put the doctor through the window - Verbally aggressive - Flying himself on the floor - Verbally abusive	Not aggressive
irreverent behaviour	desinhibited - Eating toast and reading magazine during interview - Linking her arms to staff	Socially appropriate
embarrassing behaviour	Sexually disinhibited	
mannerisms and posturing	Odd postures	No bizarre movements
behaves as if hallucinated by voices	Responding to hallucinations - Behaves as if hallucinating	Did not appear hallucinated

241

The symptoms in the Appearance and Behaviour section were inferred as absent (coding 3) when one of the following examples were reported: cooperative and cheerful; relaxed with good eye contact; good rapport; pleasant and cooperative.

SPEECH

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
slowness of speech	Speech is very slow - Monotonous with frequent pauses	Normal rate
non-social speech	Talking to himself - Muttering	
muteness	Almost mute	
restricted quantity of speech	Limited amount of speech - Few answers, nil spontaneous - Lack of spontaneity	Normal quantity
neologisms	Using strange words	No evidence of formal thought disorder
Incoherence	Unable to talk coherently - Wandering off the point - Loosing of association - Inappropriate speech	Coherent - Appropriate answers - No evidence of formal thought disorder
flight of ideas	flight of ideas	No flight of ideas - No punning or clanging

The symptoms in the speech section were inferred as absent (coding 3) when statement such as 'no abnormalities of speech', 'normal rate, pattern and quantity' were reported.

AFFECT

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
observed anxiety	Frightened and distressed - Tense - Anxious - Jittery	
observed depression	Sad and gloomy - Tearful - Subdued - Very depressed	Not subdued - Not depressed
histrionic behaviour	Theatrical manners - Frequent gesticulation	
hypomanic affect	Elated - Very aroused	Not elated
hostile irritability	Irritable - Very angry - Hostile	
suspicion	Suspicious	Not suspicious
perplexity	Perplexed - Bewildered - Bemused and puzzled	
blunted affect	Flattened - Expressionless	No blunting of affect
incongruity of affect	Smiled frequently and inappropriately - Occasional laughter seemed inappropriate	Able to smile appropriately - Affect is appropriate

The symptoms in the affect section were inferred as absent (coding 3) when statement such as 'objectively euthimic', 'good range of affect', 'euthimic', 'no evidence of affect disorder' were reported.

MOOD

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
poor concentration	Unable to concentrate on anything else - Poor concentration	
neglect due to brooding	Reduced ability to cope in everyday life - Unable to look after himself	
recent loss of interest	Anhedonia - Loss of interest	No anhedonia
depressed mood	'I am so low' - 'I feel miserable'	No feeling of depression
hopelessness	'I can't see any future' - Expressed ideas of hopelessness	
suicidal plans or attacks	She would act on suicidal thoughts - 'Life is no longer worth living'	No suicidal ideation - No suicidal intent - No active plans

morning depression	Wakes up feeling miserable - Particularly low in the morning - Worst in the morning	No diurnal mood variation
social withdrawal	Social withdrawal	
loss of weight / appetite disturbance	Decrease of appetite - loss of weight - Eating a lot	Normal appetite - Good appetite
sleep disturbance	Reduced sleep - Insomnia - Poor sleep - Early morning wakening	No sleep disturbance
subjective anergia and retardation	Feeling lethargic - Low energy	
loss of libido	Loss of libido	
irritability	Flashes of anger - Irritability - Arguing with people in the streets - Impulsive violent outbursts - Explodes easily	No irritability
expansive mood	'A bit high' - Mildly grandiose	No feeling of elation
subjective ideomotor pressure	Thoughts are speeded up - 'I can think faster'	
grandiose ideas and actions	On top of the world - 'I am a very sex man'	No grandiosity

The symptoms sleep disturbance, loss of weight/ appetite disturbance and subjective anergia or retardation were reported as absent when the statement 'no biological features of depression' was reported

\\

NON PSYCHOTIC THOUGH CONTENT

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
specific phobias	Phobia for heights	
self depreciation	Negatives thoughts about herself - Poor self image	
simple ideas of reference	Overvalued ideas of reference	No overvalued ideas
pathological guilt	Hitting himself as a form of punishment	
obsessional features	Checking taps - Stepping on pavement cracks - Obsessed with sexual thoughts - Impossible to resist own thoughts	No obsessions were elicited

When symptoms included in this group were 'reported as present' the others were 'inferred as absent'. It was assumed that the phenomenon was assessed as whole and only those symptom had been elicited. We validated this procedure (table 4.11)

246

DELUSIONS

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
thought possession (insertion, broadcast, sharing, block)	Thought insertion - Delusions of thought broadcast - His memories were wiped out by someone	Normal possession of thought - Denied thought insertion and broadcast - No thought interference -
delusions of thoughts being read	'People are reading my thoughts'	
delusions of persecution	Persecutory beliefs - Thought that her food was being poisoned - Prominent delusions of persecution -	
delusions of assistance	Says that his God was giving him special powers	
delusions of grandiose ability	'I am an apostle' - He thinks that he can read other people's thought	
religious delusions	communication with the Virgin Mary	
primary delusions	He felt a wind passing over his body and immediately he knew that this was a sign that he was destined for something very special	

The symptoms in the delusion section were rated reported as absent when statements such as 'no abnormal beliefs', 'no evidence of delusional beliefs', 'no psychotic features' were recorded. When symptoms included in this group were 'reported as present' the others were 'inferred as absent'. It was assumed that the phenomenon was assessed as whole and only those symptom had been elicited. We validated this procedure (table 4.11)

