A study of patient outcomes in an acute hospital

Thesis submitted for PhD by

Martin John BARDSLEY

London School of Economics and Political Science
University of London

July 1991

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

A study of patient outcomes in an acute hospital

Outcome measures have been developed in an acute hospital for specific patient groups (primarily cholecystectomy, diabetes, coronary angioplasty and knee replacements). For each condition a set of indicators was derived which ranged from clinical and laboratory measures to measures of general health status. All indicators attempt to show changes in patient health over time.

It has been shown possible to collect the necessary data for such outcomes measures. The costs and methods of data collection varied between conditions. Patient completed questionnaires were found to be particularly useful and in inpatient studies have given high response rates (over 95%) for postal follow-ups and have been validated by interviews.

The differing ability of the various indicators to show clinical changes has been demonstrated. In all specialties there was found to be generally high levels of association between different indicators. The information collected on patient outcomes was presented at meetings of the various clinical teams and the value of the information in promoting practical change was examined. It was concluded that different indicators have different value in such reviews and that three key characteristics are identified. The first concerns whether the measures reflect clinical or patient's perceptions of health. The second concerns the extent to which an indicator is a direct measure of health or a proxy (or process) measure. The third concerns the extent to which an observed outcome indicator can be linked to particular processes of care.

The study has generated support from the clinicians and it is suggested has changed their views on how they judge their own performance. In some instances practical changes have resulted form the presence of the outcome information. The potential future roles for outcome measurement in the health service is discussed.

CONTENTS

List of Tables

List of Figures

Chapter 1 Introduction

This thesis describes work undertaken during the CASPE/Freeman outcome project established in autumn 1988. Its aims were to...

> *"..develop and pilot.. methods of assessing the outputs and or outcomes in a way meaningful to both clinicians and managers"* (CASPE proposal to DoH 1988).

As such the project was aimed in the first instance to explore the relevant dimensions of outcome and possible methods of measurement in specific areas of hospital activity. Following on from the successful agreement and measurement of outcome, are the questions over the value of such outcomes in practice. Progress made in addressing these questions will be discussed.

The project, which has been managed by the author, attempted to develop ways in which health service information and management systems could incorporate assessments of the impact of health services on health and everyday life. There has always been a concern with the outcomes of care yet measurement of these within the service has in the past been limited. Indeed in many ways the health service is only now grasping the ways to monitor its processes and costs -something that is held to be considerably easier (and less costly) than outcome measurement.

To date the measurement of outcome has tended to be limited to research settings - as discussed later. Yet there is a growing view in the service that it is now time to tackle some of the difficult problems of outcome measurement (Mills 1987). In particular clinicians and health service managers should be able to assess the quality of their management by explicitly monitoring the patient benefit that results. The organisation and practice of care should seek to maximise those benefits when they do exist and stop doing things when no benefit exits.

The interest in outcome measurement since the project began has grown considerably.

There are many reasons for this including:-

- the technology required to process information of this sort has become more accessible (both in terms of computer hardware and software)
- the changes envisaged by the latest White Paper "Working for Patients" (Department of Health 1989) emphasised the importance of monitoring outcomes as well as including explicit quality clauses when contracting for health services
- the growth of medical audit (Royal College of Physicians 1989; Royal College of Surgeons 1989) has prompted a number of investigations into process and outcome (Hopkins 1990)
- a growing awareness of the uncertainty surrounding the benefits of some forms of medical practice and the resulting examination of variations in care (Ham 1988, Anderson & Mooney 1990) and appropriateness of care (Soloman et al 1986, Chassin et al 1987a; Chassin et al 1987b) .
- the development of health status measurement itself - and its scientific respectability (Reisenberg & Glass 1989, Brook & Kamberg 1987).

These, and no doubt other, pressures have resulted in the growing interest in outcome measurement. Despite this enthusiasm there is still considerable uncertainty of the techniques to use and very little experience in the field.

Outcome measurement is a broad term which can use a variety of methods and measurement techniques in a variety of settings - as discussed in Chapter 2. The common factor between all these approaches lies in the need for a consistent definition of outcome . Donabedian defined outcomes as:

> *" a change in a patients' current and future health status that can be attributed to antecedent health care."* (Donabedian 1985 p256)

This brief definition is important in identifying two of the recurrent themes of outcome measurement. The first is the idea of the patient's health status and how we define and measure that. Second is the notion of causality and the difficult questions of when (and how) health care interventions bring about the desired changes in health.

Aims of the study

The study has been conducted in a number of different specialties at the Freeman hospital between from 1988 to Summer 1991 using specific patient types covering cholecystectomy, diabetes, coronary angioplasty, care of the elderly, rheumatoid arthritis, urology and knee replacements.

The study has collected and reported outcomes to consultant teams or specialty management groups proposed in the resource management process (Freeman Hospital 1990). The initial aims were to integrate information into the resource management process and so use outcomes data with clinicians and managers. One of the reasons the project was based at the Freeman was the level of its information systems and knowledge in information handling - which were felt to be 'above the average' of those in the country. In addition, as a pilot site for resource management, the hospital was about to embark on a series of organisational changes which would seek to encourage the role of clinicians as managers and the use of activity and financial information.

The development of indicators should take place across a whole specialty so that the resulting measures could be integrated with resource management information. As a start down this route, this pilot study was restricted to selected conditions in secondary care.

The speed of change required for resource management to develop and become embedded in the hospital was overestimated. Often the organisational and information structures required for resource management were not yet in place (or just coming into place). There was therefore no easy marriage between the resource management process and the rapidly developing outcomes information. Review of the outcome information was therefore focused on groups of relevant clinicians and local managers in the first instance. At the same time the value of the information to a wider audience - including managers (clinical and non-clinical) has always been considered.

In many respects the examination and review of the outcome information was in setting

similar to that of clinical audit (Shaw 1989a; Shaw & Costain 1989). Certainly work with clinicians followed the cyclical process of audit (Royal College of Physicians 1989; Batstone 1990) namely identifying the standards of care, the means to measure these and reviewing results to identify necessary actions.

A major thrust of the work was concerned with evaluating how clinicians used the information and in particular its value in enabling them to 'close the feedback loop' without which audit has been memorably described as 'a pious exercise in self-congratulation'(Smith 1990).

The project used a number of working hypotheses as the basis for measuring progress towards the desired aims. More specifically these were:

a. That it is possible to agree with clinicians acceptable operational definitions of outcomes and identify the information to monitor these.

b. That it is possible to collect the agreed outcome information and report back results.

c. That outcome information would prove useful either in clinical practice, in the running of a local service or in wider management issues ? That is that outcome information could and would lead to practical change?

These have been addressed in this thesis as well as more general issues examining the similarities and differences of approach in the different specialties.

It is important to stress that the outcomes obtained were measured relative to the contribution made by the hospital. They therefore attempted to represent the marginal health benefit to patients receiving care in this setting. This may not necessarily correspond to the potential outcomes for the population of eligible patients nor can the

results describe what would have happened anyway or what happens to those patients who do not receive treatment. These are inevitable limitations to this type of approach.

Chapter 2 discusses some alternative approaches to outcome measurement that have influenced the measures and methods used in this study. The actual methods of agreeing outcome indicators are described in Chapter 4. Chapters 5 to 8 give the detailed results obtained from the various specialties and includes information on the indicators examined and the descriptions of changes in each indicator, relationships between indicators and a preliminary analysis of the effect of presenting patient characteristics on outcome expressed in these terms. Finally at the end of each of these chapters is a short section describing how the results have been used to date.

Chapter 9 gives an overall summary of the results obtained from the different specialties in general terms whilst Chapter 10 discusses general conclusions from the project and the potential for the wider development of outcome measures within the health service. The appendixes provide examples of the questionnaires used, reports generated and notes of discussions with clinical staff.

The project team has been managed by the author of this thesis who has been reponsible for discussing and agreeing outcome indicators with clinical staff, the development of methods of data collection and reporting as well as analysis of the data.

Chapter 2 Approaches and purpose in outcome measurement

Some ambiguity exists about the methods and purpose of outcome measurement. The term can often embrace a number of distinct methods in different contexts as Lohr points out:

> *"Patient outcomes are an immensely complex construct. They span the range of results that proceed from (or are presumed to be associated with) the provision of health care services. They are measured both directly and indirectly over differing periods of time and with varying degrees of objectivity, reliability and validity."* (Lohr 1988 p38)

Examples of the range of potential uses of outcome measurement includes those listed in Table 2.1 where a variety of approaches are described.

As Table 2.1 shows, the assessment of outcomes using standardised methods of assessment may be part of basic clinical practice. Whilst basic clinical care should naturally involve assessments of changes in health, the methods by which such assessments are made are typically proxies (of varying validity) for real changes in health status. There have been those who advocate using sophisticated outcomes assessment tools as part of routine clinical practice (Deyo & Inui 1984, McEwen 1988, Dickinson & Young 1989, Nelson & Berwick 1989, Nelson et al 1990).

In particular there is claimed to be an advantage in using standardised instruments, typically patient completed questionnaires, to assess general health from the patient's perspective and improve doctor-patient communication. Nelson and Berwick suggest some specific barriers that must be overcome if such tools are to be used in clinical practice and these include the practical issues of convenience of administration and acceptability to patients (they should be simple and not confusing). However they suggest that a greater obstacle lies in the lack of conviction among clinicians of the value

Table 2.1. Different approaches to outcome measurement

Approach	Case types -coverage	Issues/aims
Clinical management	Individual patients	Assessment of outcomes of individual patients for clinicians.
Clinical trial	Strictly controlled, specific patient types.	Comparing between therapies with controls possible with randomisation and typically strict entry criteria
Comparison of process	Diagnostic or treatment categories across sites	Correlation between outcome variation and different processes of care. Post-hoc standardisation for confounding variables.
Clinical audit	Diagnostic or treatment categories - usually one site.	Examination of outcome achieved against expected standards in your area.
Routine Quality controls (Provider)	May be with reference to one case type, specialty or hospital wide.	eg adverse occurrence screening avoidable deaths
Routine Quality controls (Purchaser)	Specialty/service type or resident population.	Specific clauses in contracts. Standardised Mortality Ratios. Notifiable diseases.
Planning & resource allocation	Service type or specific or new treatments	Comparison of benefits obtained. Comparisons of cost and benefit.
Confidential enquiries	Adverse events in particular services	Detailed examination of individual cases by expert panels.

of such tools (Nelson & Berwick 1989). A similar conclusion has been offered by Deyo and Patrick (Deyo & Patrick 1989) who advocated better education of health professionals and:

> *"..a 'laboratory' to provide measurement services to investigators and clinicians may make use of these scales more attractive"* (Deyo & Patrick 1989 pS254)

More commonly, outcome measurements are part of detailed clinical research or trials. These compare outcomes for a very specific subset of patients and ideally with double blind randomisation of treatment or control groups (Cochrane 1972) to assess the effectiveness of different therapies. The current view is that such trials should include not just technical measures of success but also general health status measures (Brook and Kamberg 1987; Bombardier et al 1986). However it has been pointed out that quality of life measures are too often considered secondary to 'medical outcome' measures and are *"added as an afterthought once the design, data collection and analytic techniques have been specified"* (Bergner 1989 P S148).

The observation of wide medical practice variations has prompted a variety of studies to examine the relationships between variations in process and the resulting outcomes. Roos and colleagues have usefully identified some of the alternative strategies adopted to examine this question and distinguish between studies according to the level of data collected (Roos et al 1990) ranging from cross sectional studies using hospital discharge abstract data to more detailed prospective analyses of longitudinal changes in cohorts of patients.

A popular approach at the moment is to use outcomes in evaluating non-controlled comparisons of process. A number of studies use retrospective or prospective examination of outcomes for one or more different processes of care. Thus for example one large multicentre study used sophisticated outcome measures to examine different approaches to prostatectomy (Wennberg et al 1988; WHO 1988). This model is currently favoured in the USA where national funding is supporting a number of Patient Outcome Research Teams (PORT's) (Salive,Mayfield and Weissman 1990). These teams are made up of groups of experts who are investigating outcomes of prostate disease, acute myocardial infarctions, knee replacements, cataract management and back pain.

Another approach, the Medical Outcomes Study (MOS) in the USA developed as a large multi-centre comparison of outcomes for a number of different disease categories (Tarlov et al 1989, Wells et al 1989, Stewart et al 1989). Its aims were twofold - the first being

an exploration of the effects of different organisations and styles of care on outcome (and was related to earlier RAND studies (Brook et al 1979)). The second aim was the development of more practical tools for monitoring outcomes. On this latter point the study led to the development of a relatively short health status questionnaire (Stewart,Hays & Ware 1988) as well as tools for examining patient satisfaction (Ware et al 1990). Moreover, the project championed practical outcome measurements across a range of disease types in the way advocated by Ellwood as part of 'outcomes management' (Ellwood 1988). His view is that information on outcomes should exploit sophisticated computing technology, statistical methods and measurement techniques to provide ways of assessing the effectiveness of different processes thereby guiding the practice of medicine, the management of health services and the choices made by patients. Though others are more cautious about the possible benefits (Epstein 1990) there appears to be considerable potential in the overall approach.

A more localised approach is to use outcome measurement in some form of medical audit as this study has done. There has been a growth in the field of medical audit in the past two years (Shaw 1989b; Hopkins 1989). Though most audit activities typically examine process of care (Gruer et al 1986; Spiby & Frater 1989) there is a growing interest in the addition of outcome measures. However, with a few notable exceptions there is little practical experience (Hopkins & Costain 1990) of their development and use.

Beyond the introspective processes of audit within a department there is the potential for examining outcome at a hospital or service level as part of an overall approach to quality assurance. Once again process measures are more typical, but there are some examples of where outcome indicators can be used in this setting. One is the practice of monitoring the frequency of 'adverse occurrences' recorded in hospitals (Craddick & Bader 1983; Brennan et al 1990). Such occurrences may be direct or indirect measures of patient health for example peri-operative deaths, and act as a trigger for further investigation. Other alternatives include the use of specific mortality or morbidity data (Middleton 1987).

At the level of district or regional health planning, outcome measures typically consist of mortality data either as standardised rates or relating to specific disease types as in the indicators of avoidable deaths (Charlton et al 1983; Charlton,Bauer & Lakhani 1984). Similar techniques are advised for district and regional health authorities seeking to monitor the general health and outcome of their resident population. Such indicators provide an input to planning and strategic management of services, typically by identifying potential problem areas in current provision. Though there is much speculation on the inclusion of outcomes in the contracting process, there is as yet little practical experience of which measures to use.

At regional or national level outcome measures may be associated with evaluations of specific programmes or services , for example the evaluation of heart transplant programmes (O'Brien et al 1988). Similarly national audit systems such as the Confidential Enquiry into Peri-operative Deaths (Buck,Devlin & Lunn 1987) exist. Such studies examine individual cases as well as aggregating data to consider issues on a wider scale and identify where processes of care are linked to outcome.

Finally there are questions of resource allocation using singular measures of health benefit -like the Quality Adjusted Life Year (Boyle et al 1983; Williams 1985 ; Gudex 1986). Such studies are not aimed at the level of individual patients but assess the need for investments in different health care programmes by comparing costs against expected benefits to health. Such comparisons are not without problems (Drummond,Teeling-Smith & Wells 1988; Smith 1987).

As well as variety in the applications of outcome measures, there are a number of different methodological approaches or instruments that can be used.

Clinical/laboratory measures of success

Later chapters describe a variety of measures which have traditionally been used by clinicians to determine the success of treatment, for example changes in anginal pain (Cambeau 1976) following PTCA or levels of blood glucose in diabetes (Nathan et al

1984). Such measures form the basis of most clinical research and are critical to the development of medical knowledge about the success or failure of treatment. One of the problems with such measures is the extent to which they measure a physiological process rather than a true patient outcome. As Fries observed:

> *" Process measures achieve their value only to the extent to which they serve as accurate proxies for outcome measurement. Thus sedimentation rate accurately quantifies 'fatigue', and if 'fatigue' is part of patient discomfort, and discomfort is an outcome, then sedimentation rate may have value as a surrogate outcome measure, conveniently and accurately ascertained."* (Fries 1983 p 697)

Thus the use of such clinical measures must be based on a belief that they represent (to the best of our knowledge) reasonable proxies for genuine patient health problems.

Measures of standardised mortality rates

The measurement of mortality rates is one of the best established of all outcome measures. It can be expressed relative to a geographical region (standardised for age, sex, and other possible factors) as an index of health among a population and by implication as an indicator of the performance of health services. A more specific approach is to consider mortality rates for specific conditions amenable to treatment (Rutstein 1976, Charlton et al 1983, Charlton Bauer & Lakhani 1984). The main problem with such measures include the frequency with which deaths will occur and the resulting statistical significance of variation, the relevance of mortality as the sole description of outcome of care and finally the problems of linking the processes of care to the outcome. Nevertheless mortality must always be considered part of the range of outcomes.

Measures of morbidity

In some circumstances the observation of the existence or prevalence of certain disease states within a population can be taken as an indicator of the outcomes of care provided. Typically such measures will reflect on the outcome of primary care, prevention or screening. For example admissions to hospital of patients with acute diabetic ketoacidosis may be considered to reflect on the local health service's ability to detect the signs of diabetes.

Measurement of general health status/quality of life/well-being

The past few years have seen a growth in the acceptability and application of measures which attempt to define general health status (Spilker at al 1990) as described earlier. In the first instance there were a number of measures which attempted to represent a broader definition of patient health than simple clinical indicators. One approach was that developed by Rosser and colleagues which was based on professional judgements of patient health along axes covering disability and distress (Rosser & Watts 1976) and has been used in an attempt to weight the value of different states (Rosser & Kind 1978). The resulting matrix has been used in cost utility analyses (Williams 1985).

In contrast to measures derived by a professional assessments, a variety of tools (usually patient completed questionnaires) have been developed which attempt to identify key constructs of ill health drawn directly from the patients. These are increasingly being used as part of outcomes assessments and are discussed in more detail later.

Measures of patient satisfaction

In addition to changes which are considered part of patient health, there has also been a growth in studies examining patients' views on the acceptability or otherwise of the process of care (Cleary & McNeil 1988, McIver & Carr-Hill 1989, Fitzpatrick 1990).

Quality Adjusted Life Years

The examination of cost-effectiveness requires instruments to examine both relative costs and utility as in the QALY - quality adjusted life year (Gudex 1986). Such measures are typically applied to examine different health care programmes rather than individual patients or treatments. Central to the concept is some measure of health benefit that is common (and commensurable) for different patient types. As such the calculation of QALYs is secondary to the examination of changes in patient health undertaken in outcomes measurement.

Adverse occurrences

A popular method of quality assurance that has been widely used in the USA is the

monitoring of specific health events which can be considered to be undesirable or avoidable consequences of treatment (Craddick & Bader 1983). Thus for example post-operative complications can be considered a short term outcome measure.

Readmission rates

The monitoring of readmissions to hospital has the advantage that much of the data required is already collected in some form and has been advocated as a useful measure of quality (Chambers & Clarke 1990). In this case the problem is to identify when the fact of readmission is an appropriate proxy for deleterious changes in health status.

In conclusion it appears that the field of outcomes measurement is broad and encompasses a variety of different approaches and techniques. None of these exactly matches the requirements for this study. Therefore a variety of methods from the literature have been used in the formulation of appropriate outcome measures. Given that the main purpose of the study was to develop measures which were of value to clinicians and managers, a prospective approach has been adopted and one that seeks to examine changes in patient health status following treatment. Given the uncertainties over what were the right measures to use and over the performance characteristics of the various measures, a range of measures were used including clinical and laboratory values, general health status, adverse events, mortality and readmissions. Though it was recognised that ideally patient satisfaction should also be included such measures were rarely used in the belief that the exploration of their use was being undertaken elsewhere.

Chapter 3 discusses the methods used to identify the relevant indicators and the methods of data collection.

Chapter 3 Methods - Identifying outcome indicators and methods of data collection

A. Organisation

The research team has changed in structure during the project. It eventually consisted of two research assistants (nurses) and a project manager (the author of this thesis). For the first year there was medical input from a community physician. The team was based in an office centrally within the hospital. In addition support was given by the Department of Public Health of Newcastle Health Authority.

Progress was monitored through a steering group chaired by the authority's Director of Public Health, and consisting of representatives of the various clinical groups and managers. Steering group meetings were held every 1-2 months throughout the project. The current membership of the steering group is shown in Appendix 1

B. Developing Local Outcome Criteria

The project worked within individual specialties and in each attempted to outline those changes in patient 'health' that occurred over time after a particular intervention. For acute cases the study compared observations made before hospital treatment to similar observations at discharge, 3 months and/or 12 months later. For chronic conditions observations were made at each contact with the patient and/or at annual review. The changes identified were then compared to agreed standards of what one might expect to have occurred over these time periods.

In each specialty the following steps were needed to establish the agreed range of outcome indicators and methods of data collection:

i) Choose suitable reference group of patients (or treatments) for study

ii) Identify expected changes in health over relevant time intervals

iii) Identify potential effect modifiers

iv) Identify data required to monitor changes

v) Identify best methods of data collection

vi) Pilot and undertake data collection

vii) Validate data

viii) Feedback results, review database and monitor any actions resulting

These stages are described in more detail below.

i) Choosing Suitable Reference Groups

In most specialties it was not practicable to monitor routinely a wide range of outcome measures across all possible case types. Measures such as death rates, readmission rates etc might be collected at this level but more detailed information would need to be restricted to a subset of specific treatment or patient types. The choice of a reference group was used to focus discussion, the choice of group being guided by the following criteria:

* Volumes of cases had to be sufficiently large to enable reasonable statistical results and represent a sizeable proportion of the specialty workload
* Significance in terms of health to the patient
* Capable of reference to an agreed standard
* Relevance to specific processes or areas of contention or uncertainty in care
* Practicability of data collection

In many circumstances the outcome measures were relevant across different disease groups - for example avoiding death or 'improving well-being' applies to almost all patients. However even with these common measures there will be variations between case types in, for example, the choice of relevant time periods between measurements,

the effect of different presenting characteristics and the likely standards for comparison.