247

PERCEPTIONS

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
derealisation	Derealisation	
depersonalisation	Depersonalisation	
delusional mood	'My mind was breaking apart'	
dulled perception	'The world is a very dull place'	

The symptoms in the perception section were rated as reported as absent when statements such as 'no abnormal experience', 'no abnormal perceptions' were recorded. When symptoms included in this group were 'reported as present' the others were 'inferred as absent'. It was assumed that the phenomenon was assessed as whole and only those symptom had been elicited. We validated this procedure (table 4.11)

HALLUCINATIONS

SYMPTOM NAME	REPORTED AS PRESENT CODING 1	REPORTED AS ABSENT CODING 2
affective hallucinations	Hearing his father voice	
non-affective verbal hallucinations	Voices in the second and third person - Auditory hallucinations - 'They are telling me that they hate me'	No auditory hallucinations
visual hallucinations	Describe clear visual hallucinations	No visual hallucinations
olfactory hallucinations	'My body has a strange odour'	
somatic hallucinations	'Something has been implanted in my head and stomach'	

In the hallucination section symptoms were rated as reported as absent when statements such as 'no hallucinations were elicited', 'no abnormal phenomena' were recorded. When symptoms included in this group were 'reported as present' the others were 'inferred as absent'. It was assumed that the phenomenon was assessed as whole and only those symptom had been elicited. We validated this procedure (table 4.11).

APPENDIX H

EXAMPLES OF BEHAVIOUR AND SOCIAL
ITEMS CODING

BEHAVIOUR ITEMS

ITEM NAME	REPORTED AS A PROBLEM CODING 1
overactivity	Up and down in the ward - Overactive
conversation	Muteness - Uncommunicative
social withdrawal	Insisting on eating by his own - Staying in his room all day - Isolated in the ward
leisure interests	No motivation - Very difficult to engage in any activity
Incomprehensive talking	Very difficult to get a history
Posturing and mannerisms	Crouching behind the bed - Showed movements of upper and lower limbs in a rhythmalistic way
Threatening or violent behaviour	Quite hostile on approach - Aggressive and threatening - Shouting - Verbally abusive
Personal appearance	Poor self care - Physically neglected

SOCIAL ITEMS

ITEM NAME	REPORTED AS A PROBLEM CODING 1
housing conditions	Terrible state of the flat - Homelessness -About to be evicted - Very unhappy with flat
occupation / social role	Unemployed - Difficulty in coping with baby - Unable to get public transport - Never could hold on to a job - Stop going to work
economic situation	In considerable debt - Financial difficulties -
leisure / social activities	Lack of motivation - No social activities
family and social relationship	Limited social network - Poor relationship with husband - Broken down relationship -Feeling lonely and unsupported - Little tolerance with family - Difficulty in establishing relationships

APPENDIX I

DEMOGRAPHIC AND CLINICAL VARIABLES FORM
AND A SAMPLE OF THE ACTUAL DATA COLLECTED

I.1. Demographic and clinical variables coding sheet

DEMOGRAPHIC AND CLINICAL VARIABLES

A) <u>SETTING</u>   --------
                                                                            (1)

        1. JULES THORN
        2. LAFFAN WARD
        3. NOEL HARRIS

B) <u>PATIENT CODE NUMBER</u>   ---------
                                                                          (2-7)

C) <u>DATE OF ADMISSION</u>   ----/----/----

D) <u>DATE OF DISCHARGE</u>   ----/----/----

                  (length of stay -------- 8-10)

E) <u>CONSULTANT NAME</u>   ---------------
                                                                           (11)

           1-GREENBERG
           2-FROST
           3-LITTLEWOOD
           4-GURLING
           5-HILL
           6-TURNER

F) <u>SEX</u>   --------
                                                                           (12)

           1-MALE
           2-FEMALE

G) <u>AGE IN YEARS</u>   -----------
                                                                         (13-14)

H) <u>ADDRESS</u>   --------
                                                                           (15)

           1-IN THE DISTRICT
           2-OUTSIDE THE DISTRICT
           3-NO FIXED ABODE
           4-RECORDED AS NOT KNOWN
           0-NOT RECORDED

I) <u>MARITAL STATUS</u>   -------- (16-17)

        1-SINGLE (never married)
        2-COHABITING/MARRIED (first marriage)
        3-REMARRIED
        4-SEPARATED OR DIVORCED
        5-WIDOWED (and not cohabiting)
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

J) <u>FAMILY TYPE</u>   -------- (18-19)

        1-CONJUGAL
        2-PARENTAL
        3-EXTENDED
        4-ALONE
        77-OTHER (specify................)
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

K) <u>TYPE OF HOUSING</u>   -------- (20-21)

        1-LIVING ROUGH OR WHENEVER A BED CAN BE FOUND EACH NIGHT
        2-HOTEL OR HOSTEL
        3-LODGING WITH A FRIEND
        4-PRIVATE HOUSE OR FLAT
        5-COUNCIL ACCOMMODATION
        6-RENTED HOUSE OR FLAT
        77-OTHER (specify.........................)
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

L) <u>COUNTRY OF BIRTH</u>   -------- (22-23)

        1-ENGLAND
        2-SCOTLAND
        3-WALES
        4-NORTHERN IRELAND
        5-IRISH REPUBLIC
        77-ELSEWHERE (specify......................)
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

M) <u>COUNTRY OF UPBRINGING</u>   -------- (24-25)