The scale of the made it necessary to examine a small number of well-defined patient groupings rather than seeking measures which would necessarily embrace all case types within a specialty. It was recognised that common indicators or data elements may arise. The use of selected conditions as tracers also meant that the data collection methods could be tested and improved gradually.

ii) Identifying expected changes in health

Having agreed a reference group the next step was to agree the outcomes one might expect to see and at what time periods. For some conditions - typically those resulting in an acute inpatient episode - the outcomes were the changes noted between admission to hospital and 3 months and 12 months later. In the case of chronic conditions, it was necessary to consider longer time-scales, the outcomes being concerned with marginal changes in the patient's health over time. These may be in terms of positive improvements in the patient or alternatively in the avoidance of a deterioration in health. The first time period typically started at the first contact with the service, for example newly diagnosed diabetics, and continued until a follow-up clinic or annual review. There are no clearly defined endpoints in chronic care.

The indicators that were discussed typically considered different dimensions of outcome as potential candidate measures. They ranged from specific clinical and laboratory assessments to more general descriptions of health and the patient's 'quality of life'. It is becoming increasingly accepted that the patient's perspective is critical to assessing the effectiveness of health care and there is growing experience in polling patients on different aspects of their health, quality of life or satisfaction with treatment (Fries 1983; Bergsma et al 1988; Leighton-Read et al 1987).

Where possible expected standards were agreed but given the novelty of some of the measurements - for example those attempting to reflect patient's well-being - this was not

always possible in anything other than general directional terms such as "all patients should show improvement".

iii) Identifying potential effect modifiers

In monitoring the outcome of care it was important to identify those characteristics of the patients which at presentation are known or thought likely to affect the eventual outcome. This is particularly importanat in the case of outcome monitoring in secondary care where potential improvements may be limited by a patient's previous medical history. These variables can be used to define subsets of patients which are expected to achieve different outcomes. Thus for example the expected outcomes for young, otherwise healthy cholecystectomy patients may be different from that expected for elderly patients presenting with significant co-morbidities. The practical importance of these variables in explaining observed differences in outcome were evaluated once results became available.

In this area there was a tendency to be over-precise at the outset in defining what were the relevant patient characteristics to consider. Each extra item of information has a cost in terms of its collection and analysis and this had to be balanced by the importance of that additional data in assisting the interpretation of the results. A distinction should be made here between those variables which are always likely to affect the outcome of care eg age which may require the database to be subdivided before analysis, and those where the particular presenting characteristics are important in relatively rare circumstances. These will not require separate groupings of patients but can be considered as possible explanatory variables when results emerge. For example very obese patients can make cholecystectomies more difficult and one might expect the outcomes for these patients to be worse. However in practice relatively few patients are sufficiently obese for this factor to become important. Where this is the case it may be that only a few will suffer outcomes that are significantly worse than other patients and thus though obesity can be a subsequent explanatory factor as to why any one patient achieves a worse outcome - it is not a sufficiently discriminating variable by which to subdivide the patient database *a priori*. The key issue in determining potential effect modifiers is the strength of the

causal link between the nature of the patient, the process of care and the eventual outcomes.

iv) Identifying data required to monitor changes

In most cases the relationship between the outcome indicator and the type of data required to create it is straightforward. Looking at outcomes from the clinician's perspective, it is reasonable to take the view, at least initially, that if the information is important for the clinical assessment of a patient it will already have been collected in some form. Such basic clinical observations and descriptions of patients while they are in hospital should be mirrored in the medical records. If this is not the case then one must ask "Is this information really important?" or "Is its absence attributable to poor note-keeping?"

Even if the information was already collected, some changes were needed in the method of recording or the definitions used. Such changes should improve the general consistency of recording.

It was agreed that the clinician's perspective was not the only viewpoint from which the outcome of care should be examined. The patient's perspective was also important in assessing the success of an intervention and it was here that new data collection systems were required. The choice of instrument is discussed later but in general an existing established instrument was chosen where possible.

v) Identifying the best methods of data collection

Although not constrained by the original design, the methods of data collection which have developed have shown strong similarities across specialties although not all specialties require the full spectrum. Some specialties use only a selected combination of those available. The methods used are shown in Table 3.1

Table 3.1 Summary of information capture methods

	Source
Presenting demographic /clinical details	Notes Computer systems
Patients symptoms/problems /health status	Patient completed questionnaires in hospital or by postal follow-up
Details of events in hospital	Pro-forma completed by clinicians or from notes/computer systems
Details of events after discharge	Patient, other hospital computer systems, GPs.

Though existing data sources, when supplemented with information from patients, could provide much of what was required, at some stage there had to be extra data collected by clinical staff. Additional data collection of this kind had to be justified on the grounds that:

- it replaces/improves existing data collection methods
- it provides other significant benefits as information
- it is considered essential for the review of outcomes

vi) Piloting and undertaking data collection

The pilot stages of data collection in each specialty raised questions over the feasibility of collecting certain items of information, particularly on a continuing basis. Different ways of data collection emerged and these were evaluated in terms of the reliability and consistency with which data could be recorded and the value that information will add. Issues also emerged over the definitions used, for example how is unstable angina defined or what constitutes a relevant co-morbidity? The results of

these are presented in later chapters on the data collection methods in individual specialties.

The review of data collection systems and definitions formed a part of the feedback process and the discussion of results. The data collection systems had to be capable of undergoing an iterative development which responded to demands for new items of data without losing the necessary stability of key data items needed for longer term trend analysis.

vii) Validating data.

Considerable efforts have been made to check the reliability and accuracy of the data collection systems. It became clear during the course of the project that many data items could present problems in terms of either variation in definitions, the consistency of recording, missing data or inter-rater reliability. A variety of methods were used to check on the validity and accuracy of data, the results of these exercise are presented later where appropriate.

The methods included:-

- re-abstracting data from notes to compare to questionnaires
- checks versus other computer data bases
- interviews of patients
- repeating tests (typically on patients)
- checking of empirical relationships between outcome indicators.

viii) Feeding back the results

The purpose of collecting outcome information was to inform clinical practice and identify possible areas for change. This was undertaken within a clinical framework

with results being presented at regular meetings with the clinical staff concerned and, in some circumstances, managers.

Data were fed back which consisted of:

- summaries of changes in individual patients
- specific reports on patients with 'poor outcomes'
- changes that have taken place on an aggregate basis, and when subdivided into key groups. Some of these groups will inevitably be defined by patients' presenting characteristics.

Assessing the effectiveness of outcome information.

The assessment of the effectiveness of the information itself was based on three approaches.

1. Data for each specialty were reviewed at appropriate intervals with the clinical staff concerned. The data were presented in the form of a report and discussed with a pre-arranged agenda. Minutes of each meeting were prepared and particular actions noted. In addition reports were circulated to the Steering Group or to other interested parties who may not have been at the original meeting eg nursing staff, managers, clinicians in other hospitals.

2. A diary was kept in the CASPE office to record relevant events during the course of the study. In particular a wide range of responses from the clinicians was noted including areas where they had either expressed a particular interest or concern about the information. Similarly it was noted if particular actions had resulted from the outcome information - or the presence of the study. Though not wonderfully effective, the diary helped keep track of some of the issues.

3. At the end of the project clinicians were asked some general questions about the

success or failure of the project. The issues are discussed in Chapter 9 and the responses to the survey in Appendix 3.

These various sources of information have been collated and presented in the results sections of this study.

C. Outcome information and resource management

The Freeman hospital was one of the national pilot sites for resource management. Two reasons why the study was based at the Freeman were the basic standard of the information systems in the hospital and the interest of clinicians and managers in developing information systems to assist in the management of clinical services. It was also hoped that the organisational change associated with resource management would provide a focus for the use of outcome information.

In particular the use of the main hospital information system was explored. Methods of linking the basic information on the main hospital systems (the case-mix computer) to the micro-computer used in this study were developed. Ideally the link is based on hard wiring the micro-computer to emulate a terminal to the main system. Software which periodically dumps data from the case-mix system (either directly or through floppy disks) to the micro-computer was developed. This could then be linked up to the outcome data bases.

The advantages of this type of link are that it avoids unnecessary inputting of common data - including diagnostic and demographic information. It also allows for checking on re-admissions and provides access to a larger database of patient episodes. These links have been shown to be possible though they have not become part of the basic routine this study. This was mainly because the efforts required to write the software would not be justified by the volume of cases in this pilot study. It was easier to re-type the data where necessary.

D. Selecting measures of patient function and health status

From the early stages of the project it became clear that a picture of outcomes would not be complete without some estimate of the patients overall function/health status/quality of well-being/quality of life - the terms are often used inter-changeably. Their use also differs slightly across the Atlantic, for example measures of 'Functional Status' in the USA include what this report would tend to describe as general health status measures eg the Sickness Impact Profile (Brooks et al 1990).

The chosen term for this study is usually health status - in effect attempting to operationalise a definition of health. Quality of life was considered a broader description and included elements which may not have been able to be influenced by the health service. Functional status is taken as a more limited measure of the patient's physical abilities.

There are a great many instruments which have been developed in this area - one recent review of the literature merited a supplement to Medical Care and ran to 70 pages (Spliker et al 1990). These measures vary in what are deemed to be the relevant dimensions of health though there are often common themes, for example Ware suggests five generic health concepts :- physical health, mental health, social functioning, role functioning, general health perceptions (Ware 1987). Not all measures cover this spectrum. Some measures (usually older ones) specialise in basic activities of daily living and basic patient function, for example the Barthel Scale (Mahoney & Barthel 1964). Others are almost exclusively concerned with psychological well-being .

Not all measures are claimed as appropriate for all case types, some are used in analysing only a particular disease or population groups, for example two standards in arthritis include the Health Assessment Questionnaire (Fries et al 1980; Kirwan & Reeback 1986) or the Arthritis Impact Measurement Scale (Meenan et al 1984). It is unfortunate that the use of these scales tends to cluster in certain patient types (eg

Arthritis) leaving areas where none have been tried. A great number of these indicators have been developed in one particular setting and have been tested by other researchers.

Choosing an instrument was therefore difficult. The criteria used to select an instrument were therefore:-

a. Established and reasonably well validated tool used in more than one study in the UK
b. Used in similar patient group - or if not the questions were deemed to be relevant to that patient group.
c. Expected to show longitudinal change following intervention - in many cases there was little evidence of where scales had been used to show changes after treatment.
d. As quick and easy as possible to administer - which effectively means self-administered by the patient and no longer than necessary.
e. Other things being equal a generic measure rather than a specific one.

These questions arose at some time in each specialty. There was the potential to standardise on one instrument across all cases yet this was rejected in favour of local specificity. Thus the goal of developing indicators relevant to a particular specialty was placed above the ability to make comparisons between specialties.

In fact as the project developed there was a growing interest in comparisons between conditions and by the end data was available from the same instrument on a variety of case types. In cholecystectomy and angioplasty the instrument initially selected was the Nottingham Health Profile henceforth NHP (Hunt at al 1986). This is a 38 item self-administered questionnaire scored according to the weighted sum of positive responses on six dimensions energy, pain, social isolation, emotional reactions, mobility, and sleep. In addition seven separate questions on how health affects lifestyle are included in Part II of the scale.

The NHP was selected for its ease of use and application and most importantly,

though it was originally developed as a population survey tool it had been shown to demonstrate longitudinal change (Buxton et al 1985; O'Brien et al 1988), and had been used for surgical patients (Hunt et al 1986; Black, personal communication).

The instrument has been criticised for its sensitivity at differentiating illness (Kind & Carr-Hill 1987) when its modal response in fairly healthly people is zero and because there is co-variation between its dimensions and redundancy in items. The problem that most people score zero on the scale is largely accepted , and this had to be recognised as an important limitation in some settings. However others have found the instrument to be sufficiently sensitive to differentiate within and between patients with chronic illnesses (Jenkinson et al 1988). In practice it soon became clear that the instrument did appear to be working as it should in the groups we studied and later results show the demonstrable changes that occur following treatment and differences between patient groups.

The weighting system has also come in for some recent criticism (Jenkinson Pers.Comm.) and the value of the actual weights and the methods used in weighting is questioned. The data collected in this study could be used for further investigation of this issue.

Despite this the NHP is still one of the most widely used of all such tools and held the most promise. The biggest area of uncertainty was over its ability to show change in the short term for elective surgery - a doubt expressed to us at the outset by one of the scale's developers.

The NHP provides six separate scores - one for each dimension. It is suggested that these are not aggregated but considered separately as a profile of the patients health. For most of the project results and all reports to clinicians this was the case.
In most cases the results in this report are presented in that way. However in this report aggregated scores were also calculated. There are a number of possible methods of aggregation (see Buxton et al 1985). The one chosen was simply to weight all dimensions equally and calculate a combined average. The aggregation was

used to explore when relationships to other variables may exist and simply to reduce the volume of tables.

The other measures used included the Sickness Impact Profile (Bergner, Bobbit & Carter 1981) used in diabetes. In this case the NHP was rejected largely on the grounds that too few patients would score. The complete SIP runs to 136 questions in 12 categories which was felt to be too much. However the categories are scored independently and so the diabetes results were based on only four categories selected after a trial of the full SIP on fifty patients. The four categories were chosen as representative of the full instrument and because they tended to elicit most responses.

More specifically they were:-

Ambulation (from the physical dimension)
Social Interaction (from the psychosocial dimension)
Home management/housework (independent category)
Recreations and pastimes (independent category)

In Rheumatology (and early in orthopaedics) the Health Assessment Questionnaire was selected (HAQ)(Fries 1983). This instrument has almost come to be standard in Rheumatology though it is not without its problems (Leighton-Read,Quinn and Hoefer 1987; Fitzpatrick et al 1989). In particular it is very 'functional' in the sense that it is concerned mainly with the ability to perform everyday tasks. It represents therefore a narrower view of health than either the SIP or NHP. There was also some concern expressed about its ability to differentiate between 'severely ill' patients, where the maximum score is felt to be insensitive to beneficial change. For these reasons comparative studies of the NHP and HAQ are underway in rheumatology.

After early trials with the HAQ in orthopaedics a decision was made to move to the NHP where the items were considered more relevant. The HAQ proved not a particularly inspired choice for knee replacements as much of it concerns basic function of the upper body.

Finally it is worth noting a recent development from the Medical Outcomes Study where one questionnaire (36 items) is felt to be applicable to a broad spread of case types (Stewart et al 1989). This would be an attractive option for similar work provided its applicability in the UK could be demonstrated.

Validating NHP scores through patient interviews.

Given the uncertainties about the choice and use of the self-completed health status questionnaires, studies of the validity of the NHP were undertaken to check for possible problems in completion of the forms and to identify whether observed scores were related to those made by an independent assessment. Validity can refer to a number of different characteristics of a measurement and using different terms - construct, content, face, convergent, parallel, predictive, etc. A popular cliche has it that 'there is no gold standard' with which to evaluate such measures. In this case evidence of the 'construct' validity of the scores was sought by comparing observed scores to the judgements made by raters (clinically trained) following patient interview.

In three specialties (orthopaedics, general surgery and cardiology) a random selection of patients were chosen for validating postal questionnaires and NHP forms. These results were combined with a similar exercise undertaken to examine a stratified sample of diabetes patients (selected to include extreme NHP scores) to make a total of 56 patients.

The visits utilised a semi-structured interview covering the main points of the questionnaires. For any one condition the same assessor was used. Three different assessors were used in total. The researchers (who had not seen the patients' forms beforehand) were asked to make their own assessment of the patients health-related problems using simple categories on a four point ordinal scale ranging from 'No problems' to 'Severe problems' (Appendix 3). Interviews were semi-structured around key questions representative of the different dimensions of the NHP. Ratings

were made for each dimension. The results of the researchers assessment were then compared to the previously completed patient questionnaire. Table 3.2 summarises the results when the actual score across all the different cases (taken from four different conditions) are compared to the raters assessment (on a scale 1 to 4).

The analysis was based on the Kruskal Wallis oneway analysis of variance with ranks (NHP scores are typically not normally distributed) . The mean rank of the actual NHP scores are shown for the different assessments by the raters. Significant correlations are shown on all dimensions except energy.

There were significant differences between the mean scores for each condition so scores were standardised to remove the difference between mean NHP scores of each condition. The results are similar using these standardised scores. These relationships therefore hold despite possible variation in the way different raters scored patients.

This exercise suggests that most of the dimensions are roughly in accord with the perceptions of an interviewer. The lack of a clear relationship with regard to energy has - in retrospect - been attributed to the problems in framing the question at interview. It was found difficult to discuss what may appear as a rather abstract concept. It also interesting to note that the NHP itself has only three questions in this dimension and each is correspondingly weighted rather highly. This would tend to make the scale more 'volatile' and could account for the discrepancy. The empirical analyses reported later comparing the NHP dimensions to other health indicators has not revealed energy to be especially different in its behaviour. Given the performance of the scale as a whole we have therefore concluded that it is behaving largely as expected. In cholecystectomy some test-retest studies have also been undertaken (reported in Chapter 5) with satisfactory results.

Table 3.2 Mean rank of NHP scores against raters assessment and significance of observed relationship by one way analysis of variance by (a) ranks of unstandardised scores and (b) means of scores standardised for condition.

	Mean Ranks	Raters assessments of problems			Significance.	
	None	Minor	Moderate	Severe	(a)	(b)
Energy	25.1	24.9	27.8	38.1	.110	.872
Pain	18.0	28.1	37.6	38.9	.0023	.0005
Emot	21.9	27.3	31.8	43.7	.0083	.0087
Sleep	14.3	29.7	27.9	41.0	.0001	.0055
Soc Is	22.5	30.3	34.4	44.9	.0005	.0028
Mobil	11.4	28.7	40.4	38.2	.0000	.0000
Comb	16.1	30.2	41.9	53.0	.0000	.0016

E. Statistical Methods

A variety of statistical methods have been used in analysis of the data. Guidance on the choice of tests was mainly based on one source (Blalock 1981). Where the distributions of continuous variables were considered 'normal' the test included the t-test of differences between two samples, analysis of variance for examining differences in a continuous variable between categories and product moment correlation and regression between two or more variables.

Many of the data elements studied, and in particular the health status measures, were not normally distributed which required the use of non-parametric tests where there is no assumption of 'normality'. Comparisons between paired samples which was based on the Wilcoxon rank sum test and comparisons against categorical variables based on the Kruskal-Wallis analysis of variance with ranks or the Mann-Whitney U test where only two categories were compared. Association between two variable was based on Kendall's tau_b. Comparisons of the distribution of two categorical variables was based on tests of Chi-squared, with the use of McNemar's test for 2 by 2 tables.

Chapter 4 Care of the Elderly

A. Introduction

The inclusion of inpatient care of the elderly was the last major strand to be added to the study. The reasons for its inclusion were firstly to test whether outcome data could be collected about a very different type of care to that of the other studies, and secondly the interest of the clinical staff and a desire to be involved in the project. The exploration of ways to monitor outcomes in this patient group is one of the least developed in clinical care.

The inpatient geriatric service at the Freeman Hospital is primarily based on a model of rehabilitation described by Grimley Evans (Grimley Evans 1983). There are few emergency admissions, all patients arrive at the ward either as an elective admission or through one of the other specialties in the hospital.

The study was centred on the one 30 bedded geriatric ward, for which there are two consultants both of whom have been closely involved with the study. As well as the beds in the acute hospital the department also covers two satellite hospitals and includes a day hospital.

The development of outcome measurement for this group of patients is in some ways less well advanced than for others. For example many clinical trials specifically exclude elderly patients. There are a number of studies which have sought to evaluate care of elderly patients typically by examining particular institutional settings for care. Thus for example the effectiveness and efficiency of day hospital care has been examined (Donaldson et al 1987; Gilleard 1985; Macfarlane et al 1979; Reifler et al 1981) and compared to other settings (Sherwood,Morris & Ruchlin 1986).

The second area of work concentrates on specific disease groups within the elderly - of which the treatment of stroke has served as the model for outcome assessment (Van Swieten et al 1989). Other than disease or program specific studies there is relatively little known about the general outcomes of patients discharged from geriatric wards in a hospital.

B. Data set and data collection

Identifying the Outcome Indicators

The agreement on the outcome indicators and data set in geriatrics took place over a 2-3 month period and discussions included representatives of the various professional groups in the clinical team. The involvement of disciplines other than medicine has been greater in this project then in the others.

As discussions developed it became clear that the emergent indicators were applicable to the majority of admissions and not to specific diagnostic categories. The medical classification of geriatric patients is difficult because patients typically suffer from a number of different disease conditions. Given the complexity of the medical problems it appears that in care of the elderly the types of goals for most patients tended to be similar despite differences in the nature of the underlying medical condition. There were however two important exceptions to this:

- patients admitted for terminal care.
- patients admitted for investigation only.

It was therefore agreed to develop indicators that could be applied to all patients except these two types of admission thus effectively including the majority of the specialty's workload.

Outcome Indicators

The following list of outcome indicators contains one important omission. It was noted that ideally the patient's views on their own health should be polled but administration of one of the standard health status instruments (as in the other specialties) presented particular difficulties. Neither do the measures incorporate assessment of handicap that is the effects on a persons 'role' but concentrate on disability (there ability to perform 'tasks'). Whilst such measures would be desirable no wholly satisfactory instruments have been found.

Avoidance of death

It was agreed that in general there should be as few deaths as possible for admitted patients either during the stay or shortly after discharge. It was recognised that in individual cases death may not necessarily be undesirable, yet for the group as a whole it was. It is clear that the interpretation of this information would be different in geriatrics than from the other specialties.

Improvement in basic patient functional abilities

One of the features common to the care of elderly patients is the desire to improve some of their basic abilities to perform everyday tasks. There are a number of instruments available for such assessments (Applegate, Blass & Williams 1990). There is also a growing realisation that such assessments are essential in screening (Williams 1990) and monitoring elderly patients (Dickinson & Young 1990).

The choice of an appropriate scale took some time and involved testing - and eventually rejecting the Creighton Royal scale and the Office of Population, Censuses and Surveys (OPCS) Disability scale (Martin,Meltzer & Elliot 1988). The OPCS scale was recognised to be comprehensive and much more sophisticated than the alternatives. It was tried in practice by giving it to a number of professionals to complete. Though it showed a sophisticated breadth of issues the process of completing the questionnaire proved too unwieldy to be practicable. The major problems were its length and

uncertainty about the correct process for completion. It was recognised that as the scale would be completed by a number of different personnel, simplicity was of the essence.

Eventually the Barthel scale (Mahoney & Barthel 1965) was chosen. This is a simple scale assessing ten items concerned with daily living and scoring each on between 2 and 4 categories ranging from complete dependence to independence. Despite its relative simplicity this scale is one of the most widely used (Collin et al 1988; Wade & Collin 1988) in particular its content was considered relevant, the individual items covered the main objectives independently identified by clinical staff. Though there are a number of specific rules for completion these are relatively straightforward.