        1-ENGLAND
        2-SCOTLAND
        3-WALES
        4-NORTHERN IRELAND
        5-IRISH REPUBLIC
        77-ELSEWHERE (specify......................)
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

N) <u>EDUCATIONAL BACKGROUND</u>   --------
                                                    (26-27)

        1-NO QUALIFICATIONS
        2-"O" LEVELS
        3-"A" LEVELS
        4-UNIVERSITY DEGREE
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

O) <u>OCCUPATIONAL BACKGROUND</u>   --------
                                                    (28-29)

        1-PROFESSIONAL
        2-INTERMEDIATE OCCUPATIONS
        3-SKILLED OCCUPATIONS
        4-PARTLY SKILLED OCCUPATIONS
        5-UNSKILLED
        77-NEVER WORKED
        88-NOT CLASSIFIED
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

P) <u>CURRENT EMPLOYMENT STATUS</u>   --------
                                                    (30-31)

        1-EMPLOYED
        2-TEMPORARY EMPLOYMENT
        3-UNEMPLOYED
        4-RETIRED
        77-OTHER (specify....................)
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

Q) <u>AGE AT ONSET OF FIRST PSYCHIATRIC ILLNESS</u>   --------
                                                      (32-33)

        77-NOT RECORDED
        99-RECORDED AS NOT KNOWN

R) <u>NUMBER OF PREVIOUS ADMISSIONS</u>   --------
                                                    (34-35)

        77-NOT RECORDED
        88-ADMITTED BEFORE
        99-RECORDED AS NOT KNOWN

S) <u>SUICIDE BEHAVIOUR</u>   --------
                                                    (36-37)

        1-NO ATTEMPTS
        2-ONE ATTEMPT
        3-TWO PLUS
        99-RECORDED AS NOT KNOWN
        00-NOT RECORDED

T) DURATION OF PRESENT ILLNESS   --------
                                                                      (38-39)

                1-LESS THAN 3 MONTHS
                2-MORE THAN 3 MONTHS
                99-RECORDED AS NOT KNOWN
                00-NOT RECORDED

U) HISTORY OF ALCOHOL CONSUMPTION   --------
                                                                      (40-41)

                1-DO NOT DRINK
                2-SOCIAL DRINKER
                3-HISTORY OF ABUSE
                99-RECORDED AS NOT KNOWN
                00-NOT RECORDED

V) HISTORY OF DRUG USE   --------
                                                                      (42-43)

                1-NEVER USED
                2-CANNABIS
                3-AMYTAL, ETC
                4-LSD, AMPHETAMINE, ETC
                5-COCAINE, HEROIN, ETC
                6-ADDICTION REPORTED WITHOUT MENTIONING DRUG
                99-RECORDED AS NOT KNOWN
                00-NOT RECORDED

X) CURRENT TREATMENT   -------
                                                                      (44-45)

                1-DRUGS ONLY
                2-DRUGS+GROUP ACTIVITIES
                3-DRUGS+PSYCHOTHERAPY
                4-PSYCHOTHERAPY+GROUPS ACTIVITIES
                5-DRUGS+PSYCHOTHERAPY+GROUPS ACTIVITIES
                77-OTHER (specify ....)

W) REFERRAL AGENCY   --------
                                                                      (46-47)

                1-G.P.
                2-SOCIAL WORKER
                3-ACCIDENT AND EMERGENCY DEPARTMENT
                4-GENERAL HOSPITAL
                5-C.P.N.S.
                6-POLICE
                7-SELF REFERRAL
                8-PSYCHIATRIC IN-PATIENT UNIT
                9-OUT-PATIENT SERVICE
                77-OTHER (specify.............................)
                99-RECORDED AS NOT KNOWN
                00-NOT RECORDED

Y)<u>DIAGNOSIS</u>   -------
                                                      (48-49)

                    1-SCHIZOPHRENIA
                    2-AFFECTIVE DISORDER
                    3-PARANOID DISORDER
                    4-ANXIETY NEUROSIS
                    5-PHOBIC NEUROSIS
                    6-OBSESSIVE COMPULSIVE NEUROSIS
                    7-HYSTERIA
                    8-NEUROTIC DEPRESSION
                    9-PERSONALITY DISORDER
                    10-PSYCHOACTIVE SUBSTANCE USE DISORDER (specify........)
                    77-OTHER (specify.........................)
                    00-NOT RECORDED

Z)<u>FATE AFTER DISCHARGE</u>   -------
                                                      (50-51)

                    1-G.P. (community)
                    2-OUT PATIENT FOLLOW-UP
                    3-DAY PATIENT CARE
                    4-IN-PATIENT TREATMENT
                    5-C.P.N.S
                    6-SELF DISCHARGED
                    7-NO FOLLOW UP ARRANGEMENTS
                    77-OTHER (specify.............................)
                    00-NOT RECORDED

I.2. Sample of the actual data collected on demographic and clinical information:

·10 patients for each setting and 6 variables were chosen to demonstrate the actual data collected:

Setting 1: Jules Thorn Day Hospital, St Pancras Hospital
Setting 2: Laffan Ward, U.C.H.
Setting 3: Noel Harris Ward, St Luke's Hospital

For codings see Demographic and Clinical variables form.