The simplicity of the scale also meant that it was possible to record not just the present status of the patient, but also the expected goals for that patient. Thus for each item improvement could be expressed relative to specific goals for that patient.

The disadvantages with the scale concern its lack of sensitivity for different patient groups, and modifications have been suggested, and the system of scoring. This is rather rudimentary and does not attempt to inter-relate scores across different items as some others do (Nouri 1987). As a measure of disability it is also criticised for its limitation to aspects of self-care and so does not represent the full range of problems that patients may face (Ebrahim 1990).

Finally it should be noted that the selected scale measures disability rather than handicap (Ebrahim 1990).

Independence of home environment

The home environment was recorded on admission, discharge and at follow-up on a simple ladder scale which ranged from living at home to living in long-term institutional care. A value judgement was made that, in general, the top of the ladder is 'better' than the bottom. It is important to stress that this judgement applies to the population of patients as a whole rather than any one individual.

Alleviation of acute medical condition

It was recognised that a key role of the service was treating acute conditions of the patients. However with complex disease patterns in these patients success in this respect is difficult to assess in an objective fashion - other than by the knock-on effects on patient's functional status. Therefore the clinician's judgement was used as a simple indicator. In particular staff were asked to assess the patients against three questions:

- Has the medical condition improved?
- Is the patient better than before?
- Have the overall goals been achieved?

It was intended that these be used more as a method of identifying specific patients where treatment had not been successful rather than as an aggregated indicators of performance. In practice the data has been little used but served as a useful check on change in patients when viewed with other indicators.

Minimal deterioration in mental health/awareness

In addition to changes in basic function it was recognised that there should not be deterioration in the patient's mental condition, and in some cases a positive improvement. The ten point mental health score (Qureschi & Hodkinson 1974) which forms part of the admitting routine has been used. Whilst collection of the data on admission has given few problems it was not routine to administer the test again on discharge. Consequently there were very few patients with observation at both points. Therefore the information has been used to identify patients with severe confusion as a complication of their presenting condition.

Minimise strain on main care givers where appropriate

In some cases it was recognised that an additional goal of the service was to reduce the burden on informal carers. There is considerable research into the emotional well being of care givers of the elderly demented. The evidence suggests that carers can suffer from a variety of health-related problems including psychological problems, stress and stress-related physical problems (Gilhooly 1984; Gilleard et al 1984) though the evidence is not

unequivocal (Eagles et al 1987). A variety of scales have been developed (for a review see Baumgarten 1989) but there is little consensus on which instruments to use in which circumstances and little comprehensive testing of the validity of scales in different contexts.

In this study, the stress on the care-giver was considered most important and therefore a simple questionnaire was used (Robinson 1983). This was given to the main care givers (when appropriate) and sent out at follow-up. Initially, this was used with all patients. However in many cases it was clearly inappropriate, for example where there was no identifiable main care giver or when the patient was in the charge of professional carers. Therefore it was decided that the instrument should be used on a limited selection of patents identified when:

- the patient had an easily identifiable main care-giver (and not multiple professional carers)
- the care-giver lived close enough to be directly involved in day to day care
- the main care-giver was physically able to complete the questionnaire eg not blind, or did not have severe arthritis in the hand
- the patient was not in residential or long-term hospital care
- the patient was not admitted for terminal care

Timing of observations

The relevant timing of measurements was agreed as at admission (to give the baseline values), at discharge and then three and twelve months after discharge. The identification of the individual goals for the patient related to the time of discharge and were not collected at follow-up. It was recognised that the initial baseline values would correspond to a period of the first few days of the stay rather than the actual day of admission. In the event of patients being readmitted to the Freeman Hospital, a new episode was initiated and the previous episode was considered to have ended.

Data Collection

Patient identification/baseline data

Most of the patients admitted to the care of the elderly ward were eligible for the study. The exclusions were those admitted for terminal care or investigations only. The baseline information on age, sex, diagnoses etc. was taken from the notes soon after admission. This data included the mental test score and previous medical history which are part of the normal clerking information collected by medical staff. This data abstraction onto a pro-forma was undertaken by a member of the medical staff (CASPE staff standing in when necessary).

Functional status and accommodation

The main assessments of items in the Barthel score and goals for the patients were originally designed to become part of the social round - when discussions amongst the clinical team review all patients. It was the designated task of one member of staff to note down the issues as and when they were discussed. In some cases the use of the structure provided by the Barthel was said to have helped clarify discussion on individual patients and to ensure that all the relevant issues had been addressed - rather like a checklist. There does seem to be some potential benefits for using such scales in this context (Dickinson & Young 1989). However this practice has tended to lapse with a change of personnel. If the data was not collected during the social round then it was completed afterwards by the medical staff in discussions with the primary nurse for that patient.

Care-Givers Strain Scale

It had been agreed that it was the nurse's responsibility to administer the care giver strain scale where appropriate. The relevant individuals and their addresses were recorded in a central register (something that was not in place before). Carers were given a form to complete on the ward by the nurses.

Follow up scales were sent out by post after preliminary checks on the patients whereabouts and checks to ensure that the patient was still alive.

Discharge information

At discharge the summary details of the stay were completed by the medical staff together with the further Barthel score (from the social round).

Follow-ups

The process of following up patients after discharge has proved a problem. The main difficulty has been that postal methods have not been used. Follow-up information was therefore obtained:

- at out-patient appointment for those patients who have a scheduled visit. The relevant forms are added to the notes and completed by medical staff.
- at the day hospital for those patients who are visitors.
- from satellite long-stay hospitals where medical staff will be seeing the patient or by telephone to local nursing staff.
- from local nursing homes where after initial approaches a telephone call is used to check on the status of the patients.

In some cases when the patient was not seen directly, staff at other institutions were asked about the status of the patient over the telephone. The simplicity of the Barthel scale lends itself to this method of data collection though validation of the results obtained has been necessary.

Some patients may not fall into any of the above categories. Though a routine out-patient visit after the stay has been suggested as a requirement of care it was agreed that if the only reason for the visit was to collect the data this was not acceptable. Thus an estimated 15% of patients were due for follow-up but not contacted. The whereabouts of this group are being examined for a complete 12 months follow-up.

Finally there was a problem in identifying when patients had died, as others have noted (Walters 1990). The routine was to check carefully through hospital computer systems and then all locally notified deaths in the district (most of the patients will be residents). In some circumstances local GPs were contacted.

Validation of Barthel Scoring

Whilst the simplicity of the Barthel scale is one its great advantages, there is the potential problem that the scoring depends on following some simple rules, and that some interpretation of dependence and independence have to be undertaken by the rater. There is a danger that the instrument might be unduly influenced by whoever completed it. Therefore two checks were made on the scoring of the Barthel chart when the same patient was scored by different raters or under different conditions.

The process of scoring the items of the Barthel via telephone through discussions with carers in another institution was checked for 20 patients by visiting the carers and patients shortly afterwards. In addition the assessments made during the social round were compared to those made by the primary care nurse (who knew the patient but may not have been part of the meeting) or by a nursing auxiliary on the ward.

The results of the pairs of evaluations were compared. Though not all the results were consistent they did show a generally high level of association. Table 4.1 shows the association between individual items and the total score using Kendall's Tau_b and the number of patients where the scores were exactly the same. Most items on the scale show highly significant associations between the two scores. The exceptions appear to be 'Grooming' and 'Transfer' where the ratings appear less reliable.

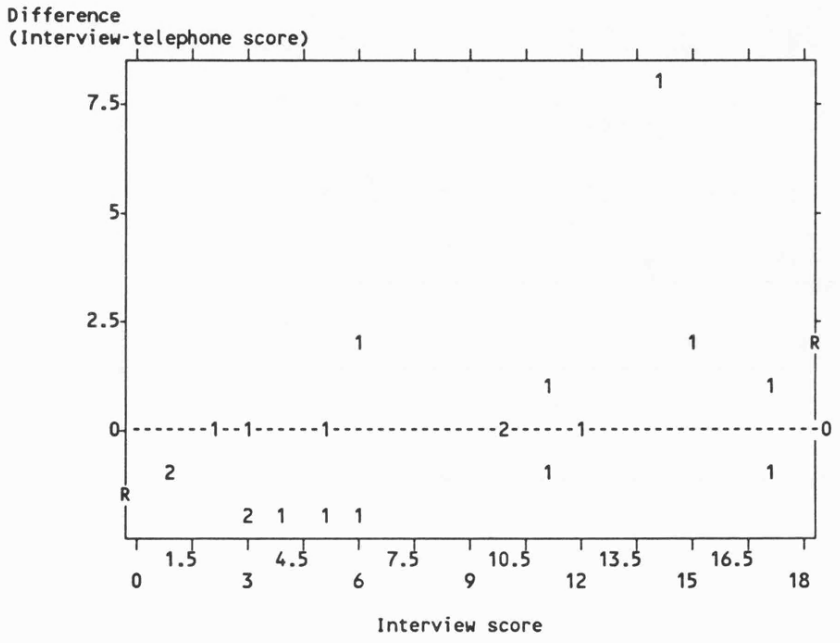
For the total score it is clear that the two ratings did not produce exact matches despite

Table 4.1 Associations between repeat administration of Barthel scale comparing telephone with direct measurement and social round with primary care nurse assessment.

	Telephone Check n=20			Ward Check n=11		
	#Same	Tau	Sig	#Same	Tau	Sig
Bowels	14	.667	.0006	10	.720	.004
Bladder	17	.909	.0000	10	.554	.019
Grooming	16	.404	.039	9	.32	ns
Toilet	16	.800	.0001	7	.57	.014
Feeding	16	.806	.0001	7	.46	.04
Dressing	16	.727	.0003	9	.69	.004
Mobility	16	.700	.0003	11	1.00	.000
Transfer	13	.571	.0023	8	.67	.005
Bathing	20	1.000	.0000	8	.72	.002
Stairs	-	-	-	9	.41	ns
Total	6	.817	.0000	2	.659	.001

the high degree of association. The results were examined using the method of Bland & Altman (Bland & Altman 1986) of comparing continuous variables, which in this case is a necessary assumption. The difference between the two scores was examined and plotted against the original assessment - Fig 4.1. It is clear that in one patient there is a major discrepancy between the two scores, this may be related to a change that occurred between the original scoring and subsequent validation, but in general the agreement was reasonably good.

Fig 4.1 Plot of difference between interview and telephone scores versus interview score (n=20).



There was no difference in the mean scores with the two methods so no apparent bias. The standard error of the difference suggest that most observations fall within the 95% confidence intervals which roughly equates to plus or minus 1 point on the full scale. This was not considered to be an important difference from a clinical perspective. It is suggested that this could be used as a 'rule of thumb' when using the Barthel score. There does not appear to be any relationship between the differences between the ratings and the absolute level of the score.

It became clear during this exercise that the potential difficulties with this instrument

were:

1. A knowledge of the guidelines for completion of the Barthel was important. The more staff that were involved in collecting data the more problems were likely to arise in this respect. In particular the measure requires distinguishing between what the patient actually does rather than what the patients is felt capable of doing.
2. Very few patients could get in and out of the bath without assistance, or if they could staff did not let them for safety reasons. Thus most cases scored zero on this item.
3. In some cases the score could change quite rapidly (within a day or two) as the patient recovered from an acute illness. In some cases it may be that this change does not really reflect improvements due to the care given. Hence the use of the expected goals for the patients used in this study, or the proposed measurement of 'pre-morbid' states suggested by others (Royal College of Physicians Working Group on Functional Assessment 1990).
4. The item on 'Stairs' created problems for completion on some patients when there was really no need for the patient to go up and down stairs either in the hospital or in their normal home environment.

B. Data analysis

The section on data analysis is concerned with three key questions:

1. Do the indicators show significant changes in hospital or during the period to follow-up?.
2. Are the indicators related to each other?
3. Are the indicators influenced by the presenting characteristics of the patients?

Data on the care giver strain scale has been analysed separately and is briefly summarised at the end of this section.

As yet only 11 patients who have not been readmitted and have survived are due for a 12 months follow-up. The detailed results of the 12 month follow-up have therefore not been presented as the numbers are too small. However it is known that some patients

have died after the three month follow-up and this fact has been used in estimating survival rates.

Changes in indicators

(i) Deaths and readmissions

It is inevitable that a substantial proportion of these elderly patients will die in hospital - even when admissions for terminal care are excluded. Table 4.2 summarises the numbers of cases studied for the sample who have reached discharge and those who have reached the time of their three month follow-up. For the sample followed to 3 month follow-up 53 of the patients died during the hospital stay and a further 39 cases up to follow-up. Fig 4.2 shows the survival curve in successive months - using cases where the date of death was known.

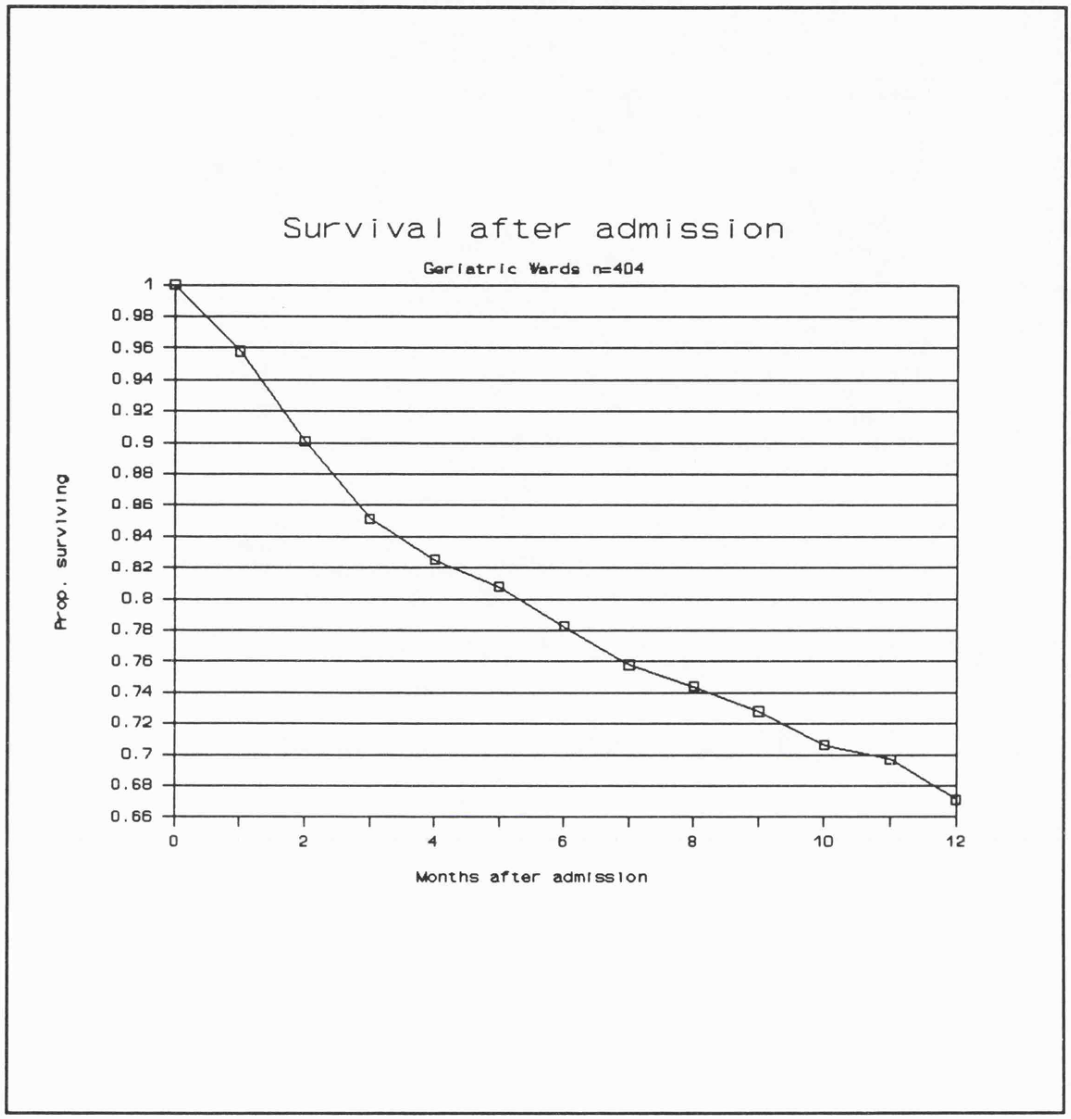
Table 4.2 Numbers of patients and response rates

Total rehab patients admitted to 1-12-90	= 404	
died in hospital	= 65	(16.1%)
discharged	= 339	(83.9%)
Total patients admitted to 1-9-90	= 292	
died in hospital	= 53	(18.2%)
died to follow-up	= 39	(13.4%)
readmitted to 3m	= 15	(5.1%)
ok at follow-up	= 139	(47.6%)
not contacted	= 46	(15.8%)

There were a number of patients discharged who did not receive a follow-up as discussed earlier. It is assumed that these patients are still alive and have not been re-admitted to another hospital. Fig 4.2 suggests that the rate at which patients died was fairly constant over the three months following admission (average length of stay is of the order of one month). These numbers probably slightly underestimate the numbers that have died.

In addition 15 patients (5.1%) had been readmitted to the Freeman Hospital. The procedure in these cases was to consider this event as an endpoint of the first admission. Data collection was then started to assess the outcomes of the second inpatient episode.

Fig 4.2 Estimated survival following admission to geriatric ward.



(ii) Functional status

Table 4.3 summarises the changes in Barthel scores between admission and discharge. There was a significant improvement in aggregate scores ($p < .001$ using Wilcoxon ranks sum). The mean scores increased from 11.7 on admission to 14.5 at discharge. Table 4.3 also clearly shows that in terms of the total Barthel score few patients who survived had a lower score at discharge and even fewer had scores greater than 2 points lower. The distributions of scores at both admission and discharge showed a high proportion of

scores over 17 - indicating relatively less disability. As a result of this there were clearly some end effects with this scale potential for improving scores was considerably reduced if the initial scores are high.

These improvements in the total Barthel score were matched by highly significant improvements on each of the individual items as shown in Table 4.4. which summarises the changes in the individual items of the Barthel scale between admission and discharge. The proportion of cases recording a particular score are shown together with the means scores for that item and the proportions getting 'better' the 'same' or worse'. The higher the score the greater the independence of the patient.

There were certain problems of specificity with this scale as can be seen fairly clearly from this table in that the most common scores were those indicating the greatest independence, that is the highest score. Typically most patients had the same score on any one item at admission and discharge. Those items with a scale of four categories benefit in this respect in that they allow a finer gradation of independence and consequently cases were more likely to show change. Despite these limitations on any one item there was an observed 15-30% of patients who showed an improvement. When all items were added together 60% of cases showed some improvement which suggests that the total scale performed better than any one part.

Interestingly the change in ability to cope with stairs showed little overall improvement - problems were found in scoring this item as either stairs were not available for testing or the task was deemed to be irrelevant. The other items that showed less improvement were 'Eating' and 'Bathing'. The aggregate score appeared to be more sensitive than the individual items. Given the breadth of coverage of the scale, it is perhaps not surprising that few patients suffered from problems on all the items. The pattern, as expected, was that most patients had only one or two items on the scale that were amenable to change.

Table 4.3 Changes in Barthel scores - admission to discharge (n=314)

Admission: Mean = 11.74 Std Dev = 5.58
Discharge: Mean = 14.45 Std Dev = 5.01
Sig Difference p<.0000 (Wilcoxon Signed ranks tests)

Crosstabulation:

		Distribution of scores on discharge										
	Count Tot Pct	< 2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	Row Total
Scores on Admission	< 2	2 .6	3 1.0	4 1.3	1 .3		2 .6	2 .6	1 .3	1 .3	1 .3	17 5.4
	3-4	1 .3	5 1.6	2 .6	3 1.0	1 .3	4 1.3	2 .6		2 .6	1 .3	21 6.7
	5-6	1 .3	3 1.0	1 .3	3 1.0	8 2.5	1 .3	4 1.3	3 1.0	4 1.3	1 .3	29 9.2
	7-8	1 .3		2 .6	5 1.6	10 3.2	5 1.6	1 .3	4 1.3	7 2.2	2 .6	37 11.8
	9-10		1 .3		1 .3	2 .6	7 2.2	6 1.9	5 1.6	4 1.3	5 1.6	31 9.9
	11-12				1 .3	2 .6	5 1.6	5 1.6	9 2.9	9 2.9	5 1.6	36 11.5
	13-14							6 1.9	6 1.9	8 2.5	6 1.9	26 8.3
	15-16		1 .3			1 .3		1 .3	5 1.6	13 4.1	5 1.6	26 8.3
	17-18					1 .3		1 .3		25 8.0	12 3.8	39 12.4
	19-20									1 .3	51 16.2	52 16.6
	Total	1.6	4.1	2.9	4.5	8.0	7.6	8.9	10.5	23.6	28.3	100.0

Table 4.4 Changes in the scoring of items of the Barthel admission to discharge
Percentage of cases with particular scores, means at admission and discharge and the percentage who are 'better'(B), the 'same'(S), or 'worse'(W).

% cases n=339	Admission Score				Discharge Score				Changes in score				
Item	0	1	2	3	0	1	2	3	Mean Admit	Mean Disch	B	S	W
Bowels	20.1	8.6	71.4	-	8.0	6.8	85.3	-	1.51	1.77	19	77	4
Bladder	29.8	10.9	59.3	-	18.3	10.0	71.7	-	1.30	1.53	22	73	6
Grooming	57.5	41.9	0.6	-	31.0	67.0	0.6	-	0.43	0.68	28	68	4
Toilet	25.7	35.7	38.6	-	10.0	26.5	63.1	-	1.13	1.53	34	63	3
Eating	8.0	25.4	66.4	-	2.9	20.4	75.8	-	1.58	1.71	15	82	3
Dressing	23.3	44.8	31.6	-	9.1	33.3	56.6	-	1.09	1.47	36	60	4
Bathing	1.2	92.3	6.5	-	87.6	10.6	0.0	-	0.05	0.11	5	94	1
Mobility	31.6	19.5	5.0	44.0	11.2	14.2	8.0	66.4	1.76	2.29	35	62	3
Transfer	12.7	20.4	15.9	51.0	1.8	11.2	13.0	73.7	2.05	2.58	34	63	3
Stairs	0.6	46.3	6.8	23.9	.6	34.5	8.6	31.3	1.32	1.47	12	84	4
Combined	-	-	-	-	-	-	-	-	11.74	14.45	63	32	5

This can be seen when the individual goals are examined. Table 4.5 shows for individual items of the Barthel score at admission and discharge, the proportion of cases who were better than, equal to or less than the stated goal. It can be seen that in the majority of cases (60-75%) the goals on any one item were equal to the admission score. The potential to show improvement is only therefore possible in the remaining 25-30%. The results at discharge suggested that there were few patients (excluding deaths) where the desired goals had not been achieved (less than 10%) leaving most cases with a discharge status equal to or above the individual goals.