PATCODE	SETTING	CONSNAME	SEX	AGE	FAMTYPE	DIAGNOS
FR8917	1	1	2	27	4	8
FX9303	1	1	2	19	4	9
FV7683	1	1	2	21	4	9
FY1891	1	1	1	39	1	1
FX2137	1	1	2	25	4	3
FS7351	1	1	1	32	4	77
EH0001	1	1	1	32	4	1
FW7191	1	1	1	24	2	6
EE7939	1	1	2	39	1	9
BK5545	1	1	2	39	1	77
EB7258	2	3	1	40	4	8
438996	2	2	2	25	77	1
337029	2	2	2	40	1	2
FV5545	2	2	2	27	1	2
BJ2076	2	2	1	48	1	77
447948	2	3	1	36	4	1
440526	2	3	2	37	3	77
425441	2	2	1	45	4	1
445877	2	2	2	37	4	1
BZ5212	2	2	2	49	4	1
030001	3	5	2	22	3	1
442355	3	6	2	22	1	2
436445	3	5	2	42	1	2
353510	3	5	1	29	3	1
438990	3	5	2	77	1	77
190735	3	6	1	39	1	1
439005	3	5	2	20	77	77
442319	3	4	1	33	3	77
438998	3	6	1	38	1	8
433468	3	5	2	29	4	4

259

APPENDIX J

DISCHARGE INFORMATION

J.1. Discharge information coding sheet

SETTING NUMBER: _____ (1)

INCLUSION CRITERIA: --------------(2)

PATIENT NAME: ------------------------------ (3-8)

INFORMATION COLLECTED ON ADMISSION (SYMPTOMS 'REPORTED AS PRESENT')	SYNDROME DERIVATION (name, number and column)	SYNDROME RATING	DISCHARGE STATEMENT	LINK
S108	NG	2	missing	88
S111	AG	2	less agitated	1
S120	GA	2	less distressed	1
S22	IC	2	missing	88
S23	SD	1	missing	88
S34	OD	3	stopped eating	-1
S3537	OD	-	wandering at night	-1
S63	AH	2	still complaining of verbal	0
S66	VH	2	and visual hallucinat.	0
9 symptoms		SCL score=16		

GENERAL STATEMENT OF CHANGE:

Rating:(9)

LINKSY:(10-12) 66%   LINKSI:(13-15)   LINKBH:(16-18)

COMMENTS:

J.2. Equation for the calculation of 'non-response' symptoms

$$NRV = X_p + (X_i - X_{\Sigma i})$$

where NRV is the value to be estimated, X_p is the patient summed mean score without considering codings 88, X_i is the item mean score for all patients and, $X_{\Sigma i}$ is the overall mean of the symptoms for all patients. Let us show how we used this equation based on the real research subject presented in the previous page.

The patient had 9 symptoms reported as present on admission with the respective ratings on discharge - S111(1), S120(1), S34(-1), S3537(-1), S63(0) and S66(0). There were no statement of change at discharge related to symptoms S108, S22 and S23. Before deriving the FSC we would like to estimate the value of symptom S108 (NR108), S22(NR22) and S23(NR23).

We have $X_p = 0$, $X_{\Sigma i} = 0.71$ and X_i for S108 = 0.53, for S22 = 0.60 and for S23 = 0.70. Replacing these values in the equation we have the following results:

$$NR108 = 0 + 0.53 - 0.71 = -0.18$$
$$NR22 = 0 + 0.60 - 0.71 = -0.11$$
$$NR23 = 0 + 0.70 - 0.71 = -0.01$$

The coding 88 is replaced by these values and a FSC (summed mean score) of - 0.03 is derived for this patient.

J.3. Sample of the actual data collected (Number of symptoms 'reported as present' and the Syndrome Check List score) and at discharge (General Statement of Change - GSC, and % of symptom linkage)

10 patients for each setting were chosen to demonstrate the actual data collected:

Setting 1: Jules Thorn Day Hospital, St Pancras Hospital
Setting 2: Laffan Ward, U.C.H.
Setting 3: Noel Harris Ward, St Luke's Hospital

GSC codings and ratings:

-1 statement of 'deterioration'
 0 statement of 'no change'
 1 statement of 'improvement'
 2 statement 'unratable'
 3 GSC not reported

PATCODE	SETTING	GENSTAT	LINKSY	NUMBSYMP	SCLSCORE
FR8917	1	1	0	10	16
FX9303	1	2	17	10	13
FV7683	1	3	100	6	8
FY1891	1	2	0	3	6
FX2137	1	3	50	4	7
FS7351	1	1	0	2	4
FW7191	1	1	66	3	6
BK5545	1	-1	0	4	7
FW8565	1	3	100	5	10
BY6102	1	3	57	7	12
EB7258	2	0	80	16	10
438996	2	1	69	13	23
337029	2	1	90	10	11
FV5545	2	1	10	10	18
BJ2076	2	1	100	6	11
447948	2	1	60	10	17
440526	2	0	0	3	3
G57685	2	2	0	7	13
423871	2	1	69	13	23
425441	2	3	80	5	10
030001	3	2	33	3	7
442355	3	1	70	10	19
436445	3	1	63	11	17
353510	3	2	38	13	23
438990	3	3	33	6	10
190735	3	1	30	10	17
439005	3	3	40	15	25
442319	3	1	77	9	16
438998	3	1	70	10	14
433468	3	1	100	6	11

APPENDIX K

EXAMPLES OF GENERAL STATEMENT
OF CHANGE (GSC)

K.1. For *'improvement'* in mental state:

- Her mental state improved.

- Much improved.

- He made good progress at the hospital and now the acute episode has passed.

- It was felt she made some improvement.

- By the end of his admission he was less depressed.

- He had made progress.

- He did improved his mood.

- On discharge her depressive symptoms had cleared and

her personality problems were not apparent.

- His mood was stable, neither elated, neither depressed.

- His mental state did improve during his admission.

- His mental state and behaviour improved considerably.

- He was not depressed and successfully abstained from

alcohol consumption.

- At the time of discharge she was totally symptom free.

- Overall his symptoms have greatly reduced.

- He responded very well to treatment.

- He made gradual improvement.

- Her parents and the ward doctors have all noticed a

definite improvement in her mental state.

- She has done incredibly well.

- Showed dramatic improvement.

- Quick recovery.

- Improved +++.

- Fully recovered.

- She felt considerably better.

- She was no longer suffering from an acute episode of depression.

K.2. For *'no change'* in mental state.

- It was felt by both, the patient and the team, that his attendance made little difference to his symptoms and behaviour.

- As she can not cooperate with any offer there is little else we can do for her.

- No changes, the pattern of previous admissions will continue.

- Although her mood lifted slightly at time of discharge, she continued to feel low and unable to concentrate on work.

- He was not benefiting for being here, so he was discharged.

- At discharge we concluded that this patient had not made satisfactory progress.

- It was felt that she did not benefit from treatment here.

- Any efforts to investigate his problems and help him had failed.

- She is still hypomanic in her mental state.

- No changes at all.

- No changes in mental state.

- Mental state much of the same on admission.

- Appear to have not changed.

- Continued to express psychotic phenomena at discharge.

K.3. For *'deterioration'* in mental state

- Her mental state deteriorated further.

- She seemed to regress as her social and personal problems persist.

- She got worst and was admitted to a in-patient unit.

- Felt to be increasingly elated and fragile in her mental state.

K.4. GSC that were not possible to be rated

- Mental State has fluctuated.

- Very difficult to assess his mental state.

- Her symptoms fluctuated throughout.

- Mental state stable enough to go.

- Well enough to be transferred.

- She was discharged because she was not a severe management problem.

APPENDIX L

EXAMPLES OF STATEMENT OF CHANGE
FOR SYMPTOMS REPORTED PRESENT AT ADMISSION

APPEARANCE AND GENERAL BEHAVIOUR

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
self-neglect	Personal hygiene still precarious ('no change') - Now is able to care for herself ('improvement') - Self-care has deteriorated ('deterioration') - Personal hygiene improved ('improvement')
bizarre appearance	Towards the end of his admission he was no longer bizarre ('improvement') - Dressed appropriately ('improvement')
slowness and underactivity	No evidence of motor abnormality ('improvement')
agitation	Less agitated ('improvement') - Very restless at discharge ('no change') - His agitation responded well to medication ('improvement') - Very calm at discharge ('improvement') - still very agitated ('no change')
gross excitement and violence	Very pleasant at discharge ('improvement') - Very relaxed at discharge ('improvement')
embarrassing behaviour	Still very disinhibited ('no change') - Less disinhibited ('improvement')
behaves as if hallucinated by voices	Still behaving as responding to voices ('no change') - No hallucinations ('improvement')

SPEECH

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
slowness of speech	Pace of speech has markedly improved ('improvement')
muteness	Articulated and confident ('improvement')
restricted quantity of speech	Still not verbally responsive ('no change' - More articulated ('improvement')
incoherence of speech	It is possible now to follow his train of thought ('improvement') - Coherent ('improvement') - No formal thought disorder ('improvement') - Able to talk more coherently ('improvement')
flight of ideas	No evidence of flight of ideas ('improvement') -

AFFECT

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
observed anxiety	Less distressed ('improvement') - Still anxious ('no change') - No anxiety ('improvement') - Less anxious ('improvement') - Anxiety has reduced ('improvement') - Very anxious ('no change')
observed depression	She does not appear to be depressed ('improvement') - Not clinically depressed ('improvement') - Episodes of tearfulness ('no change') - Free of crying boots ('improvement') - Cheerful ('improvement') - Brighter ('improvement') - less episodes of tearfulness ('improvement')
hypomanic affect	Settled and euthimic ('improvement') - Still hypomanic ('no change') - Very elated ('deterioration')
hostile irritability	She is not hostile ('improvement') - She is still irritable ('no change') - Very friendly ('improvement') - Seem less angry ('improvement') - She is very angry
suspicious	Pleasant and sociable ('improvement') - Not suspicious ('improvement')
perplexity	Still perplexed ('no change')
blunted affect	Flat ('no change')
incongruity of affect	Appropriate affect ('improvement')

271

MOOD

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
poor concentration	Unable to concentrate ('improvement') - Concentration still poor ('no change') - Started concentrating ('improvement')
recent loss of interest	Able to initiate activities ('improvement')
depressed mood	She feels well ('improvement') - Improvement in depressed mood ('improvement') - His mood gradually lifted ('improvement') - Continue to feel low ('no change') - No change in subjective dysphoria ('no change') - His depression has gone ('improvement') - No evidence of depression ('improvement')
hopelessness	Able to contemplate future plans ('improvement') - Able to look at life in a more positive way ('improvement') - able to set goals ('improvement') - Can't make any positive plans ('improvement')

suicidal plans or attacks	Became actively suicidal ('deterioration') - No suicidal thoughts ('improvement') - Expressing suicidal ideation ('no change') - no suicidal ideation ('improvement') - Several suicidal attempts ('deterioration') - Still expressing thought of self harm ('no change') - Brighter and stable ('improvement') - No longer suicidal ('improvement')
loss of weight / appetite disturbance	Stopped eating ('deterioration') - His appetite has improved ('improvement') - Eating well ('improvement') - Began eating ('improvement') - Poor appetite ('no change')
sleep disturbance	Disturbed sleep ('no change') - Sleep improved ('improvement') - Sleeping pattern is normal ('improvement') - Sleeping well ('improvement')
subjective anergia	Going for walks ('improvement') - More active ('improvement')
loss of libido	Sometimes there ('improvement')
irritability	Appeared calm ('improvement') - Still irritable ('no change') - No irritability ('improvement') - Still complaining of violent impulses ('no change')
expansive mood	Decreased elation ('improvement') -
grandiose ideas	Still showing grandiose ideas ('no change')