The best performance against goals were seen for items covering 'Bowels' and 'Bladder'. The worst were for 'Grooming', 'Dressing' and 'Transfer'.

Table 4.5 Individual goal setting by items of Barthel scale.
Percentage of cases where goals are greater than or less than admission or discharge Barthel ratings n=339.

	Goal >admit	Goal =admit	Goal <discharge	Goal =disch	Goal >disch
1 Bowels	24.5	75.5	1.5	87.6	12.9
2 Bladder	27.7	72.0	3.6	80.8	15.7
3 Grooming	30.4	69.6	6.8	80.2	13.0
4 Toilet	39.3	60.2	7.0	81.1	12.1
5 Eating	12.1	87.9	5.3	88.5	6.0
6 Dressing	39.8	59.9	7.4	79.1	13.6
7 Bathing	7.1	92.9	2.1	93.5	4.4
8 Mobility	40.1	58.4	6.8	80.5	12.7
9 Transfer	36.6	63.4	7.4	80.2	12.3
10 Stairs	13.3	85.5	5.6	86.4	7.9

Table 4.6 charts the discharge and follow-up Barthel scores for those patients that

Table 4.6 Changes in Barthel scores from discharge to follow-up (n=138).

Mean at admission = 12.386 Std Dev=5.05
Mean at discharge = 14.73 Std Dev=4.97
Mean at follow-up = 13.500 Std Dev=5.59
Follow-up sig greater than admission p=.0061
Follow-up sig less than discharge p=.0002

Score at discharge	Count / Tot Pct	Distribution of scores at follow-up: <2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	Row Total
<2		2 1.4	2 1.4			1 .7		1 .7		1 .7	1 .7	8 5.8
3-4		1 .7	1 .7	2 1.4	1 .7	1 .7	1 .7				1 .7	8 5.8
5-6				2 1.4	1 .7	1 .7	2 1.4		1 .7			7 5.1
7-8				1 .7	2 1.4	1 .7						4 2.9
9-10						3 2.2	2 1.4	2 1.4	1 .7	2 1.4		10 7.2
11-12			1 .7		2 1.4	3 2.2	2 1.4	1 .7	2 1.4	2 1.4	1 .7	14 10.1
13-14					1 .7	1 .7	1 .7	2 1.4	5 3.6		2 1.4	12 8.7
15-16								2 1.4	2 1.4	6 4.3	8 5.8	18 13.0
17-18					1 .7			2 1.4	2 1.4	12 8.7	12 8.7	29 21.0
19-20							1 .7		2 1.4	7 5.1	18 13.0	28 20.3
Column Total		3 2.2	4 2.9	5 3.6	8 5.8	11 8.0	9 6.5	10 7.2	15 10.9	30 21.7	43 31.2	138 100.0

survived to the three month time period. The general pattern appeared to be of little change after discharge - the scores at follow-up remained significantly higher than at admission for paired data but there was no significant change from discharge. It appeared that the beneficial effects of the hospital spell in terms of improving Barthel scores were limited to the in-patient stay with patients remaining fairly stable after that. It is interesting to observe that there were a handful of quite severely disabled patients on discharge who had subsequently improved to 3 months (presumably most will be in some other care setting). Very few show residual scores less than say 5.

(iii) Accommodation status

Table 4.7 summarises the accommodation status of patients at admission and discharge for those patients who survived. The four categories used here were abbreviated from the ten categories that were used on the data collection form and represented a scale, as expressed by the clinicians, of 'desirability' of the home environment, in effect a preference for the patient to be independent.

There was relatively little change between the various categories. The most common category for the survivors was the 64% of cases who were admitted from home and discharged home. All changes were 'for the worse' according to the pre-determined

Table 4.7 Changes in accommodation status - Admission to discharge (n=138)

		Accommodation on discharge				
On Admission	Count / % Total	Home	ResCare	Rehab	LT care	Row Total
Home + (Shelt)		88 63.8	10 7.2	7 5.1	10 7.2	119 86.2
Resid Care			11 8.0		4 2.9	15 10.9
Rehab Hosp				1 .7		1 .7
LT Care					2 1.4	2 1.4
Column Total		88 63.8	21 15.2	8 5.8	16 11.6	138 100.0

preferences expressed by clinical staff and in particular a number of patients admitted from home are inevitably discharged somewhere else. The fact that no cases move up the scale of independence from long term care to home, lends a little validity to the

ranking of these categories. Comparing the accommodation at discharge to three months reveals relatively little change amongst patients once they had left hospital. The biggest changes were for 8 patients who moved between rehabilitation in hospitals and long term care.

(iv) General ratings

At discharge clinical staff were asked to assess whether goals of patients had been achieved, whether the medical problems had been resolved and if the patient was generally better. The results (shown later in Table 4.9) show that for the majority of cases (around 80%) the answers to these cases were positive. More specifically on the assessment of medical goals 74% of patients were scored as 'Better', with only 20% the 'Same' and as few as 2% as 'Worse'.

Almost identical results were seen for the assessment of general status 71% 'Better', 24% 'Same' and 2% as 'Worse'. The questions on whether the overall goals had been achieved showed less confidence in the benefits of care, for 78% of cases the answer being 'Yes', 9.1% 'Unsure' and 9.1% 'No'. This item was originally included as a screening method to pick out non-responding patients. The relationships between these assessments and the other indicators are discussed later as well as the relationship between them.

Relationships between indicators

Barthel score and accommodation

There were strong and statistically significant links between the admitting Barthel scores and the accommodation status of patients. This is as one would expect, patients who have come from other institutions score on the whole lower than those who have been living at home. The mean values of the Barthel scores by admitting accommodation status are shown in Table 4.8(a). The gradient in scores roughly matches the gradient developed by clinical staff when developing a classification of accommodation. One contentious point in these results was the status of sheltered accommodation, the Barthel scores for this group of patient suggest an independence level equivalent to those living

at home (it has been included in this group in the simpler four point scale used earlier). Scores for patients admitted from residential care are notably lower.

Table 4.8(a) Mean Barthel scores on admission by home environment

	Barthel Score Mean	Std Dev	Cases
For Entire Population	11.7420	5.5841	314
	16.0000	0.0	1
1 Home alone	12.4286	5.5233	63
2 Home+social services	13.7358	4.8799	53
3 Home +prof	12.2500	4.3995	20
4 Home + other	11.4706	5.7649	119
5 Sheltered accommodation	12.6250	6.4323	24
6 Residential care LA	8.0000	3.3040	25
7 Residential care Private	6.0000	0.0	1
8 LT care rehab	7.5000	.7071	2
9 LT care	5.0000	4.9666	4
P LT care private	3.0000	2.8284	2

Table 4.8(b) shows the Barthel scores according to accommodation on both admission, discharge and at 3 month follow-up using the simpler 4 categories. The significant differences between Barthel scores according to accommodation status at admission to hospital were also found to exist on discharge from hospital. Thus patients discharged 'home' had significantly higher scores than those discharged to residential care. Considering the net change in Barthel score there are no differences between the groups - all show an improvement of between 1.5 and 3 points on the scale.

Table 4.8(b) Mean Barthel scores at admission, discharge and change according to home environment at admission and discharge.

		Mean Barthel scores by admission environment			Mean Barthel scores by discharge environment		
	No.	Admit	Disch	Change	Admit	Disch	Change
Home	289	12.2	15.1	3.0	13.1	16.4	3.4
Resid	30	8.0	10.1	2.6	9.9	13.6	3.7
Rehab	2	7.5	9.0	1.5	8.7	10.5	1.9
LT Care	5	6.2	8.4	2.2	8.5	7.2	-.61
Significance		p<.001	p<.001	ns	p<.001	p<.001	p<.001

Table 4.8(c) classifies patients into those who have remained at the same level of accommodation status against those who have 'slipped' down the ladder. There are significant differences between admission scores, discharge scores and change in Barthel for these two groups. The improvements in patients who remained at the same level of accommodation was slightly better.

Barthel scores vs survival

The Barthel scores for patients who die during the hospital stay were significantly lower than for the survivors (Table 4.8(c)) the average score for survivors was 11.7 against an average of 4.4 for those that die ($p < .0001$). When the Barthel scores of patients who die after discharge were compared with those that survive there were no significant differences in either admitting score, discharge score or change in score.

The identification of lower Barthel scores in those who die in hospital was as expected and the absence of such an association for those that die after discharge presumably reflects the fact that lowered scores tend to be associated fairly closely with the short times just before the patient dies. It is dangerous to read too much into the observed relationship between low Barthel scores and survival - and in particular to infer some form of causality.

Table 4.8(c) Mean Barthel scores and accommodation changes against other indicators (n=314)

Kruskal-Wallis analysis of variance by ranks (Barthel scores);
Chi-square (accommodation) (ns = p>.05)

	Admission Barthel	Discharge Barthel	Change in Barthel	Accom Same	Accom Worse
Ok to 3m Died in hosp Died to 3m	11.67 8.39 12.33 p=.0000	14.5 - 13.8 ns	3.0 - 2.3 ns	-	-
Accom Same Lower	12.2 9.9 p=.0017	15.4 11.3 p=.0000	3.2 2.0 p=.016	-	-
Gen Better Same Worse	11.6 12.0 11.6 ns	15.0 13.2 7.4 p=.0012	3.7 1.5 -4.2 p=.000	82% 63% 20% p=.001	19% 38% 80%
Medical Better Same Worse	11.9 11.0 11.3 ns	15.1 12.7 9.3 p=.0002	3.5 1.9 -2.0 p=.0001	80% 66% 29% p=.004	20% 34% 71%
Goals Yes Unsure No	12.1 11.0 8.5 p=.0018	15.1 12.9 9.7 p=.000	3.2 2.7 1.6 ns	81% 74% 36% p=.000	20% 26% 65%

Barthel scores and accommodation vs General ratings

The results of the summary questions completed at discharge by clinical staff are shown in Table 4.9. The three questions were:

- Has the medical condition improved?

- Is the patient better than before?
- Have the overall goals been achieved?

Table 4.9 shows the high degree of inter-relationships between the three simple summary questions. In 84% of cases there was direct agreement between the achievement of medical goals and improvements in general health.

When the Barthel scores were compared to the summary made at discharge there was in general agreement between the two assessments (Table 4.8(c)). Thus the discharge Barthel scores and the change in Barthel scores were significantly worse for cases where the medical problems were classified as 'Worse' and where the general status was 'Worse'. Similarly these

Table 4.9 Relationships between summary assessments (n=339)
(All assessment significantly associated. Chi-squared p=.0000)

	Overall improved			Medical Goals		
	Yes	Unsure	No	Better	Same	Worse
General status?						
Better	214	18	8	226	14	0
Same	51	11	17	23	54	2
Worse	0	1	6	0	0	5
Medical Goals?						
Better	223	16	9			
Same	41	13	14			
Worse	0	1	6			

patient tend to have slipped down the ladder of accommodation status which indicates some consistency between the ordinal scale of accommodation and the perceived goals of care.

With respect to the general achievement of goals it is interesting to note that patients where the goals were not met tended to have lower scores on admission as well as at discharge. Presumably this category of patient includes ones who were relatively dependant and though some improvement was achieved the expectations were higher. The reasons for this need to be explored further.

Barthel vs Mental Test Score.

When the admitting mental test scores were examined there was a weak link to overall Barthel scores. Patients with low mental test scores (below 5) showed in lower Barthel

scores, the mean values on the Barthel are shown in Table 4.10. There is little variation in mean Barthel scores between the relatively less confused patients, scoring greater than 5 on the mental test score.

Table 4.10 Mean admitting Barthel scores by scores on abbreviated mental test

		Mean	Std Dev	Cases
For Entire Population		11.7420	5.5841	314
Test score =	0 or not recorded	11.6126	5.6844	111
Test score =	1	7.2727	5.6761	11
	2	9.7143	4.8206	7
	3	11.6250	5.9507	8
	4	9.9167	4.3161	12
	5	12.7000	5.3222	20
	6	12.3158	5.5783	19
	7	12.4783	5.6638	23
	8	11.8276	5.6446	29
	9	11.8750	5.6383	32
	10	12.8571	5.4753	42

Once again this relationships is expected as patients with severe confusion will tend to be more dependent. This can be seen as supporting evidence for the validity of the Barthel scoring.

The effects of presenting characteristics.

There were relatively few presenting characteristics which were thought to be possible effect modifiers. The effects of these variables against the mean Barthel scores from admission to discharge, as well as the relationship to survival and accommodation status are shown in Table 4.11. As in other specialties the analyses presented are basically exploratory bivariate comparisons between outcome indicators and presenting variables designed to assess which show the strongest relationships to observed outcomes. There is undoubtedly some interaction between the presenting variables and more sophisticated analysis could minimise these effects and explain the results more fully.

There appeared to be some links between age and the outcome indicators though the differences were not large. Younger patients (under 70) showed a greater improvement during the stay together with slightly better survival rates and fewer accommodation changes. Presumably this reflects a proportion of these patients where the acute illness is more significant than chronic longer term problems.

Table 4.11 Barthel scores and accommodation status by presenting characteristics (n=404)
(ns = p>.05)

Characteristic (% cases)	Admit. Barthel	Disch. Barthel	Change Barthel	Survive	Died in Hospital	Died to 3m	Accom Same>worse
Marital Married(31%)	11.4	13.9	2.8	66%	28%	5%	89% 10%
Status Single (12%)	11.5	15.5	4.1	88%	6%	6%	80% 21%
Widow (57%)	12.1	15.0	3.1	69%	23%	8%	73% 27%
	ns	ns	ns	p=.006			ns
Consult =A (58%)	12.6	13.6	2.6	73%	21%	5%	80% 20%
=B (42%)	10.4	15.0	3.5	72%	20%	9%	72% 28%
	p=.004	p=.012	p=.036				ns
Sex Male (48%)	11.5	13.6	2.5	68%	27%	5%	80% 20%
Female (52%)	11.9	14.9	3.2	75%	17%	8%	75% 25%
	ns	p=.022	ns	p=.04			ns
Age (Mean years)	-	-	-	81.4	81.4	85.0	81.1 82.9
				p=.0253			p=.037
Age Band <70 (5%)	9.1	14.8	5.7	82%	14%	5%	75% 25%
71-80 (38%)	11.4	14.8	2.9	77%	22%	1%	86% 15%
81-90 (47%)	12.7	15.6	2.5	70%	20%	11%	71% 29%
>90 (9%)	11.0	15.1	3.1	69%	20%	11%	66% 35%
	p=.023	ns	p=.01	p=.011			p=.019
Complics =1 (19%)	11.8	14.5	2.9	81%	17%	3%	73% 27%
=2 (28%)	12.5	15.2	3.1	80%	13%	7%	79% 22%
=3 (29%)	11.8	14.4	2.8	75%	20%	5%	74% 26%
=4 (18%)	10.6	14.3	3.8	60%	28%	7%	81% 19%
=5 (6%)	12.3	13.0	1.2	59%	33%	1%	90% 10%
	ns	ns	ns	p=.013			ns

Differences between the sexes mean Barthel scores can be seen on discharge, where women score slightly higher than men (p=.022) and in survival where a higher proportion of women survive hospital (p=.04) these effects were not strong. Marital status appeared not to affect the Barthel scores or accommodation but was related to survival. Single patients appeared less likely to die in hospital or at three months afterwards. Though it is interesting to speculate on the reasons why this is so, it does not seem likely to be causally related to the processes of care.

There appeared to be some differences in the admitting patterns of the consultants. Consultant B admits patients with on average lower Barthel scores yet discharges them with higher ones. The survival rates and changes in accommodation status are similar for the two consultants.

Finally the results on the index of complications, based on the number of different organ systems with active disease, did not appear to be related to mean Barthel scores or

accommodation status, though there was a link with survival where patients with more complex disease were more likely to die (p=.013). It is recognised that this initial method of classifying complications is crude and that a more sophisticated approach may reveal stronger relationships. Alternatively it may be argued that the link between functional ability (as measured by the Barthel) and diagnostic classification will always be weak and confounded by the natural variability in the severity of any one disease as well as the mix of diseases present in any one patient.

Care Giver Strain Scale

The care giver strain scale was only applied to a subset of patients, and the results to date have been rather limited. The analysis of this particular indicator is therefore considered separately. The basic distribution of the strain scale is shown in Fig 4.3 and clearly demonstrates that despite problems in selection of appropriate patients and carers the scores were fairly evenly distributed around the mean of 5.9 (Std Dev=3.3). Relatively few carers score either zero or over 10 - that is at the extremes of the scale.

When the individual items were examined the frequency with which any one item ticked range from 22% "It is a financial strain" to 75% "Some behaviour is upsetting".
Apart from these two extremes most items were ticked in around 50% of carers answering 'yes'. The scale was weakly related to the admitting Barthel score in the type of relationship one might expect: the more independent the patient the lower the score on the strain scale. No strong relationships were observed between the strain scale and changes in Barthel, accommodation or mental test score.

Fig 4.3 Distribution of total scores on care giver strain scale (n=131).

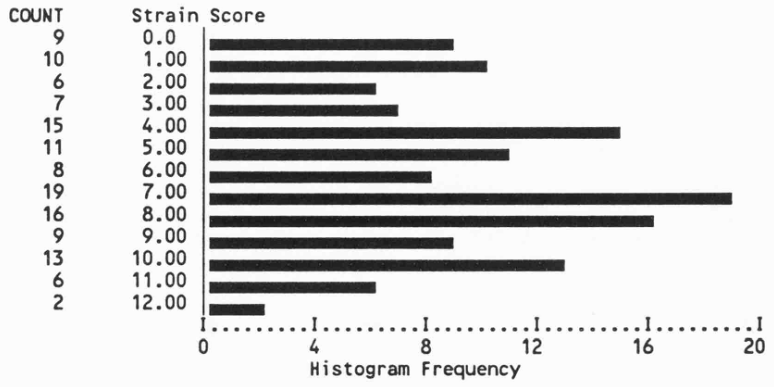
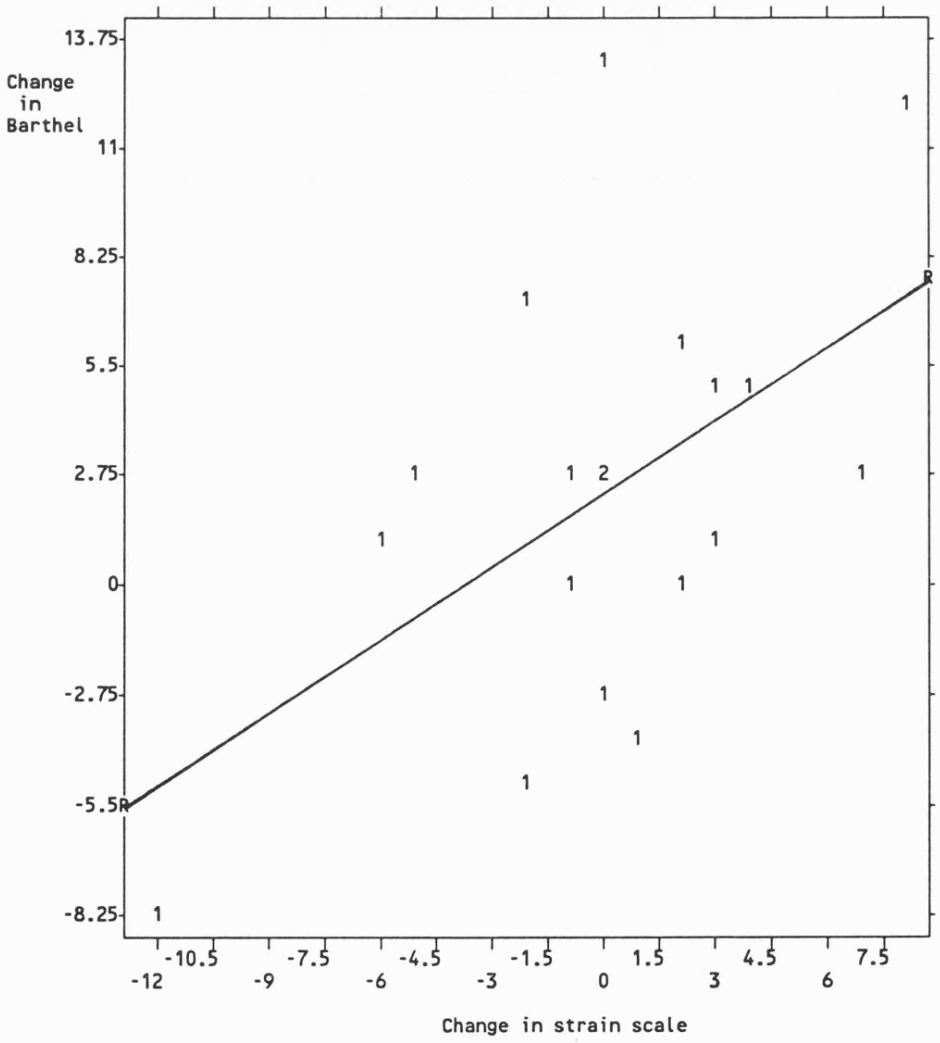


Table 4.12 Responses to a sample of care giver strain scale n=73

	%'Yes'
1. Sleep is disturbed	52.1
2. It is inconvenient	47.9
3. It is a physical	63.0
4. It is restricting	65.8
5. There have been family disruptions	47.9
6. There have been changes in personal plans	46.6
7. There have been other demands	45.2
8. There have been emotional	43.8
9. Some behaviour is upsetting	75.3
10. It is upsetting that	54.8
11. There have been work adjustments	28.3
12. It is a financial strain	21.9
13. I feel completely overwhelmed	54.8

In cross section it would appear that the items selected in the scale are appropriate for considering the problems of carers in this type of population.