\\

NON-PSYCHOTIC THOUGHT CONTENT

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
worrying	Still worrying ('no change') - Appeared not to be worried as before ('improvement')
self depreciation	Feelings of self depreciation and self hatred has exacerbated ('deterioration')
simple ideas of reference	Started wearing sunglasses to 'hide my self consciousness' ('deterioration')
obsessional ideas and ruminations	Ruminations are worse ('deterioration') - Continued to ruminate ('no change')

DELUSIONS

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
thought insertion	denies thought insertion ('improvement')
thought broadcast or thought sharing	Remained thought disordered ('no change')
delusions of thought being read	denies any delusion ('improvement')
delusions of reference	Free of his delusions of reference ('improvement')
delusional misinterpretation and misidentification	Still having delusions ('no change')
delusions of persecution	Her paranoid ideas has receded ('improvement') - Persistent paranoid ideas ('no change') - Remain with his persecutory ideas ('no change') - No current delusions ('improvement') - No longer experiencing persecutory beliefs ('improvement') - No evidence of paranoid features ('improvement')
religious of grandiose ability	Still thinks he is an apostle ('no change') - Continued talking about his healing power ('no change')
religious delusions	Still with religious delusions ('no change')
delusional explanations	Persistent ('no change')
morbid jealousy	Still convinced of partner infidelity ('no change')

\|

HALLUCINATIONS

NAME OF SYMPTOM REPORTED AS PRESENT AT ADMISSION	EXAMPLE OF STATEMENT OF CHANGE REPORTED AT DISCHARGE
non-affective verbal hallucinations	Continued disturbed by audit hallucinations ('no change') - No longer complaining of hearing voices ('improvement') - Still complaining of audit hallucinations ('no change') - Denies hallucinations ('improvement') - No hallucinations ('improvement')
visual hallucinations	Still complaining of visual hallucinations ('no change')
other hallucinations	Continued disturbed by tactile hallucinations ('no change') -

APPENDIX M

CATEGO INFORMATION

MRC Social Psychiatry Unit

CATEGO CLASSES		CATEGO SUBCLASSES	
'Schizophrenic psychoses'			
	S+	1.	NS+
		2.	NS?
		3.	DS+
		22.	NSMN/DSMN
	S?	4.	DS?
'Paranoid psychoses'			
	P+	5.	DP+
		23.	DPMN
		30.	DPPD
	P?	6.	DP?
		31.	DP?/AP?
'Borderline and doubtful psychoses'			
	O+	7.	CS+
		26.	CSMN
		28.	RSMN
		33.	RSPD
		34.	CSPD
	O?	8.	RS+
		9.	RS?
		10.	SS
		17.	UP+
		18.	UP?
		35.	XP
'Mania and mixed affective psychoses'			
	M+	13.	MN+
		21.	MAP+
		15.	HM+
		–	MAN+
		25.	MNCS
		27.	MNRS
	M?	14.	MN?

CATEGO CLASSES   CATEGO SUBCLASSES

'Depressive psychoses'

D+   11. PD+
                                            19. AP+

D?   12. PD?
                                            20. AP?
                                            32. PDRS
                                             -   DP,?,RD+
                                                 (see Appendix 9)

'Retarded depressions'

R+   42. RD+

'Neurotic depressions'

N+   38. ND+
                                            40. SD+

N?   39. ND?
                                            41. SD?
                                            43. RD?

'Anxiety States'

A+   44. AN+
                                            46. PN+

A?   45. AN+
                                            47. PN?

'Obsessional neuroses'

B+   36. ON+

B?   37. ON?

'Hysterical' conditions

H   48. HT

Symptoms NOS

X   49. XN

No abnormality in PSE

NO   50. NO

279

M.2. Table of Catego report for each patient and rater

The symptoms reported as present by both raters were used to fill in the Syndrome Check List (S.C.L.) to derive diagnoses and catego classes for each patient.

Pat.	Catego subclass		Catego class		Diagnosis	
	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2
1	ND+	ND+	N+	N+	300.4	300.4
2	SD+	SD+	N+	N+	300.4	300.4
3	DP?AP?	DP?AP?	P?	P?	297.9	297.9
4	SD+	SD+	N+	N+	300.4	300.4
5	RD+	RD+	R+	R+	296.2	296.2
6	NS+	NS+	S+	S+	295.3	295.3
7	SD+	SD+	N+	N+	300.4	300.4
8	DP?	PD+	P?	D+	297.9	296.2
9	RS+ND+	UP+ND+	O? N+	O? N+	300.4	300.4
10	UP?ND+	RD+	O? N+	R+	300.4	300.4
11	MN+	MN+	M+	M+	296.1	296.1
12	SD+	SD+	N+	N+	300.4	300.4
13	DS?	NS+	S?	S+	295.3	295.3
14	MN+	MN?	M+	M?	296.1	296.1
15	RS+	RS+	O?	O?	299	?

This table shows an overall agreement of 73% for catego subclasses and classes, and 86% for tentative diagnosis. A mean kappa of 0.85 is derived when examining the number of syndromes found as present or absent for each patient in a 2 x 2 matrix.

APPENDIX N

TABLE AND MATRICES OF KAPPA
FOR BEHAVIOUR AND SOCIAL ITEMS

N.1. Table of stratified kappa

PATIENT AND SETTING	KAPPA BEHAVIOUR	KAPPA SOCIAL
Setting 1		
	DO NOT APPLY (DNA)	
PAT. 1	DNA	1
PAT. 2	DNA	1
PAT. 3	DNA	-0.36
PAT. 4	DNA	0.62
PAT. 5		1
		mean kappa 0.65
Setting 2		
	0	
PAT. 6	0	0.55
PAT. 7	0	0.55
PAT. 8	0.55	1
PAT. 9	0.62	1
PAT. 10	mean kappa 0.24	1
		mean kappa 0.82
Setting 3		
	1	
PAT. 11	1	1
PAT. 12	1	0.29
PAT. 13	0.63	0.29
PAT. 14	0.75	0.55
PAT. 15	mean kappa 0.88	0.62
		mean kappa 0.55

mean stratified kappa for behaviour data = **0.56**
mean stratified kappa for social data = **0.67**

N.2. Matrices and kappa of collapsed ratings

Due to the huge variability of stratified kappa for each patient, two matrixes for behaviour and social data were examined to check if collapsed kappa would have a different behaviour.

Collapsed ratings for behaviour data

	'Not reported as a problem'	'Reported as a problem'
'Not reported as a problem'	105	4
'Reported as a problem'	2	9

105 agreement on behaviour items 'not reported' as a   problem
9   agreement on behaviour items 'reported as a problem'
2/4 disagreement on behaviour items 'reported as a   problem'

collapsed kappa = 0.72

Collapsed ratings for social data

	'Not reported as a problem'	'Reported as a problem'
'Not reported as a problem'	42	6
'Reported as a problem'	6	21

42 agreement on social items 'not reported as a problem'
21 agreement on social items 'reported as a problem'
6/6 disagreement on social items 'reported as a problem'

collapsed kappa = 0.65

APPENDIX O

CROSS-TABULATION OF PSE AND
CASE NOTE RATINGS

PSE and case note ratings were cross-tabulated to explore rating 'inferred as absent' and 'not reported' made from case notes (see sections 3.3.3.1 and 4.5.2.1). We considered 100 symptoms for N=34 giving us a total of 3400 ratings.

PSE codings and ratings:

0 - symptom is absent
1/2/3 - symptom is present

Case note codings and ratings:

1 - symptom is 'reported as present'
2 - symptom is 'reported as absent'
3 - symptom is 'inferred as absent'
4 - symptom is 'not reported'

	PSE codings		
CASE NOTE codings	0	1	2/3
1	133 a	152 e	62 i
2	903 b	22 f	7 j
3	1179 c	64 g	19 k
4	689 d	139 h	31 l

From this table we deduced that 93% (table 4.12) of the symptoms 'inferred as absent' were really absent as shown by the PSE ratings. To reach this result we made the following calculation:

$$\frac{c \quad \text{(total agreement)}}{c+g+k \quad \text{(all 'inferred as absent' ratings)}} = \frac{1179}{1262} = 0.93$$

By the same token, we gathered from the table that 80% (table 4.12) of the symptoms 'not reported' were really absent as shown by the PSE ratings. The following calculation was made:

$$\frac{d \quad \text{(total agreement)}}{d+h+l \quad \text{(all 'not reported' ratings)}} = \frac{689}{859} = 0.80$$

A calculation like that was made for all symptom headings presented in table 4.12.

APPENDIX P

TWO WAY TABLE PRESENTING THE CALCULATION OF SS, SP, PPV FOR ALL SYMPTOMS COMBINED

A two-way table derived from the cross-tabulation of the PSE and case note ratings presented in the previous appendix was constructed and measures of sensitivity (SS), specificity (SP) and positive predictive value (PPV) were calculated (N=34):

PSE ratings

CASE NOTE ratings	symptom present		symptom absent	
symptom present	214	m	133	o
symptom absent	282	n	2771	p

m = e + i = 214
n = f + g + h + j + k + l = 282
o = a = 133
p = b + c + d = 2771

$$SS = \frac{m}{m+n} = 0.43 \qquad SP = \frac{p}{o+p} = 0.95 \qquad PPV = \frac{m}{m+o} = 0.62$$

The same calculation was made for all symptom headings and results are shown in table 4.12.

We also present a coefficient based on Manchanda and Hirsch (1986) where all ratings are combined leading to the following coefficient of

$$Validity = \frac{m + p}{m + n + o + p} = 0.88$$

APPENDIX Q

TWO WAY TABLE PRESENTING THE CALCULATION OF SS, SP, PPV FOR ALL SYNDROMES COMBINED

Using the same method presented in the previous appendix we derived a two-way table for all syndromes combined and measures of sensitivity (SS), specificity (SP) and positive predictive value (PPV) were calculated (N=34):

PSE ratings

CASE NOTE ratings	syndrome present		syndrome absent	
syndrome present	173	m	79	o
syndrome absent	138	n	767	p

$$SS = \frac{m}{m+n} = 0.56 \quad SP = \frac{p}{o+p} = 0.91 \quad PPV = \frac{m}{m+o} = 0.69$$

The same calculation was made for 34 syndromes of the SCL and results are shown in table 4.13.