Fig 4.4 Changes in Barthel score versus changes in strain scale from admission to 3 months (n=19)



Only a handful of cases were available to look at the changes in the scale that may have occurred between admission and the three month follow-up. The mean score on admission for the twenty cases was 5.25 (Std Dev=3.6) and at follow-up 5.35 (Std Dev=4.3). Though the mean showed no change, individual scores have changed, by as much as eight points in some cases.

Fig 4.4 shows that there was some correlation between changes in stress and changes in dependency for the few cases which have follow-up results (Correlation coefficient=.59 $p<.01$). This is a promising result when considering the validity of the stress scale and changes in the scale.

Conclusions on Data Analysis

The analysis of the data has shown that the indicators used to study outcomes in care of the elderly can rely heavily on two issues - survival and functional ability as measured by the Barthel scale. The other indicators, that is monitoring changes in accommodation status, mental test score and strain on the main care giver have shown little longitudinal change though they may well be valuable as information which describes, in cross-section, the patient population.

It is clear that significant improvements can be seen in the Barthel scores of patients from admission to discharge. The scale suffers from the fact that it is insensitive to changes in the more independent patients (approx 30% of the admitting population to this hospital). For the study of care of the elderly in other settings eg the day hospital, this is liable to be a more serious problem. One reason the scale was originally chosen was because it was felt that the individual items represented the basic problems that patients presented and which were likely to change as a result of the hospital stay. At the same time it was recognised that not all patients would suffer from all the possible problems. It would have been possible to consider improvements in terms of the specific items rather than aggregated score though the analysis would have been correspondingly more complex. On the other hand, by using the whole scale especially when only one or two items may be relevant, requires the assumption that the relative weighting of the items

were representative of the relative importance of the particular problems that make up the scale. Thus a one step improvement in bladder control is considered equivalent to a one step improvement in ability to dress oneself. Compared to other measures of general health status this weighting is certainly crude but at a simple level the total scale appears to work reasonably well.

The Barthel scale behaves largely as predicted in that it is related to survival, accommodation status and mental test score. Thus patients who die in hospital tend to have lower scores on admission, those admitted from home tend to have higher scores than those admitted from residential care, and those with severe confusion tend to have lower scores than those without. Similarly the Barthel scores, and changes in these scores broadly agree with the summary judgements made by the clinical staff at discharge.

There are advantages in using the combined Barthel score which appears to work better in showing improvement during the hospital stay than any one of its individual components. Very often patients will score the maximum on a number of individual components and the potential for improvements in scores are limited to the few remaining items. The combined score on the other hand aggregates across dimensions to give potentially a finer gradation for the assessment of improvement or deterioration in function. The observed changes in patient's function are in most cases limited to the course of the hospital stay. Changes after discharge (other than deaths) are much more limited though the improvements observed in hospital tend to be maintained and there is no deterioration. The assessment of patients' home environment does show the types of changes expected as patients become more dependent, though these changes are largely mirrored by information in the Barthel score. The changes in accommodation status are greatest between admission to hospital and discharge with relatively little change after discharge and to the three month follow-up. Using an ordinal scale of different forms of accommodation with some preferred to others, a few patients fall down the scale from admission to discharge. This fact in itself is not necessarily bad as long as discharge destinations are deemed appropriate. It is more useful in a comparative sense to ensure that either over time or against other institutions, the proportion of cases

unable to return home is not very much lower than expected. Perhaps this information is best used when specifically linking levels of disability with particular forms of accommodation.

The indicators have not been found to be particularly sensitive to some of the basic presenting characteristics of the patients. The age of the patient appears to be the most important with the younger age group (under 70) faring better in terms of function and survival. The fact that no strong relationships are seen between the simple index of active disease and the various indicators may be a reflection of the rather simple way in which active disease are recorded and in this case summarised. It would be useful to have a more sophisticated way of dealing with information on the medical problems of the patients yet when the combinations and interactions of diagnoses can be so complex this will prove difficult.

D. Review Process

To date the feed-back of the results in geriatrics has been limited. One problem has been to identify an appropriate forum for discussing methods and results. Amongst the various alternatives have been individual meetings with clinical staff, and larger multi-disciplinary audit groups.

The project involved a variety of clinical staff in some role and they have all offered their support and have undertaken much of the basic data collection and organisation. In fact, at the request of the staff the project has now been extended to the day hospital where considerable efforts have already been made in identifying what data should be collected and how to achieve that (pilot questionnaires are shown in Appendix 7). There has been a commitment to the project from staff at a number of levels and this has been very encouraging.

Data collection and analysis has not been without problems and changes have had to be

made during the course of the project to improve these. It is difficult to know how much of this process of data collection, which has changed and evolved over time, would be applicable in other hospitals. The actual process of the study itself has had a number of effects on the practice of information collection within the department. In particular:

1. The identification of the main care givers is now done explicitly by nursing staff and recorded centrally
2. The identification of where the patient goes on discharge is now collected centrally whereas before the information could be in a variety of different locations or not known at all.
3. At one stage, the recording during the social round of the Barthel score and accommodation status and the goals for that patient was felt to have improved and helped focus discussions. However, changing personalities have meant that completion of these forms has been carried out in a variety of different ways and that this improvement may not have been sustained.
4. Nurses have become more involved in assessing patients and in particular have undertaken, where relevant, the completion of mental test scores on discharge. Previously MTS scores were not noted on discharge.

In addition to these changes in information processing there are some other examples of where the study has had some, albeit limited, impact. Though there were problems with the administration of the care giver strain scale and it has yet to show benefits as a longitudinal measure, there has been an increasing interest especially amongst nurses in the responses received. In fact the forms appear to be used as a screening tool to help identify carers under particularly high levels of stress and to draw this to the attention of the ward nurses. In some cases the severity of the responses has surprised staff - the proportion of carers who have ticked 'I feel completely overwhelmed' has caused comment. Since the start of the project, there has been a move initiated by the manager of services for the elderly (covering the Freeman Hospital and, as was, two satellite hospitals) in collaboration with clinical staff, to set up a support group for relatives and carers.

Discussions with medical staff have to date been inconclusive. There has been a growing confidence in the validity of the results and the initially rather defensive reactions have been largely overcome. There now appears to be considerable potential in using the data rather more aggressively, in particular to monitor the inter-relationships of functional status, accommodation environment and length of stay in the hospital. The observed lengths of stay for some patients was slightly longer than expected and the expected differences in lengths of stay between consultants has not emerged. However suspected differences between the two consultants in the admission patterns has been confirmed. In many ways the data has prompted a variety of questions about how the service operates, and whether this can be improved.

For example the relationship between average dependency on the ward, nurse staffing levels and the current methods for assessing nursing workload are being explored. As yet there are not clear links between the outcomes achieved and particular process of care that can be changed to improve results. As one consultant put it in response to our questionnaire (Appendix 3):

> *"It has focused our attention on what we are doing and will help us to think about it more and audit our work more often and more carefully".*

The interest in continuing outcome monitoring is felt to provide a good opportunity for ensuring that changes in service delivery enhance (or at least do not detract) from patient care.

Chapter 5 Cholecystectomy

A.Introduction

The choice of cholecystectomy patients as an appropriate group for the outcomes study was made at its outset. It was chosen as a high volume surgical case type which involved relatively standard operating room procedures and techniques, and was thought to yield a fairly homogeneous set of outcomes. It was thought to provide a typical example of surgical workload in an acute hospital that would be suitable as a test bed for producing outcome measures. There were four general surgeons who perform the operation at the Freeman Hospital (though one relatively infrequently) and all were involved in the project and analysis of results.

Cholecystectomy (removal of the gall bladder) is one of the most common surgical procedures performed. The procedure is recommended when gall stones block or partially obstruct either the cystic duct or common bile duct (choledocholithiasis) producing symptoms such as biliary colic, jaundice etc.

The severity and acuity of symptoms can vary, in some cases the patient will present with very severe biliary pain as they are undergoing an acute attack of cholecystitis. With acute attacks surgery may be performed as urgent or emergency though where possible medical management to control pain is preferred until the patient is stabilised. Early surgery is preferred to a delay of a few weeks or even months. In some cases the patient will experience a series of such attacks . In chronic cholecystitis the symptoms may be less acute though rumble on for some time. When the patient experiences only relatively mild symptoms (eg belching, fatty food intolerance) the use of surgery is controversial. In some cases gallstones may be asymptomatic or silent. Under these circumstances it is unclear whether it is best to operate or not, the current wisdom falling down just in favour of waiting until symptoms appear (For a review, see Soloman et al 1986).

There have been links demonstrated between the presence of gall stones and the emergence of subsequent cancer of the gall bladder (Diehl et al 1989) though the risks of this are balanced by the incidence of surgical complications (Finlayson 1989; Ransohoff et al 1983) and again the use of surgery is equivocal.

In the Freeman Hospital, which does not routinely admit emergencies, most of the approximately 150 operations a year are elective. The operation itself is intermediate in scale, taking about 30-60 minutes in theatre, and with a fairly long length of stay post-operatively (1-2 weeks).

The rates with which the procedure is performed (relative to a base population) have been shown to vary by geographic regions (Opit & Greenhill 1974; Roos & Roos 1981; Cageorge,Roos & Danziger 1981) and by country (Vayda 1973). One analysis of the literature (Soloman et al 1986) summarised the possible explanations for this variation as follows:

- imprecise understanding of the disease itself and the differences between physicians in interpreting and recognising the signs of gall bladder disease;
- the use of different surgical indications, when making the decision to operate. For example one study found consultations before surgery reduced the rates.(Bunker 1970);
- the mix of patients and possibility that there may be greater gall bladder disease or 'need' in some areas;
- consumer characteristics, for example education, pain tolerance - though findings on the effects of these issues are inconsistent;
- access to care;
- method of payment and organisation of care, for example in the US fee for service tends to increase usage compared to prepaid group practice (LoGorfo 1979);
- availability of resources (mainly hospital beds) and skilled physicians, and

- the characteristics of the physicians.

In view of this variability it is not surprising that there are a number of studies looking at the criteria for appropriate cholecystectomy (Soloman et al 1986; Scott & Black 1991).

Given the right indications the procedure is felt to be effective. Potential problems of surgery include an operative mortality rate of 1-2% (Mitchell & Morris 1982) and a variety of post-operative complications. Of the alternatives to the procedure, the use of chemical therapy to dissolve stones has been the subject of a number of trials, though it has proved not to be effective (Bateson 1984). More recently the technique of endoscopic removal of gall stones (ERCP) has been used, typically on patients for whom open surgery is not advised. The latest alternative to surgery, using lithotripsy to smash gall stones is currently being evaluated (Milner et al 1989, Chiverton et al 1990).

B. Data set and data collection

Identifying the Outcome Indicators.

Initial discussion with the general surgeons, drawing on the relevant literature, sought to identify the expected changes in patient health before treatment and following the procedure. This formed the basis for the identification of the outcome indicators and the recognition, where possible, of the standards to be used for comparison. The discussions focused on the changes in health that were expected, the appropriate time periods at which to take measurements, possible confounding variables, data definitions and methods of collection as discussed in Chapter 3. The starting data set was developed through a series of discussions with one or two of the surgeons over a period of about two months. Some minor changes have been made during the course of data collection.

The outcome indicators are summarised in Table 5.1 which gives a brief summary of how data was obtained and the standards that were agreed at the start of the project.

Table 5.1 Summary of Key Outcome Indicators for cholecystectomy

Indicator	Information collected by	Standard/Goal
Successful treatment	Pro-forma completed by surgeon	Minimal perioperative complications
Treatment complication	Pro-forma on discharge - key list including MI, infections,DVT etc.	Minimal complications
Symptom/problem relief	Patient completed checklist of symptoms on admission and postal follow-up	No residual symptoms or Net improvement
General Well-being	a. NHP form on admission and postal follow-up. b. Summary health status question	Improvement on baseline
Adverse events after discharge	a. Readmissions (from patient and/or hospital computer systems) b. General questions on improvement	No related readmissions No additional problems at follow-up
Deaths	Pre-follow-up checks on hospital PAS and DHA lists.	Minimal

Outcome Indicators

1. Mortality

Though small it was agreed that operative and post operative mortality would be observed. Typically the literature refers to deaths in hospital (Mitchell and Morris 1982, Solomon et al 1986), where rates of 1-2% are expected, or alternatively to studies of surgical mortality considered up to 30 days post-operatively, with the choice of time period here having a significant effect on the results (Bradbeer 1989). In this study it was thought necessary to follow-up at fixed points after discharge (agreed to be 3 and 12 months) though the actual date of death could be recorded if needed for comparative purposes. Deaths may not be related to the procedure or the disease, therefore the details of such cases were reviewed individually with additional details taken from patient case records.

2. Relief of symptoms

There are a number of typical symptoms of cholecystectomy which cause varying degrees of discomfort and distress to patients. Following the procedure, it was assumed that any pre-operative problems of the patients should have 'cleared up' or at least improved. Not all patients would have every symptom so baseline data on presenting symptoms would

be required. The symptoms used were pain, abdominal distension, flatulence, vomiting, bowel problems, intolerance of fatty foods and loss of appetite.

The most important symptom was agreed to be pain. In order to provide some clear guidance on the severity of pain it was agreed that a working definition would be 'abdominal pain for which the patient takes some form of analgesia'. The questionnaire was worded accordingly.

During the course of the study additional information was collected on the nature of any bowel problems. The original non-specific question was retained but two extra questions were added to determine whether these problems were due to either diarrhoea or constipation.

3. Operative and post-operative complications.

It was recognised that some patients presented problems while actually in theatre either as a surgical or anaesthetic complication. This can be regarded in one sense as a very short term outcome, and in another as a potential predictor of later problems (Pettigrew, Burns and Carter 1987). Additionally there are a number of complications that the patient may experience following surgery. An agreed list of such post-operative complications was drawn up and included:

- Wound infection
- Deep Vein Thrombosis/Pulmonary Embolism
- Urinary Retention
- Cardiovascular complication
- Intra-abdominal infection
- Post-operative bleeding
- Respiratory infection
- Central nervous system complication
- Septicaemia
- Renal failure
- Other (to be specified)

4. Major health events following discharge

It was agreed that following discharge there should be few major health problems - in particular problems serious enough to warrant readmission or visits to accident and

emergency departments or to a GP should not occur as a result of this procedure. The presence of these events was therefore used as a proxy for a deleterious change in patient health status (Chambers & Clarke 1989).

5. Improvement in general well-being

In addition to the obvious clinical changes one of the expected outcomes of the procedure was that the patient would feel better in themselves and resume normal activities, undisturbed by their earlier health problems - "joie de vivre" as it was described by one surgeon. It was agreed that this outcome should be captured by using a general health status instrument such as a patient-completed questionnaire. Chapter 2 discusses the issues around choosing such an instrument, the selected one being, in this case, the Nottingham Health Profile (NHP). Once again due to the degree of variability expected between presenting patients it was agreed that an improvement in NHP score would be a better indicator of success than a single post-operative score. This therefore required the pre-operative assessment of health status to act as a baseline. Additional to the NHP was a single summary question (with 5 point reply scale) and two validating statements "I have health problems I did not have before" and "I feel no better than I did before" seeking patients' agreement or disagreement.

Timing of observations.

The timing of when outcome measures are taken is critical to their interpretation. It was agreed that pre-operative baseline data were needed as well as details around the time of the operation and at discharge. There was much discussion about the timing for follow up measurements post-discharge. Routine clinical practice is for patients to come to an outpatient appointment six weeks after the operation Using this as the outcomes follow-up was considered (with a view to less expensive data collection), but it was considered that patients at this stage still suffered from a variety of transitory complaints which would confuse the picture of their overall health. Therefore a follow-up at three months was chosen as a point at which it would be safer to make judgements about their long term health status. In addition a second follow-up at 12 months was used to check that any good results at three months were maintained and not simply a short term artefact.

Other data, and process and confounding variables.

In addition to the data required to actually measure outcomes it was recognised that information on potential confounding variables would need to be collected in order either to explain particular results or at least to check that variations in outcomes had not been caused by variation in another factor for example age. The identification of which factors to include has to be pragmatic and experience suggests there is a tendency to include too many. Additionally some basic items of data concerning process need to be included since there is genuine interest in whether any relationship can be observed between certain process measures and outcomes obtained.

Further additional *administrative information* was required including the patient's address and telephone number if possible. These were used for post-discharge follow-ups and validation checks.

Information on the *presenting characteristics* of patients included basic details of age (or date of birth) and sex and were taken from the standard hospital identification label used on all request forms. The hospital number was used as the key identifier. The most important clinical descriptors of the presenting patient were agreed to be:

- Indications for surgery - classified into one of eight mutually exclusive categories ranging from acute to asymptomatic. This classification has caused some problems as discussed later.
- Elective/urgent/emergency admission
- Frequency of abdominal pain
- Time since onset of symptoms
- Co-morbid disease by organ system, typically based on active co-morbid conditions
- Medication. This was collected in the early stages of the project as a check on the active co-morbidities. Though it fulfilled this role reasonably well it did prove time consuming to collect and was therefore dropped in later stages of the project - a more precise definition of an active problem was used instead.

- Occupation. Again this was collected in the early stages in order to determine social class. However this proved difficult to collect reliably and was dropped. It was agreed that any investigation of class effects could use post-code data and a socioeconomic grouping such as ACORN.
- Obesity
- Smoker/Non Smoker
- Assessment of surgical risk

Process measures of the operating surgeon/consultant such as pre and post-operative length of stay were also examined. The analysis section later in this chapter considers the relationships between the most important of these variables and the outcome indicators.

One 'composite' patient characteristic was introduced as a result of problems found in the recording of the indications for surgery. The variable called 'acuity' was a simple ordinal scale intended to differentiate between patients using data from the indication for surgery, whether surgery was elective or emergency, the time since onset of symptoms and the frequency of reported pain. The variable was defined as:-

Acuity: = 1 Emergency/urgent surgery, acute pancreatitis, cholangitis
= 2 Acute or history of acute cholecystitis with more than four periods of pain reported in past year.
= 3 Biliary pain less than 4 times in past year
= 4 Asymptomatic (No biliary pain).

Data Collection

Data collection was based on a series of questionnaires completed either by the patient, the surgeon, the consultant, or a research nurse. The study has always sought to encourage data collection using existing processes, resources and ideally existing data in an attempt to make the collection of outcome information as easy as possible. In general surgery, the team also had the services of an established part-time research nurse who participated in this project and who proved invaluable in the co-ordination of basic data collection.

1. Patient identification

The process of identifying appropriate patients for the study was carried out on the wards by talking with one of the two key ward sisters and checking with consultants' secretaries if any eligible patients had been admitted. Cholecystectomy patients at the Freeman Hospital tended to be concentrated on two wards on the same floor, which contributed to the success of this approach. The system by and large worked well though some problems in patient capture were experienced if communications broke down temporarily. project staff covered for absence, holidays etc and checked the relevant wards every week. There were only a few cases where the procedure was performed before the research team had managed to contact the patient. Two patients were subsequently excluded from the study due to the presence of cancer being known in advance in one case, and in the other, because of a complex operation involving a variety of other procedures confounding any results.

With potentially four independent consultant surgeons to consider, making sure all patients were recruited was a problem. Identification of patients was labour intensive and sometimes rather late. However other approaches were determined to be impractical (and experience in the validation studies has verified this). More specifically, identifying patients as they are taken from the waiting list was wasteful in that many patients never arrived (for one reason or another), and others appeared unannounced. Similarly using the theatre lists as a check on who was to have the operation was not possible as the lists were prepared less than a day before the sessions, leaving too little time for the baseline measures to be taken. Theatre lists, obtained from the computer system, were however scanned to check if any cases had been missed.

2. Pre-operative patient questionnaire and background information

Once identified, patients were asked if they would like to cooperate with the study by the research nurse who handed them a letter of explanation and the symptom and NHP questionnaires for completion. These were collected by ward nursing staff.

Usually information concerning the indications for surgery and medication was collected at the same time from the medical notes. If necessary this data could be collected

retrospectively.

3. Procedure details

Any problems encountered during the operation were recorded by the surgeon on a simple form kept in theatres. Though not in the original data set, during the course of the study an assessment of risk (of adverse outcome) was added to the data at the request of the surgeons. This was completed preoperatively using a standard pro-forma kept in theatres. The same sheet identifies any operative or anaesthetic problems that may arise.

4. Discharge information - post-op complications

The simple check list of possible post-op complications encountered was completed on discharge. Originally this form was attached to the summary abstract form completed by all consultants before cases could be coded. This was later changed so that the form was completed by the research nurse based on evidence in the notes.

5. Follow-up information

The follow-up information came mainly from the patients, who were asked to complete questionnaires (the same as the pre-operative ones) looking at symptoms and general health status. Patients were also asked additional questions covering re-admissions to hospital, and other adverse events. If no response was forthcoming within a month a reminder with duplicate form was sent out automatically. Follow ups were carried out at three and twelve months after discharge.

Some patient outcomes were examined in more detail within the reporting system and where patients had died, their notes were used as the basis for short individual reports.

6. Data Inputting

All the data were input to a micro-computer using compiled dBase software. All input program were written by the research team. The estimated costs of data collection are discussed in Chapter 9.

Validation

Three checks on the accuracy and validity of the collected information were used:

1. *Retrospective comparison to notes*

   An audit was carried out on 20 sets of notes to ensure completeness and accuracy of recording of all the basic data elements in the presenting characteristics, and in the operative and post-operative complications of patients. The results of that process are described below.

   Indications for surgery :- there were some significant differences between the indications when re-abstracted from the notes which highlighted some ambiguity in the precise meaning of the terms 'chronic cholecystitis' and 'history of acute cholecystitis'. In the analysis therefore an additional reclassification on an acuity scale was used to overcome this. However the most important distinctions for the acute conditions were accurately identified.

   Co-morbidities :- the forms were largely accurate though there was a tendency not to include disease which occurred a number of years before.

   Pain in last year :- general agreement on the frequency of biliary pain (typically at the most extreme end of the questionnaire's scale).

   Symptoms :- typically the questionnaires recorded more symptoms than the notes. This in itself may not be a significant problem.

2. *Patient interview*

   A random sample of patients was selected to validate the postal questionnaires and NHP results. Visits by CASPE researchers utilised a semi-structured interview covering the main points of the questionnaires. The researchers (who had not seen the patients' forms) were asked to make their own assessment of the patient's problems using simple categories (see Appendix 2). The researcher's assessment was then compared to the actual patient questionnaire results and the following points noted.

Time since diagnosis/symptom onset:- Agreement on the time since symptom onset was found in all cases while the timing of diagnosis agreed in 9 out of 13 cases. It was concluded that the time since onset of symptoms was a rather more reliable guide than time since diagnosis .

Overall health rating:- In most cases the general assessments on the scale (poor-fair-good-very good) agreed between those taken at interview and the questionnaire results (9/13 preoperatively 7/13 postoperatively). When differences did occur they were never greater than one point on the scale. The change in health rating showed similar agreement. There was no observed bias for questionnaires to record either higher or lower than interview.

Events after discharge:- There was no case that had been readmitted among the sample (as shown by interview and questionnaire). In three cases there had been visits to a GP which were not recorded on the forms. In fact in one case this was due to a visit after the questionnaire had been returned, while in the others the visit was part of a 'routine' which was not related to the operation (the questionnaire asks specifically about visits relating to complications following the operation).

Nottingham Health Profile :- Results were pooled with similar interviews from other studies and are reported in Chapter 3. The comparisons of actual NHP scores to the assessments made by the interviewer were satisfactory (and highly significant) for all except one dimension showing a high degree of association. The exception was the score for 'Energy', where little association was shown, however problems were noted at interview with the phrasing of the validation question for this dimension. It was concluded that by and large higher scores on the NHP were associated with the interviewers' assessment of more severe health problems.

Symptoms :- There were some discrepancies in the recording of symptoms. The results, pooled for all symptoms, are shown in Table 5.2(a). There were significant numbers of cases pre-operatively where symptoms revealed at interview were not on the form, as well as vice versa. The follow up comparison showed greater agreement

but there were fewer positive symptoms to report. It appears that the reliability of recording a particular symptom is not too high and that great reliance cannot be placed on it.

Table 5.2 (a) Comparison of questionnaire with interview responses - Presence of individual symptoms as reported by patients.

		Interview			
		Pre-op		Follow-up	
		Yes	No	Yes	No
Questionnaire	Yes	17	14	6	6
	No	11	49	3	76

Table 5.2(b) Comparison of NHP scores before admission to those obtained in hospital (n=19)

NHP Dimension	# agree	Mean Difference	Tau_b	Probability
Energy	15	8.6	.715	.0001
Pain	11	3.3	.832	.0000
Emotional Reactions	8	6.6	.740	.0000
Social Isolation	10	-3.1	.789	.0000
Mobility	12	-.1	.801	.0000
Part II	13	6.8	.735	.0001
Combined Score	12 *	3.1	.856+	.0000

+ Pearson correlation coefficient=.957 (p=.0000)

However across the two methods (questionnaire and interview) the total number of problems reported was significantly correlated pre-operatively (Kendall's tau_b=.415 p=.04) and at follow-up (Kendall's tau_b=.546 p=.02). There was a clear improvement in the total numbers of symptoms reported between pre and post-operative conditions.

Interpreting the differences in symptoms reported is complicated by differences that one might expect when eliciting information through interview and through questionnaire. The questionnaire explicitly prompts a number of conditions in a way that is more insistent than at interview. There is some tendency for the questionnaire to record more individual problems. Additionally, the elapsed time between completing the form and interview was up to 6 weeks during which time problems -

or their perceived significance to the patient may change. Finally, there must be some uncertainty about how consistent the patient will be with respect to reporting any one symptom as being problematic although with pain showing a greater degree of pre-operative agreement than other symptoms, it is possible that the patient's perception of pain as a problem may provide a useful marker.

3. Comparison of pre-admission and admission NHP

One of the possible dangers of giving the NHP form to patients who were already in hospital (in bed) was the danger that the form may 'over-state' their problems as they adopted a 'sickness' role. Thus comparisons of a form completed in hospital to the follow-up, completed at home, may become invalid. Therefore 19 patients were sent the form in advance to complete at home , and then given a repeat form (with apologies and explanation for the repetition) once in hospital. The scores on the two forms were compared.

The results in Table 5.2(b) show that there was no significant difference between the pre-admission scores and admission scores. The two scores were highly correlated. The tendency was for pre-admission scores to be slightly higher than those completed in hospital. A few individual patients did show sizeable variation between the scores, typically across a number of dimensions.

4. Internal consistency on empirical results

There were some items of data collected which duplicated other measurements. It was hypothesised that where this happened the two results, using different approaches, would be significantly associated. The results section shows these relationships largely supported internal consistency. In particular, as Table 5.6(a) shows the mean NHP score for Pain was significantly associated with pain reported elsewhere in the questionnaire.

C. Data Analysis

The analysis of results in cholecystectomy concentrates on a number of key questions:

1. Do the indicators show significant changes following cholecystectomy?
2. Are there relationships between the indicators that confirm they are behaving as expected? Can they be thought to represent the outcome of the care given?
3. Are the indicators sensitive to the presenting characteristics of the referring population?

Deaths

At the start of the study the numbers of patients who died was expected to be very low (1-2%). Only two cases died in hospital after the procedure. In fact, with mortality assessed at 3 months and 12 months post-op the numbers have been higher than expected. These cases tend to have been reviewed individually and not used in further statistical analyses. The deaths tended to be restricted to patients with cancer of some form, patients with complex disease where the Freeman Hospital had been used as a tertiary referral or to the elderly.

Changes in the indicators to 3 months (Tables 5.3(a) - 5.3(e))

Symptoms

Changes in the presence of individual symptoms are shown in Table 5.3(a). There was clearly some variability in the prevalence of the symptoms before operation, ranging from 30% of cases reporting "loss of appetite" to 68% of cases who "cannot tolerate fatty food". It is clear that patients did not appear to suffer from all the symptoms, the average number reported being 3.3. For each symptom there were some cases who appeared to suffer from the problem after the operation but not before - the most marked symptom in this respect being "bowel problems" . The conclusion here would seem to be that either this represents uncertainty in the reporting of these symptoms (which seems more likely) or that problems have been created by the hospitalisation itself. The observed prevalence of any one symptom must therefore be expected to include a number

for which confidence intervals can be calculated.

Table 5.3(a) Symptoms/Problems before and 3 months after cholecystectomy (n=149)

Mean No. Symptoms/problems per patient: Pre-op=3.30 sd=1.77
Follow-up=1.67 sd=1.6

ns=p>.05

% cases	No-No	Yes-No	No-Yes	Yes-Yes	Sig	Prevalence [95% Conf. limit]	Net change	Prop change
1 Pain	55.7	27.5	6.7	10.1	.000	37.6 [30-45]	20.8	55.3
2 Flatulence	34.9	33.6	8.1	23.5	.000	56.5 [49-65]	25.5	45.1
3 Dist. Abdomen	46.5	28.2	4.7	21.5	.000	49.7 [42-58]	23.5	47.3
4 Vomit	39.6	46.3	3.4	10.7	.000	57.0 [49-65]	42.9	75.3
5 Bowel	54.4	18.1	14.8	14.8	ns	32.9 [25-40]	3.3	10.0
6 Appetite	59.1	18.8	10.7	11.4	ns	30.2 [23-38]	8.1	26.8
7 Fatty Food	27.5	43.6	4.7	24.2	.000	67.8 [60-75]	38.9	57.4

Table 5.3(b) Events in hospital

Peri-operative complications = 29 cases (20.1%)
Explore Common Bile duct = 7 cases
Empyema/pus = 3 cases
'Anaesthetic/recovery' = 3 cases

Post-operative complications = 33 cases (22.9%)
Urinary Retention = 10 cases
'Drain problems' = 7 cases
Resp infection = 4 cases

Table 5.3(c) Events after discharge

Visits to Accident & Emergency = 12 cases (8.3%)
Visits to GP = 31 cases (21.5%)
Readmission = 8 cases (5.6%)

New Health Problems = 10 cases (6.9%)
Feels no better = 35 cases (24.3%)

Table 5.3(d) Summary Health Status - Pre-operative vs follow-up assessments

N/R=Not recorded Sig Chi-squared p<.0001

Pre-op assessment versus follow-up (3 months)

Pre-op \ Count	Follow-up assessment N/R	Poor	Fair	Good	Very G.	Row Total
N/R			1	2		3
Poor	1	3	10	5	1	21
Fair	2	1	23	25	3	54
Good	2	1	17	22	15	55
Very Good			1	5	10	16
Column Total	3 2.0	5 3.4	52 34.9	60 40.3	29 19.5	149 100.0

Table 5.3(e) Changes in Nottingham Health Profile. Mean scores before and after cholecystectomy

	Pre-op				Follow-up				
	Mean	Std Dev	Median	%Zero	Mean	Std Dev	Median	% Zero	Sig(p)*
Energy	35.48	38.30	24.0	44.3	20.88	32.08	0.0	63.1	<.000
Pain	27.80	29.28	19.7	32.9	10.13	20.99	0.0	68.5	<.000
Emotional R.	16.58	19.89	9.3	40.9	7.55	16.34	0.0	70.5	<.000
Sleep	29.09	28.44	21.7	27.5	18.02	25.77	0.0	53.7	<.000
Soc. Isol	7.51	16.93	0.0	79.2	6.01	15.33	0.0	81.9	.31
Mobility	14.25	21.36	0.0	54.4	10.90	17.46	0.0	61.1	.12
Part II	25.10	28.75	14.3	43.6	10.72	21.40	0.0	73.8	<.000
Combined	21.78	18.53	18.1	8.1	12.24	15.21	6.1	34.2	<.000

* Wilcoxon signed pairs test.

It is worth noting that there was a high proportion of residual symptoms reported 3 months after the procedure - even if the problem was present at admission it had not necessarily disappeared by the first follow-up.

Table 5.3(a) shows improvements expressed in two ways, the first is the net change in percentage of patients reporting a problem, comparing before and after the operation. This represents the percentage of patients showing 'Yes-No' movements less the percentage showing 'No-Yes' changes. The second method is to express this change relative to the proportion of patients with that problem before the operation.

The results show clear and significant reductions in the reported incidence of pain, flatulence, vomiting, abdominal distension and tolerance of fatty foods. The changes were tested using McNemars test. There was little or no overall change in the incidence of bowel problems or loss of appetite.

In addition to changes in individual symptoms, an aggregate score was calculated to represent the total number of symptoms/problems reported. This score weights all symptoms equally - in the absence of any better knowledge. Using this score patients were observed to report significantly fewer symptoms/problems at 3 month follow-up (1.67 compared to 3.3 $p < .0001$).

This pattern of results has been consistently demonstrated during the course of the project. One area of concern has been the lack of change in the incidence of bowel

problems which prompted the question of whether this was due to constipation or diahorrea - the latter having significant clinical implications. The forms were amended to ask directly about these two problems (whilst retaining the original question). To date there appears to be no clear pattern of either complaint diarrhoea or constipation, predominating at follow-up. The results are in general internally consistent in that patients who report a 'bowel' problem also tend to report either one of diarrhoea and/or constipation (not both).

Events within hospital

The incidence of complications at the time of operation was simply based on the comments made by surgeons at the time. As this was not based on a specific schedule of questions it must be considered to be relatively sensitive to differences in recording practice. Most commonly cited was that exploration of the common bile duct had been necessary. Otherwise the comments made were fairly diverse and included problems with adhesions, necrosed or shrunken gall bladders etc.

Post-operative complications were based on a checklist of common problems that may arise and hence recording could be considered more reliable. In most cases the reported problems were relatively minor in clinical terms - with urinary retention being the most common. There were a number of reported problems involving t-tubes and drainage during the post-operative period.

Events after discharge

Following discharge, there was a relatively high incidence of events though the value of quantifying some of these was questioned. Perhaps the most significant indicator is the 5.6% of cases readmitted. Some of the readmissions were related to the original cholecystectomy whilst others were not.

The reasons quoted included :

- 'Replace blocked arteries in stomach'
- 'Infection in wound'
- 'Road traffic accident'
- 'Pain and sore around wound'

- 'Scar tissue turned septic'
- 'To convalescence'
- 'Pain and vomiting'
- 'Pain and coughing up blood'
- 'Heart attack'
- 'Distended stomach'

Not all of these were relevant to the prior surgery and it was clear that assessing the success of the procedure based simply on the total of re-admissions would inevitably include some false positives. The distinction between relevant and irrelevant re-admissions was found to be hard to make in advance and the reasons for readmission given by patients tended to be limited. It was decided that the best way to analyse these events is to consider them individually although with such an approach retrieving the relevant data can be time consuming.

Summary health status

The simple grading of patients health from poor to very good showed a significant change after the procedure. There was some natural variability between patients in how well they felt, with a substantial number even rating their health as 'very good' just before surgery. After surgery, 45 out of 144 patients gave the same rating as before though the net trend was to improve (Chi-squared $p < .0001$) , typically by one step up the ordinal scale. A few patients reported their health at 3 months as worse than pre-operatively.

Two simple questions were asked of patients as a check on this and other information. 7% of patients reported 'health problems they did not have before' and 24% reported that they felt no better than before. The links between this information and the other indicators are discussed later.

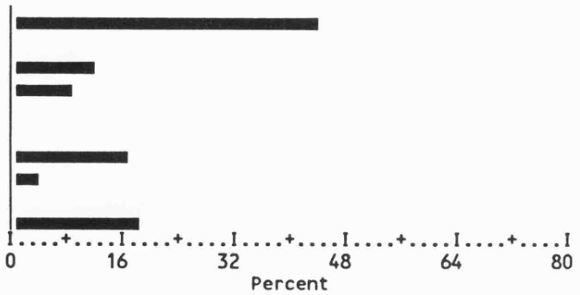
Nottingham Health Profile

The Nottingham Health Profile is scored as the sum of weighted responses in each of six dimensions, the range being from 0% to 100% on any one dimension (higher scores indicate 'worse health'). In addition seven questions on Part II asked about the effects of health on everyday life. The results of Part II questions are shown in similar fashion to the other dimensions by giving each positive response an equal weight. In addition the

Fig 5.1 NHP distributions before cholecystectomy. Percentage of cases with given scores for each dimension.

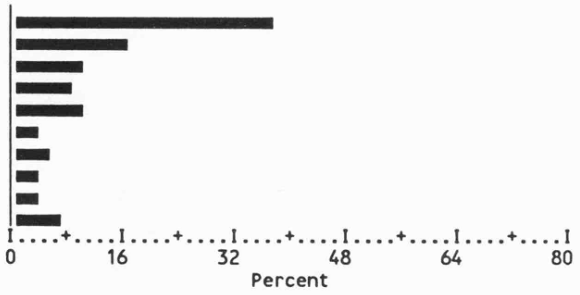
Energy

Count	Midpoint
67	5.00
0	15.00
17	25.00
13	35.00
0	45.00
0	55.00
25	65.00
6	75.00
0	85.00
28	95.00



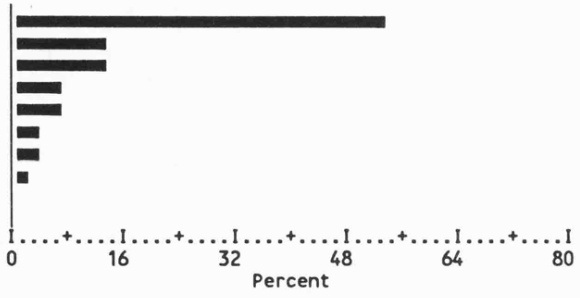
Pain

Count	Midpoint
57	5.00
25	15.00
14	25.00
13	35.00
15	45.00
6	55.00
8	65.00
4	75.00
5	85.00
9	95.00



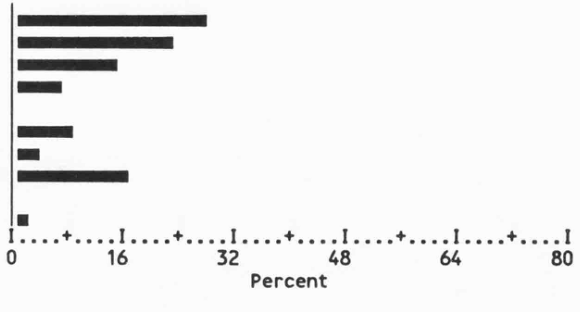
Emot.Reactions

Count	Midpoint
83	5.00
19	15.00
20	25.00
10	35.00
11	45.00
6	55.00
4	65.00
3	75.00
0	85.00
0	95.00



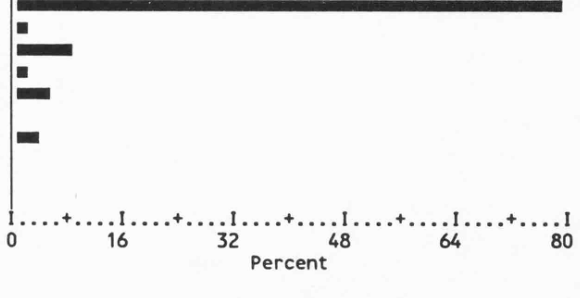
Sleep

Count	Midpoint
43	5.00
34	15.00
22	25.00
11	35.00
1	45.00
12	55.00
6	65.00
24	75.00
1	85.00
2	95.00



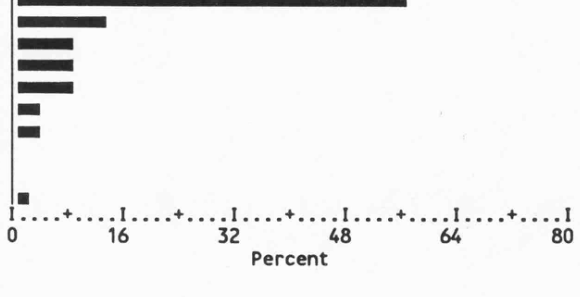
Social Isolation.

Count	Midpoint
123	5.00
2	15.00
13	25.00
2	35.00
8	45.00
1	55.00
6	65.00
0	75.00
1	85.00
0	95.00



Mobility

Count	Midpoint
87	5.00
20	15.00
12	25.00
12	35.00
13	45.00
4	55.00
5	65.00
0	75.00
1	85.00
2	95.00



six scores were combined into a single score - based on weighting each dimension equally (as discussed in Chapter 3). It is noted that summaries of Part II responses and combining score was not recommended by the developers of the profile (Hunt,McEwen & McKenna 1986) and is used in this context in order to simplify the presentation of results.

The results showed significant improvements following the procedure in most of the dimensions and in the combined score. The general pattern of scores on admission were broadly in line with those expected for this type of patient with higher scores for 'Energy' and 'Pain', than say for 'Mobility' and 'Social Isolation'. The dimensions with higher scores on admission were also those that showed significant changes compared with the follow-up. The variance of any one dimension was typically fairly large the standard deviation being approximately equal to the mean. One would expect greater confidence in estimates of the mean score with larger sample sizes.

The distributions of scores on all dimensions (Fig 5.1) were highly skewed with a substantial proportion of cases scoring zero - especially for the follow-up score. The effects were seen in the most extreme in the dimension covering 'Social Isolation' where for the admitting scores 79% of patients scored zero. A number of patients (8%) failed to score on any of the main dimensions. The skewed distribution resulted in the follow-up scores on all dimensions having a median score of zero (see Table 5.3(e)). The highly skewed nature of the distributions required non-parametric tests of statistical significance.

Changes in Indicators to 12 months (Table 5.4)

Table 5.4 summarises the indicators for the 80 patients who have received a 12 month follow-up in addition to that at three months. The changes to three months of this subset of cases echoed the changes seen in the overall population. It is clear that subsequent to the three month period there was little change in the symptoms/problems reported by patients, the summary health status or in the NHP. The observed differences were not significant between three and twelve months.

Table 5.4 Change in indicators to twelve months (n=80)

	Admission	3 months	12 months
Prevalence of symptoms % cases			
Pain	41.3	17.5	17.5
Flatulence	60.0	36.3	32.5
Distended abdomen	53.8	28.8	35.0
Vomiting	62.5	16.3	11.3
Bowel problems	25.0	32.5	36.3
Appetite	33.8	27.5	26.3
Fatty Food	70.0	33.8	35.0
No. Symptoms (mean)	3.46	1.93	1.94
Readmitted (% cases)	-	6.3	17.5
Go GP (% cases)	-	22.5	28.8
A&E visit (% cases)	-	6.3	16.3
More Health problems (%)	-	2.5	13.8
No Better (%)	-	26.3	27.5
NHP means Energy	38.9	24.4	23.6
Pain	26.0	12.1	10.5
Emot. R	15.5	8.1	7.7
Sleep	27.1	20.7	20.8
Soc.Isolation	7.5	5.9	5.8
Mobility	13.3	13.2	12.7
Part II	24.2	12.3	11.0
Comb	21.4	14.1	13.5
Summary =Poor (%)	17.5	2.5	3.8
Health =Fair	42.5	41.3	25.0
Status =Good	32.5	43.8	50.0
=V.good	7.5	12.5	16.3

Of the other indicators there was a notable rise in reported readmissions from 6.3% of cases at 3 months to 17.5% at 12 months after discharge. Similar rises in the reported incidence of visits to A&E departments occurred during this period. Looking at the reasons given by patients for these events they tend to be unrelated to the cholecystectomy procedure. It appears that other health events are creating some false positives on these two indicators.

The reasons quoted for re-admissions include:

- 'Cyst in throat'
- 'Broken collar bone due to fall'
- 'Severe chest infection'
- 'Heart attack'
- 'Tear duct op - for glaucoma'
- 'Broken ankle'

- 'Bladder repair, prolapse leg/back'
- 'Day patient - twice for bladder'
- 'D&C'
- 'Wound open due to infection'

It would appear that to monitor specifically the effects of the cholecystectomy a three month follow-up is appropriate - a longer time interval tends to decrease the specificity of some of the indicators.

Relationships between indicators

There are a variety of ways in which the relationships between indicators can be examined though not all can be reported here. The most important observations about the relationships within and between indicators are described below.

There were high correlations between individual dimensions of the NHP and between symptoms.