We did not consider syndromes 8,9,28 and 38 (see Appendix D)

We also present a coefficient based on Manchanda and Hirsch (1986) where all ratings are combined leading to the following coefficient of

$$Validity = \frac{m + p}{m + n + o + p} = 0.81$$

APPENDIX R

MATRIX AND CALCULATION
OF THE VALIDATION PROCEDURE
FOR THE CLINICAL DIAGNOSIS

A weighted kappa coefficient was calculated as a mean of validating clinical diagnoses reported in the case notes. They were compared with CATEGO-PSE diagnoses derived from independent PSE interviews.

We considered only 23 diagnosis because PSE does not discriminate some diagnosis reported in the case notes (e.g. personality disorder). A matrix was constructed as follow:

Diagnoses considered:

Schizophrenia   - 1
Affective disorder - 2
Neurotic depression - 3
Paranoid disorder   - 4

CATEGO
diagnosis

CASE NOTE
diagnosis

	1	2	3	4
1	9a 1b 5.2c	0 .5 .4	1 .25 3.9	0 .75 .43
2	1 .5 1.6	0 1 .13	1 .75 1.2	1 .5 .13
3	0 0 4.2	1 .75 .35	7 1 3.1	0 0 .35
4	2 .75 1.0	0 .5 .09	0 .25 0.8	0 1 .09

a observed cell proportion
b agreement weight
c chance-expected cell proportion

Weighted kappa is given by

$$Kw = \frac{\Sigma S - \Sigma Sc}{Sm - ESc}$$

Where S is a x b, Sc is b x c, and Sm is Σa x maximum weight. In this matrix Kappa is

given by

$$Kw = \frac{20.3 - 13.0}{23.0 - 13.0} = 0.73$$

APPENDIX S

PAIRED OBSERVATIONS FOR GSC, FSC AND CPRS

S.1. Categorical rating scales

Paired observations of GSC, FSC and CPRS scores were derived and compared by means of outcome category. Two scales are proposed based on the scores born by each outcome measure: a 3-point rating scale and on a 5-point rating scale. Table S.2. shows paired observations of the GSC and CPRS scores compared on the basis a 3-point rating scale and table S.3. presents the paired observations for the FSC and CPRS scores compared on the basis of a 3-point and a 5-point rating scales.

3-point rating scale

coding	scores	outcome category
1	-1.00 -- -0.01	deterioration
2	0	no change
3	0.01 -- 1.00	improvement

5-point rating scale

coding	scores	outcome category
1	-1.00 -- -0.06	deterioration
2	-0.05 -- 0.05	borderline/no change
3	0.06 -- 0.36	minimal improvement
4	0.37 -- 0.66	moderate improvement
5	0.67 -- 1.00	marked improvement

S.2. Table of paired observations of the GSC and CPRS scores compared on the basis a 3-point rating scale (N=18 †)

PATCODE	GSC score	3-point rating scale codings	CPRS score	3-point rating scale codings
01	1	3	0.54	3
02	1	3	0.55	3
03	1	3	0.80	3
04	1	3	0.88	3
05	1	3	0.33	3
06	1	3	0.91	3
07	1	3	0.74	3
08	1	3	0.68	3
09	1	3	0.33	3
10	0	2	-0.33	1
11	1	3	0.41	3
12	1	3	0.05	3
13	1	3	0.17	3
14	1	3	0.58	3
15	1	3	0.88	3
16	1	3	0.56	3
17	1	3	0.35	3
18	1	3	0.95	3

† Only 62% of the Manchester case notes had a general statement of change reported at discharge

From this table we found 94% of the pairs being matched on the basis of a 3-point rating scale.

S.3. Table of paired observations of the FSC and CPRS scores compared on the basis a 3-point and 5-point rating scales (N=29)

	FSC score	3-point rating scale codings	5-point rating scale codings	CPRS score	3-point rating scale codings	5-point rating scale codings
01	-0.22	1	1	-0.37	1	1
02	0	2	2	-0.33	1	1
03	0.14	3	3	0.12	3	3
04	0.40	3	4	0.62	3	4
05	0.52	3	4	0.53	3	4
06	0.52	3	4	0.54	3	4
07	0.53	3	4	0.55	3	4
08	0.60	3	4	0.45	3	4
09	0.61	3	4	0.95	3	5
10	0.64	3	4	0.84	3	5
11	0.68	3	5	0.33	3	3
12	0.78	3	5	0.88	3	5
13	0.84	3	5	0.35	3	3
14	0.85	3	5	0.50	3	4
15	0.85	3	5	0.56	3	4
16	0.86	3	5	0.41	3	4
17	0.91	3	5	0.17	3	3
18	0.91	3	5	0.80	3	5
19	0.92	3	5	0.54	3	4
20	0.92	3	5	0.20	3	3
21	0.92	3	5	0.20	3	3
22	0.93	3	5	0.58	3	4
23	0.94	3	5	0.05	3	2
24	0.97	3	5	0.11	3	3
25	0.97	3	5	0.74	3	5
26	0.99	3	5	0.91	3	5
27	1.00	3	5	0.88	3	5
28	1.00	3	5	0.33	3	3
29	1.00	3	5	0.68	3	5

From this table we deduced that 96% of pairs matched on the basis of the 3-point scale and 45% on the basis of the 5-point scale.