Correlations within NHP

With the NHP there were highly significant correlations between rank scores of different dimensions. The combined score was also closely related to the individual dimensions. Table 5.5(a) shows the rank correlation coefficients between the NHP dimensions and their statistical significance (all except one correlation being $p < .001$). The results show considerable overlaps between the dimensions, patients scoring high on one dimension are also likely to score high on others. In general the dimension

Table 5.5(a) Correlations (rank) between dimensions of the Nottingham Health Profile (n=143)

	Scores at admission							
Correlations:	Energy	Pain	Emot.Reac	Sleep	SocIsol	Mobility	Part II	Combined
Energy	1.0000	.4704**	.5501**	.3670**	.4167**	.5323**	.5237**	.8254**
Pain	.4704**	1.0000	.5872**	.4195**	.2988**	.5194**	.4695**	.7556**
Emot Reac	.5501**	.5872**	1.0000	.3712**	.4378**	.4103**	.5474**	.7201**
Sleep	.3670**	.4195**	.3712**	1.0000	.2235*	.3614**	.2790**	.6440**
Soc Isol	.4167**	.2988**	.4378**	.2235*	1.0000	.3584**	.2924**	.5320**
Mobility	.5323**	.5194**	.4103**	.3614**	.3584**	1.0000	.4944**	.7187**
Part II	.5237**	.4695**	.5474**	.2790**	.2924**	.4944**	1.0000	.5978**
Combined	.8254**	.7556**	.7201**	.6440**	.5320**	.7187**	.5978**	1.0000

1-tailed test: * p<.01 ** p<.001

Table 5.5 (b) Frequency of cases with symptom pairs - percentage of cases on admission suffering from two symptom/problems (n=149)

Percentage cases with two problems	Prevalence	Pain	Flat.	Abdo. Disten.	Vomit	Bowel	Appet.
1 Pain	38%	-					
2 Flatulence	61%	27	-				
3 Distended Abdomen	50%	26**	42**	-			
4 Vomiting	56%	25	39*	35**	-		
5 Bowel problems	32%	13	20	19	19	-	
6 Appetite	32%	13	18	14	19	17**	-
7 Fatty Food	67%	29	48***	41***	46***	21	19

* p<.05 Chi-squared
** p<.01
*** p<.001

covering 'Social Isolation' showed weaker correlations with the other dimensions. The correlations with the summary Part II scores were also fairly low - although given that this dimension was not supposed to be scored in this way this is hardly surprising. It had been hoped that the Part II score may give a general summary of the whole of the NHP, but it appears that the combined score performs much better in this role as a statistical representation of the six dimensions. This is not to suggest that the six scores are unnecessary for interpretation of changes (see Chapter 3) and in outcomes monitoring in general.

Similar high correlations between the dimensions of the NHP were observed comparing the follow-up NHP scores.

Links within Symptoms

Table 5.5(b) shows the prevalence of presenting symptoms/problems and the proportion of cases suffering from two problems in the various pairing combinations. When the presence of symptoms were compared (in this case in pairs) with what might be expected by chance, there were once again a high number of significant relationships.

It appears that the symptoms fell into two groups each with high internal links - one group covering pain, abdominal distension, flatulence vomiting and tolerance of fatty

food. These symptoms tended to cluster in patients. The other group covered poor appetite and bowel problems which show no close correlations with any other symptoms other than themselves. It should be noted that these two questions require a negative answer to indicate a problem - the reverse of the other group which may partly explain the difference. In addition they are the least specific of the symptoms and the only symptoms not to show a significant change to three months. The fact that their behaviour is so different from the rest must raise questions over their validity or reliability of capturing the data.

There was general agreement between the overall symptom scores, the summary health status and NHP scores.

The relationships between these indicators were all positive and tended to be statistically significant as shown in tables 5.6(a)-5.6(c). Thus in general, high values on the NHP tended to be accompanied by higher incidence of reported symptoms and lower summary health status as described by the patient. These relationships held when comparing across either the data sets at admission or at follow-up.

Table 5.6(a) shows the significance of observed relationships between individual symptoms and the dimensions of the NHP using the Wilcoxon rank sum test. It is reassuring to observe highly significant relationships between pain scores on the two measures. In fact pain, distended abdomen, vomiting and intolerance of fatty foods all appeared to be linked to the NHP. It should be noted that flatulence did not appear to be linked to any dimensions of the NHP. Interestingly, given earlier provisos, poor appetite and bowel problems showed links but these were much weaker. The NHP dimensions covering 'Social Isolation' and 'Mobility' showed weaker relationships with the symptoms scores than the others. These were also the dimensions with the lowest mean scores which may suggest that they are less relevant as health problems - in this patient group - than the other dimensions.

Table 5.6(a) Significance of relationship between NHP scores and individual symptoms (n=149)

	NHP dimensions							
Symptom	Ener.	Pain	Emot Reac	Sleep	Soc Isol	Mobil	Part II	Comb.
1 Pain	.0015	.0000	.0037	.0016	ns	.0334	.0018	.0000
2 Flat.	ns	ns	ns	ns	ns	ns	ns	ns
3 Dist Abdo	.0005	.0000	.0000	.0041	.0261	.0039	.0110	.0000
4 Vomit	.0117	.0017	.0001	ns	ns	.0498	.0011	.0014
5 Bowel	ns	.0139	.0208	.0290	ns	ns	ns	.0295
6 Appetit	.0209	.0432	.0094	ns	ns	.0547	.0113	.0225
7 Fatty F	.0172	.0002	.0000	ns	.0065	ns	.0214	.0013

Wilcoxon Rank sum test. ns = p>.05

Table 5.6(b) shows highly significant links between the total symptom count and a consistent gradient of decreasing NHP scores as the patients summary of their health improves from 'Poor' to Very good'. Similarly Table 5.6(c) shows that the better the patient feels the fewer symptoms they appear to report.

Table 5.6 (b) Significance of association between select dimensions of NHP and symptom scores and summary health rating. Mean NHP scores by summary health rating (n=153).

	Energy	Pain	Emot R	Sleep	Soc Isol	Mobil	Pt II	Combined
Symptoms p	.0011	.0000	.0000	.0007	.0651	.0163	.0001	.0000
Summary Health rating Sig.	.0000	.0000	.0000	.0006	.023	.0000	.0000	.0000
Means Poor	81.7	54.4	35.0	41.7	17.6	28.1	45.7	43.1
Fair	41.9	29.8	20.4	35.4	9.5	10.0	28.3	25.8
Good	21.1	20.6	8.5	20.0	3.6	9.0	17.8	13.8
Very Good	7.0	5.1	5.4	21.7	5.2	5.0	14.3	8.2

The significant relationships between the NHP, total symptom score and summary health status adds supporting evidence for the validity of all three measures. It also suggests that in order to identify the 'sicker' patient at presentation , or the successful outcome of care, they may not all be necessary and that there is some redundancy in the data.

Table 5.6(c) Frequency of cases reporting symptoms by summary health status (n=153)

	% cases	Poor	Fair	Good	V Good
% cases	-	13.5%	34.8%	35.4%	10.3%
Mean No. symptoms p<.0000	-	5.10	3.41	2.93	2.13

Table 5.6(d) Mean changes in NHP scores in selected dimensions against mean changes in symptom score and summary health status (n=153).

		Change Energy	Change Pain	Change Combined
Change in no. Symptoms	-2 n=5	-15.2	-7.2	-12.9
	-1 n=14	11.5	8.7	9.6
	0 n=22	11.3	9.1	5.1
	1 n=31	8.5	19.3	10.2
	2 n=32	18.2	15.9	8.7
	3 n=18	27.2	24.4	15.1
	4 n=22	22.6	32.9	16.1
	5 n=8	33.4	35.3	23.2
Change in summary rating	Worse n=32	7.0	14.1	8.8
	Same n=58	13.2	13.7	8.1
	Better n=63	21.3	24.5	12.7

As well as comparing these indicators in cross-section, it is also possible to examine the relationships between changes in indicators. In general when these changes are compared they broadly agree - though the correlations between movements in the three indicators are naturally not as strong. Table 5.6(d) shows the mean change in selected NHP dimensions ('Pain', 'Energy' and the combined score) against changes in total number of reported symptoms (admission minus follow-up) and changes in summary health status. For the few patients whose symptoms were worse or the same, the mean change in NHP scores were less than for those patients who improved on a number of symptoms. Thus it appeared that the magnitude of change in the total symptom score is related to the magnitude of change in these NHP dimensions. The relationships between changes in symptoms and changes in the other dimensions are not as strong.

Similarly the patients who are 'worse' on the summary health status show a smaller change in mean NHP though these differences are not significant. This may indicate that the definitions of 'better', 'same' and 'worse' used for the

summary health status indicator are not sufficiently specific. If the definitions were altered such that 'better' is defined as a change of two steps in the ordinal scale then clearer links were visible. Interestingly the mean change for these patients is still positive and not negative which is what would be expected if the two measures were in perfect agreement. That is patients who were 'worse' on the summary health status would score higher on the NHP at follow-up than at admission.

Finally there were also significant differences between the change in the number of reported symptoms for those patients defined as 'better' on the summary health status (average -2.1 symptoms per patients) against those defined as 'worse' (average -0.97 symptoms per patient, $p=.015$).

These results therefore lend evidence for the validity of observed changes in the indicators - though the links between indicators are less strong than when they are compared in cross-section. Nevertheless there were the predicted relationships between improvements in the symptoms reported, improvements in the NHP and improvements in the summary health status.

The fact that the relationships between changes in the indicators are not as strong as the relationships between the indicators when compared in cross-section suggests that using these indicators it is easier to define a patient as 'sicker' than to define a patient as 'better'. There are a number of reasons why the change in the indicators show more variability. The obvious one is that the degree of change requires two measures each of which will have inherent variability which when combined will tend to make the answers even more variable. A second difficulty with measures of change is that there may be problems with 'end-effects' of the scales used, that is the NHP and symptom scores both tend to zero after the procedure. This may 'underestimate' the relative improvement in patients with low initial scores. It may be that simple linear comparisons of changes in indicators are not appropriate. Finally it may be that assessment of changes in score is limited by the construction and weighting of the instruments

which may not be appropriate to measure longitudinal change with sufficient specificity. Further analysis would be required to examine these alternatives and improve the association between one indicator and another.

Absolute changes in NHP and symptom scores and NHP were strongly related to the presenting score.

With the NHP there are clearly some 'end-effects' when a very high proportion of patients do not score post-operatively. If the magnitude of change is taken as the indicator of outcome then the sicker the patient on admission the more chance there is of improvement.

The correlations between the initial NHP and subsequent change in each dimension were typically high ranging from .62 for 'Sleep', to .75 for 'Pain'. Fig 5.2 shows as illustration the mean change in combined scores according to the initial score. In this case it can be seen that average change in score is very much lower when the initial score is low. The maximum possible change is shown for comparison. The same relationship was observed between the change in total number of reported symptoms/problems and the total symptoms/problems present at admission - the correlation coefficient being 0.60.

The importance of this observation is that using the change in these indicators as a measure of success, results will be extremely sensitive to the profile of admitting scores. This may be valid in clinical terms. The patient who presented with very severe health problems and had none at follow-up, obtained a greater benefit in outcome terms than the patient who was 'fairly healthy' before the operation. For comparison between changes in scores for different populations of patients some standardisation may be necessary to exclude the variability in admitting score.

Post-operative and peri-operative complications were significantly related to each other - but not to any of the other indicators.

Table 5.7 shows the proportion of patients who experienced adverse events either during the hospital stay or afterwards. It can be seen that patients with 'peri-operative complication' were also more likely to experience post-operative complications ($p < .001$). However these events in hospital were not directly related to the later events after discharge. One implication of this result is that events within the hospital may complicate cases but will not necessarily result in worse outcomes over a longer term. Earlier analysis of the assessments of risk made by the surgeons (for the relatively few patients where data were available) suggested that this was strongly linked to the incidence of peri-operative problems but not to longer term outcomes.

There were correlations between events after discharge such as visits to GP's, A&E departments and re-admissions.

The second part of Table 5.7 shows the proportion of cases who experience adverse events after discharge. The results suggest that these events tend to be related and that patients who are readmitted will be more likely to be those patients that have been to the GP or to Accident & Emergency departments.

The final section of Table 5.7 suggests that the responses to the statement "I feel no better than I did before" did not show any relationship to post-operative complications. Experience during validation interviews has suggested that there were some problems with the interpretation of this question and that the results are therefore not reliable. The other summary of transitions in health was the statement "I have health problems I did not have before" where positive responses to this question are associated with a higher incidence of visits to GPs and A&E - which is as predicted.

Fig 5.2 The mean change in combined NHP score by admitting score.

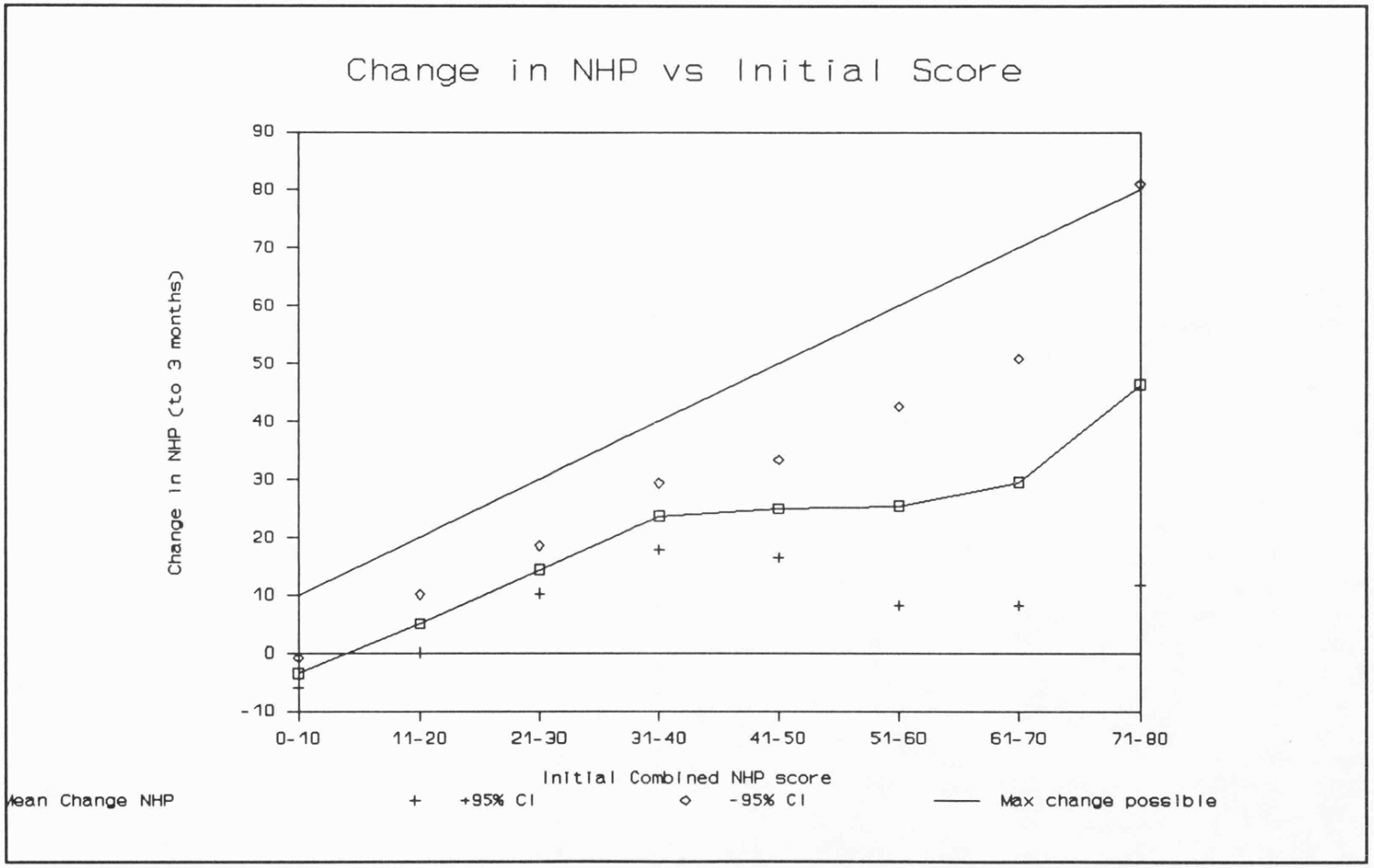


Table 5.7 Co-incidence of events during stay and to follow-up (n=149).

Percentage of patients reporting particular events/problems after operation.

	Prevalence % cases	Periop	Postop	Readmit	Go GP	Go A&E	No better
Periop complics	17.8	-					
Postop complics	21.7	10.2***	-				
Readmitted	5.1	0.6	1.3	-			
Go GP	20.4	4.5	6.4	3.2**	-		
Go A&E	7.0	1.3	1.3	2.5***	3.8**	-	
'No better'	28.0	4.5	6.4	2.5	6.4	1.3	-
'Health problems'	5.7	1.3	1.3	1.3	3.8**	1.9*	2.5

*** = p<.001 Chi-squared test
** = p<.01
* = p<.05

The relatively small proportion of cases readmitted to hospital after three months showed significantly worse symptom scores, NHP scores and summary health status.

Table 5.8 records the mean follow-up NHP scores for each dimension and number of symptoms according to whether patients experienced one of the possible adverse events after discharge.

In general patients with peri-operative and post-operative problems showed no higher NHP or symptom scores than those without. This confirms the earlier observation that events during the hospital stay have not necessarily translated into longer term problems.

On the other hand patients who had been readmitted showed significantly higher NHP scores. This relationship is important if readmission is to be used as a proxy measure (Chambers & Clarke 1990). It appears that using these criteria readmitted patients are genuinely in worse health although the causal link to the earlier procedure still needs examination. Similar, though weaker relationships were also seen for the group of patients who reported visiting their GP.

Table 5.8 The relationships of health events and health status (n=149).

Mean NHP scores and number of symptoms by various post-operative events.

	Energy	Pain	Emot Reacs	Sleep	Social Isol	Mobil.	Pt II	Comb	Mean # symps
Periop Yes	20.3	9.1	7.6	18.6	5.8	10.0	9.5	11.9	1.7
No	23.0	12.7	5.9	11.1	5.8	14.3	16.5	12.1	1.6
	ns	ns	ns	ns	ns	ns	ns	ns	ns
Postop Yes	18.4	9.2	6.9	17.8	5.9	8.7	9.1	11.2	1.7
No	29.3	11.6	8.9	15.5	5.3	17.9	16.0	14.8	1.7
	ns	ns	ns	ns	ns	p=.050	ns	ns	ns
Read- Yes	18.6	7.8	6.3	15.8	4.8	9.8	8.7	10.5	1.6
mit No	59.0	45.0	26.3	44.8	23.2	27.4	46.4	37.6	3.0
	p<.001	p<.001	p<.001	p=.012	p<.001	p=.040	p=.001	p<.001	ns
Go GP Yes	18.2	7.5	6.9	14.8	5.3	10.5	8.6	10.6	1.5
No	30.4	17.3	9.0	27.1	7.9	11.3	18.8	17.2	2.5
	p=.046	p=.026	ns	p=.015	ns	ns	ns	ns	p=.001
Go A&E Yes	19.7	8.1	7.2	16.6	5.5	10.1	9.5	11.2	1.6
No	33.8	31.5	9.2	26.6	9.8	18.7	26.0	21.6	2.2
	ns	p=.001	ns	ns	ns	ns	ns	ns	ns
Better Yes	16.5	7.6	5.0	14.0	4.4	7.8	8.6	9.2	1.5
No	31.3	15.0	13.1	25.7	9.2	18.0	15.9	18.8	2.0
									ns
H Prob Yes	19.8	8.8	6.8	17.4	5.5	10.3	10.1	11.4	1.6
No	36.3	24.8	15.6	15.3	9.7	17.7	20.6	19.9	3.6

Mann-Whitney U test. ns=not significant (p>.05)

Effects of presenting characteristics.

The relationships between the various indicators and some basic presenting characteristics have been examined using simple bivariate analyses. A summary of the results are presented in this section. The importance of examining variation in outcome indicators by such presenting characteristics is that see if the indicators are unduly sensitive to changes in the mix of patients. If so then comparisons of indicators - between sites or over time - will require some form of standardisation for these patient characteristics. A second purpose was to consider whether relationships turned out as expected and if possible to reduce the size of the data set collected by eliminating variables which proved to be of less value.

The presenting characteristics used were:-

- Age
- Sex

- Indications for surgery using the groupings of 'Acute', 'History of acute', 'Chronic', 'Others'.
- Acuity scale as described earlier (1=Emergency, very acute; 2=Acute with frequent biliary pain; 3= Biliary pain less than 4 times/year ; 4=asymptomatic, no biliary pain).
- Co-morbidities - patients were classified according to the presence of active co-morbid conditions. Groups were for no co-morbidities, one condition, greater than one condition.
- Mean time since onset of symptoms - less than 6 months, 6-12 months or greater than 12 months.
- Obesity - defined as BMI > 30 kgm^{-2} (Bray 1978).
- Smoking

There were no significant differences in the age profiles of the two sexes though the mean age of men (60.95 years) is slightly higher than women (58.89 years).

Symptoms/problems

Table 5.9 summarises the mean number of reported patient symptoms/problems according to a variety of presenting characteristics. Results are shown for pre-operation scores, follow-ups to 3 months and the change in score. There were significant differences between the sexes in the number of reported symptoms/problems on admission - and a lesser difference at follow-up. However the mean change in number of reported symptoms was not significantly different.

Most of the differences observed for the other variables were not significant - with the exception of the acuity scale which showed higher symptoms scores for the most acute patients on admission although at follow-up there was no clear pattern across the acuity scale. The result is that the greatest change in reported symptoms/problems is found amongst the acute patients who tended to have more symptoms pre-operatively - as one might expect. The fact that the pattern of results were different for the Acuity scale and the indications for surgery (as recorded on the forms) raises questions over the recording of the latter variable (as was discussed earlier).

Table 5.9 Mean number of reported symptoms/problems pre-operatively, at follow-up and change in between by presenting characteristics (n=149)

	Mean No. reported symptoms/problems		
	Pre-op	Follow Up	Change
Female	3.59	1.87	1.72
Male	2.61	1.20	1.41
	p=.0018	p=.025	ns
%Age<40	3.66	1.55	2.10
40-59	3.25	1.89	1.36
60-69	3.43	1.86	1.57
>=70	2.85	1.21	1.65
	ns	ns	ns
% Acute	2.83	1.08	1.75
History	3.12	1.61	1.51
Chronic	3.64	1.96	1.68
Others	3.63	1.37	2.25
	ns	ns	ns
Acuity =1	3.69	1.00	2.69
=2	3.67	2.00	1.67
=3	2.51	1.51	1.00
=4	2.68	2.00	0.67
	p=.0003	p=.04	p=.0022
Comorbids.=0	3.39	1.58	1.82
=1	3.21	1.76	1.44
>1	3.26	1.71	1.55
	ns	ns	ns
Mean Time <6m	3.00	1.73	1.27
since 6-12m	3.00	1.30	1.70
symptom >12m	3.73	1.87	1.86
onset	ns	ns	ns
Obese No	3.37	1.68	1.70
Yes	3.20	1.47	1.73
	ns	ns	ns
Smokers No	3.30	1.64	1.66
Yes	3.30	1.74	1.56
	ns	ns	ns

ns = not significant (p>.05)

Events in Hospital and Events after discharge

Table 5.10 summarises the proportion of cases associated with a particular adverse event after discharge according to patient characteristics. Once again the most significant variable appeared to be the acuity scale. The more acute cases had a greater number of peri-operative and post-operative complications. Fewer of the more acute cases tended to visit their GP after discharge.

Table 5.10 Percentage of cases recording adverse event (during and after hospital stay) by various presenting characteristics (n=149).

% cases	Adverse event					
	Periop complic	Post-op complic	Readmit	Visit GP	Visit A&E	No Better
Incidence(%)	16%	19.0%	5.6%	21.8%	7.7%	28.9
Female	16.2	21.9	3.8	20.5	11.4	29.5
Male	15.9	13.6	9.1	21.0	5.7	25.0
	ns	ns	ns	ns	ns	ns
Age: <40	10.5	23.7	7.9	34.2	7.9	18.4
41-60	14.3	17.9	3.6	21.4	7.1	28.6
61-70	14.3	9.5	0	16.7	4.8	31.0
>70	23.5	26.5	11.8	14.7	11.8	38.2
	ns	ns	ns	ns	ns	ns
Indications:						
Acute	33.3	33.3	8.3	8.3	8.3	41.7
History Chronic	7.1	11.4	7.1	22.9	10.0	24.3
Other	23.2	25.0	3.6	25.0	5.4	28.6
	25.0	25.0	0	0	0	37.5
	ns	ns	ns	ns	ns	ns
Acuity =1	30.8	23.1	3.8	3.8	3.8	38.5
=2	12.7	19.0	7.6	32.9	8.9	26.6
=3	15.4	15.4	2.6	10.3	7.7	28.2
=4	0	0	0	0	0	0
	p=.0003	ns	ns	p=.001	ns	ns
Co-morbidities						
None	16.7	21.2	1.5	25.8	3.0	16.7
=1	15.4	17.3	7.7	15.4	11.5	28.8
>=2	16.1	19.4	9.7	19.4	9.7	51.6
	ns	ns	ns	ns	ns	p=.002
Time <6m	12.2	14.6	7.3	24.4	12.2	24.4
since <6-12m	6.1	15.2	9.1	12.1	6.1	24.2
onset >12m	23.8	27.0	3.2	23.8	4.8	34.9
Symptoms	ns	ns	ns	ns	ns	ns
Obese No	18.6	21.6	2.9	19.6	6.9	29.4
Yes	13.3	20.0	6.7	20.0	13.3	20.0
	ns	ns	ns	ns	ns	ns
Smokers No	19.8	19.8	5.7	17.9	9.3	26.4
Yes	7.0	18.6	4.7	27.9	2.3	32.6
	p=.05	ns	ns	ns	ns	ns

Chi-squared. ns = not significant (p>.05)

In most cases the co-morbidities are not showing the type of relationship that one might expect. The exception to this is that patients with active co-morbid conditions are more likely to feel 'No Better' after the operation. It is difficult to know if the lack of a relationship between co-morbid problems and these other events is due to the recording and classifying of the co-morbidities, or that the predicted relationship does not exist.

Nottingham Health Profile

Table 5.11(a) summarises the mean NHP scores by age and sex category. There are some differences in the NHP scores by age and sex though the pattern is not consistent.

Table 5.11(a) Mean NHP scores by categories of age and sex.

	SEX							
	F				M			
	AGE				AGE			
	<50 YRS	51-60YRS	61-70YRS	>70YRS	<50 YRS	51-60YRS	61-70YRS	>70YRS
No. Cases	31	17	28	23	7	11	14	11
Energy								
Pre-op	42.53	36.71	34.49	45.05	3.43	35.20	31.54	19.20
Post-op	28.36	27.91	14.37	25.63		14.62	16.17	13.45
Pain								
Pre-op	33.92	19.35	30.91	25.99	17.32	40.45	12.98	29.26
Post-op	11.98	5.56	9.73	15.90	4.77	6.94	4.82	15.75
Emot.Reac								
Pre-op	19.56	21.17	17.22	14.23	10.42	24.81	3.41	8.95
Post-op	9.84	8.20	7.42	3.84	1.01	5.12	7.55	4.82
Sleep								
Pre-op	25.15	40.52	27.20	27.05	22.38	28.98	28.12	25.23
Post-op	15.97	29.54	19.35	18.27	7.20	3.18	16.38	17.10
Social I								
Pre-op	5.41	13.69	8.93	12.30	3.22	4.05		1.45
Post-op	4.59	12.71	6.84	4.83		4.05	7.57	1.83
Mobility**								
Pre-op	9.72	7.46	15.29	29.57	8.32	17.68	5.72	15.57
Post-op	6.38	8.00	12.44	27.09		6.69	6.24	12.52
Part II								
Pre-op	37.80	18.49	22.97	18.84	30.62	25.98	9.19	27.28
Post-op	16.13	15.13	4.76	16.78		10.39	5.10	6.50
Comb								
Pre-op	22.72	23.15	22.34	25.70	10.85	25.20	13.63	16.61
Post-op	12.86	15.32	11.69	15.93	2.16	6.77	9.79	10.91

** significant correlation with age p<.001) women only
* significant correlation with age p<.01)

It appears that women tend to score higher (ie 'worse health') than men on most dimensions though this is most significant for 'Social Isolation' and 'Energy'. These differences are observed for pre-op and follow-up scores - but there are no differences in the absolute change in score.

There is no consistent pattern with respect to age though mobility scores tend to be higher with older women than with elderly men. Table 5.11(b) shows the significance of the relationships between some other presenting characteristics and the NHP scores at admission, at 3 month follow-up as well as the change in NHP to 3 months. Table

5.11(b) shows the observed statistical significance (using Kruskal Wallis test) between the NHP scores according to categories of sex, indications for surgery, the acuity scale, and the number of co-morbid conditions. The majority of tests revealed no association between the variables other than that expected by chance.

Table 5.11(b) NHP scores by presenting characteristics: Significance of relationships between NHP scores and presenting characteristics (n=149).

	Presenting characteristic			
NHP scores	**Sex**	**Indic-ations**	**Acuity scale**	**Co-morbid-ities**
Pre-op				
Energy	.014	ns	.042	ns
Pain	ns	ns	.0000	ns
Emot R	ns	ns	.005	ns
Sleep	ns	ns	.086	.035
Soc Isol	.008	ns	ns	ns
Mobility	ns	.0002	.085	.0006
Part II	ns	ns	.0017	ns
Combined	ns	.043	.0006	.031
Follow-up				
Energy	.007	ns	ns	.013
Pain	ns	ns	ns	ns
Emotional R	ns	ns	ns	ns
Sleep	.030	ns	ns	ns
Soc Isol	ns	ns	.033	ns
Mobility	.04	.0008	ns	.001
Part II	.03	ns	ns	ns
Combined	.001	.023	ns	.020
Change				
Energy	ns	ns	ns	.024
Pain	ns	.0002	ns	ns
Emotional R	ns	ns	ns	ns
Sleep	ns	.034	ns	ns
Soc Isol	.07	.049	ns	ns
Mobility	ns	ns	ns	ns
Part II	ns	.027	ns	ns
Combined	ns	.006	ns	ns

Kruskal Wallis analysis of variance.

Variation in the acuity scale showed the highest number of significant associations and seems the most important characteristic with respect to the admitting NHP scores. The pattern is that the more acute patients had significantly higher NHP scores on all dimensions (worse health) except 'Social Isolation'. There were no differences related to acuity at follow-up nor in the observed change in NHP. Given the earlier observations that higher pre-operative NHP scores tended to lead to higher changes in score, it is surprising that the higher NHP scores seen for the more acute cases did not also mean that higher changes in the overall score were observed. The fact that there were no significant differences in the change in NHP and the levels of the acuity scale may be an

artefact of the data or an indication of the power of the statistical test.

The relationship between the NHP scores and the classification of indications for surgery was rather different. The largest changes were noted in the change in NHP rather than in the admission or follow-up scores. The observation that more acute cases, as judged by the indications for surgery, exhibit greater changes in mean NHP scores, differs from the acuity score which failed to show similar links. Further investigation of the behaviour of these two variables is needed.

The presence of co-morbidities seemed to be related to the sum of the NHP dimensions at admission ('Sleep', 'Mobility' and combined score) and follow-up ('Energy, 'Mobility' and combined score). However the change in NHP between admission and follow-up was affected only for the score on 'Energy'.

Finally , though not included in Table 5.11(b), obesity and smoking were not significantly related to any of these NHP scores. It should be noted that this analysis is only a simple exploration of the strongest links between variables and does not necessarily imply a direct causal link. It is quite possible that some of the observed relationships may be due to combinations of variables, for example it may be that NHP scores for 'Mobility' are related to both age and the presence of co-morbidities - variables which tend to be associated anyway.

Though none of the variables examined show consistent relationships across NHP dimensions, it appears that sex and the balance of acute versus chronic patients are most important for making comparisons. In practice it has been found that the relative proportions of the sexes have been fairly constant during the course of the study. Thus comparisons of scores which have not been standardised for sex would be acceptable. However the proportion of acute cases has shown some fluctuations between quarters and may be the explanation of why poor outcomes are observed at certain time periods. The same cautions must also apply if NHP scores for cholecystectomies are to be compared between institutions.

Main conclusions of the data analysis - Cholecystectomy.

1. Most indicators show a general improvement following cholecystectomy. An early concern of the study was whether the various instruments used would be able to detect change following the operation. But the NHP, the presence of symptoms and the overall health rating all showed significant change for the population of patients concerned. This results is especially important in that it has demonstrated the value of the NHP in this particular context.

2. There are significant inter-relationships between the various indicators and some duplication of data. By and large the indicators when viewed as cross-sectional descriptions of the patient's health tend to support each other and give a validity to their use. These relationships are less clear when comparisons are made of movements in the indicators where they are clouded by variability in the instruments and more importantly 'end-effects' in using some of these scales.

3. There is sufficient variability in incoming patients to make the phenomenum of regression to the mean a problem especially when scores such as the NHP and symptoms scores tend toward zero in healthy patients. Thus, when considering the relative improvement in health associated with the procedure, a critical determinant is the admitting health status of the patient.

4. Despite the observed improvements in the population of patients there are clearly a number of patients who show little or no improvement and others where there are indications of health problems after the procedure. The proxy measures of readmissions, and to a lesser extent GP visits, are related to poorer health - though this in turn may not be related to the cholecystectomy.

5. Though the presence or absence of individual symptoms was reported, they tended to be rather unpredictable as outcome measures when viewed individually and the behaviour of questions on bowel problems and loss of appetite was confusing. A simple index of the total number of reported symptoms/problems appeared to be a more reliable indicator

of improvement.

6. Of the patient characteristics which appear to be related to the outcome indicators the most important was the acuity of the patients condition before surgery. Many of the other presenting variables showed some relationship to the indicators but there were few consistent and clear patterns in the results which necessarily suggest the need to standardise for these factors at admission. Larger sample sizes would give a clearer picture of the relevance of these factors and it is suggested that age, sex and the presence of co-morbidities (in some form) be retained in the dataset.

7. The major changes in outcome indicators occur within three months of the procedure and little change is seen thereafter to 12 months. Indeed events between 3 and 12 months may serve to confuse the picture as there is an increasing likelihood that deleterious changes in health are not related to the procedure. Interestingly problems during the hospital stay are not significantly associated with poorer outcomes at 3 and 12 months.

8. Further work could eliminate some of the variables collected in this study and operate on a reduced data set. The following is suggested:-

Outcome Indicators

Change in NHP or similar health status measure (baseline to follow-up).
Survival
Readmissions (and reasons)
Optionally:-
- Change in number symptoms/problems reported
- Peri-operative and post-operative complications

Patient descriptors and other information

Administrative details - identifier, address, etc.
Age and sex
Acuity/Indication for surgery classifications
Simple list of active co-morbid diseases
Optionally:-
- Clinicians assessment of patient risk.
- Process data, pre & post-operative lengths of stay.

D. Review Process

The results were fed-back to clinicians in the form of brief reports covering changes in the main outcome indicators, together with some more specific questions. As the results became available they were examined by quarters and reported back (see Appendix 5).

These reports were circulated to the consultant surgeons and then discussed at a subsequent meeting. As the recruiting of cases was slower than initially expected it was agreed that results should be reviewed every 3-4 months. In total four meetings have been held which all the consultant surgeons attended. As the consultant surgeons rarely all met together these meetings were specially organised, and there was clearly an interest in the approach and the results.

One attempt was made to introduce the study and interim results to the agenda of an existing meeting which included the whole surgical team. This was successful in disseminating the presence of the study though less productive in terms of discussing results. It was agreed that in future meetings with the four consultants would be the best forum.

Each meeting had an agenda and minutes were taken. The aims were to discuss both the process of data collection, the results to date and possible ways to improve the outcomes. A number of issues relating to data collection were raised and changes made as a result of these meetings. In particular the recording of the indications for surgery (the definition of which continues to be problematic) was examined, and the idea of the surgeons assessment of the risk of an adverse outcome introduced.

To summarise the reaction of the consultants:

a. The changes in symptoms scores were broadly as expected, though the high degree of residual bowel problems did raise some questions. Additional questions on diarrhoea or constipation were added to the forms but these have so far proved inconclusive.

b. The changes in NHP were broadly in line with those expected - that is significant improvements in the scores for 'Energy' and 'Pain' with little change in 'Social Isolation' or 'Mobility'. After some initial caution, the surgeons were happy to accept this instrument as a measure of general health status - indeed these results and their general acceptance are an important product of the cholecystectomy study.

c. A number of patients died, or had serious problems after surgery. Individual reports on these 'non-responding' patients were prepared. Early review meetings discussed these individually though no general lessons emerged. In some cases the history and subsequent course of these patients was known and the issues had been discussed in other settings. At the most recent meeting the value of group discussions of these individual cases was raised.

d. The lengths of stay were longer than expected - though they have fallen steadily throughout the course of the study. This was one clear example of where it is possible to monitor the effects of a change in resource use directly on patient outcomes In this case there was no deleterious effect on outcomes observed as the length of stay decreased

e. The interpretation of the results was hampered by the lack of comparative data. This made the focus of the discussions more concerned with the processes and feasibility of data collection than with potential improvements in practice.

f. The results did serve to highlight the main area of concern with regard to outcomes from this procedure, namely the process of patient selection. Though no direct changes in practice appeared there was a growing consensus that the key to good outcomes lies in patient selection rather than the actual execution of care. It has been agreed that the issue of appropriateness of the procedure will be looked at using two examples of criteria drawn from the literature. By implication other aspects of care were judged to be satisfactory.

g. The data was linked to the case mix system and in particular to the DRG classification scheme (Bardsley, Coles & Jenkins 1987). This raised some important questions

concerning the classification and the coding process. However it was felt that as the DRG classification was still in a state of flux (DRG definitions are currently being revised) and it was not worth investing too much effort to link these further.

h. The surgeons are keen to continue outcome studies of this form though the value of cholecystectomy as a relevant group for study was questioned. In particular it was expected that a major change in techniques will make the traditional procedure become obsolete. Indeed laparoscopic cholecystectomies have just started - the intention is to include these cases, suitably identified, in outcomes monitoring. Beyond this other proposed changes in the delivery of care such as the increase in day surgery have been suggested as opportunities for further work in outcomes monitoring.

j. The study has been seen by the surgeons as useful in exploring the relevant dimensions and techniques for outcome measurement and as one surgeon reported the study has been:

> *"Interesting as it showed elements of the patient's response to treatment we would not normally have access to" (Appendix 4).*

However the surgeons do not feel it has pointed to obvious ways in which their practice could be improved. The study has helped the surgeons to identify the process of patient selection as critical to good outcomes and was seen to offer a tool for ensuring that future changes in practice do not have deleterious effects on outcome.

In summary the cholecystectomy study has given some useful insights into the process of outcomes monitoring. There has been a clear commitment to the project from the clinical staff and a willingness to devote their time and resources to data collection and analysis. The surgeons concerned have expressed a wish to continue outcomes monitoring, possibly on other case types if the resources are available.

Chapter 6 Diabetes

A. Introduction

Diabetes was the second condition chosen in the early stages of the project. As a chronic disease where most care is based in the outpatient (or GP) sector it represents a completely different model to that of the acute inpatient episode of the cholecystectomy. The Department of Diabetes Medicine at the Freeman Hospital has close links with the University and other teaching hospitals in the city and has an international reputation for research in diabetes medicine. The key medical and nursing staff have inpatient beds, run outpatient clinics and education programmes to cater for a population of approximately 1300 diabetic patients.

Diabetes Mellitus - background information

It is estimated that diabetes affects over 1% of the general population and absorbs 4-5% of health care spending (Laing & Williams 1989). The disease itself is caused by the body's inability to produce sufficient amount of the hormone insulin which controls the metabolic balance of blood glucose. Whilst extreme imbalances in blood glucose can give rise to acute problems of hypo- and hyper-glycaemia, perhaps the major threat comes from the long term complications of the disease which affect many different organ systems. If untreated these complications can lead to very serious problems, blindness, amputations, kidney failure, heart attacks and strokes etc. These factors combine to give observed excess mortality in people with diabetes (Dorman et al 1984; Panzram 1987). The severity of some of these problems has been linked to the degree of blood glucose control in the preceding 10-20 years (Pirart 1978; Rand et al 1985). The aims of diabetes care are therefore to control blood glucose levels , to minimise immediate and long term risks, and treat (or oversee) complications as they arise. On this latter point it has been estimated that early detection of retinopathy followed by laser treatment can prevent blindness (Kohner & Barry 1984), and that intensive management of foot problems can reduce the amputation rate by 50% (Edmonds et al 1986).

The control of blood glucose is achieved by a controlled diet, the use of certain tablets, or through injections of exogenous insulin. In addition patient education in how best to manage their condition is important. In general the least biologically disruptive regimes are preferred (ie diet to tablets to insulin).

Most diabetes patients fall into one of two categories; Type I or insulin dependent diabetes (IDDM) typically appears during adolescence and requires insulin treatment from the start and for the rest of the patient's life. Type II or late onset diabetes occurs mainly over the age of 35 and is characterised by an increasing imbalance in blood glucose (the definition of when patients become diabetic is critical). The treatment typically starts with dietary control only (if possible) moving onto tablets and then insulin only when necessary. The largest single group are the non-insulin treated patients (NIDDMs) with type II diabetes.

During the past few years the preferred locus of care has been shifting away from hospitals towards more convenient and accessible community based clinics or with GPs (Day & Spathis 1988, Wood 1990). It is generally accepted that if the quality of care can be maintained then such forms of provision are preferable for the bulk of the basic care of patients with diabetes.

B. Data set and Data Collection

Identifying outcome indicators

The diabetes study required the differentiation of the key patient types. Though the actual outcome indicators across these types are similar, the typical results that one might expect would be very different.

The key patient groups were defined as:

Type I Diabetes - Newly diagnosed
- Historical caseload

Type II Diabetes - Newly diagnosed
- Historical caseload
- Treatment by insulin
- Treatment by tablets
- Treatment by diet only

The pattern of outcomes expected from newly diagnosed patients is different from those who have been diagnosed for some time. The data collection and analysis for these patients was therefore different from that used to examine the long term changes in the returning clinical population ie the historical caseload.

The analyses in diabetes therefore split into two main groups.

1. Newly diagnosed/referred patients - required new data collection systems to monitor changes during the first year of treatment at the Freeman Hospital. The study specifically examined non-insulin treated diabetes. Most patients were newly diagnosed though others were referrals from elsewhere or patients whose contact with the Freeman (or other hospitals) had lapsed. It was agreed to include all these cases as representing a group about to start a new treatment regime at the Freeman. These were followed for the year after their first appointment.

2. Data on patients from the historical caseload were drawn from the system of annual reviews that was already in place at the Freeman Hospital. This scheme attempts to see all patients at least once a year for a battery of checks. Data on these visits has been collected on microcomputer since 1984. Patients in this part of the study included those with Type I diabetes as well as Type II cases treated by insulin, tablets or diet alone. The results for these separate patient groups are treated separately.

The Outcome Indicators

Table 6.1 shows the outcome indicators and the agreed standards. The indicators fall into four groups described in detail below.

Indicators of metabolic control

eg HbA1, blood lipids, body mass index, number of inpatient admissions for destabilisation.

These indicators are predictive of longer term health problems. They are relatively easy to measure and control in the short term. It is arguable whether some of these measures are process or outcome. In particular they may not necessarily reflect ill health from the patients perspective but are rather indicative of the long term problems (which are also included). At present the belief is that control of these factors is significantly related to long term health, although the proof may not be certain. However it is clear that they represent key objectives in the current treatment and control regimes in diabetes. It is for this reason that they have been included and in practice proved to be the most valuable indicators.

1. HbA1 (glycosylated haemoglobin) assays have developed since the late 1970's (Nathan et al 1984) as a way of assessing mean blood glucose concentration during the preceding 6-8 weeks. It therefore proved a better measure of blood glucose control than the blood glucose levels which were subject to wide fluctuations. It has now become so widely adopted that *"many physicians could not imagine treating diabetes without it"* (Nathan 1990). One recent study (Larsen,Horder & Morgensen 1990) has observed the effects of routine monitoring of HbA1 itself and concluded that compared to a control group regular measurement leads to changes in diabetic treatment and improved metabolic control. As such its value is not only as a guide to treatment but also in predicting when long term complications may result from extended periods with high blood glucose (Klein et al 1984). For example one recent study (McCance et al 1989) has shown that mean HbA1 levels were correlated with an increasing risk of developing proliferative retinopathy in insulin dependant diabetes - independent of the time since diagnosis. Similarly a relationship has been shown between mean HbA1 levels and the progression towards

kidney failure (Feldt-Rasmussen et al 1986; Morgensen 1988).

Table 6.1 Examples of the Outcome Indicators at the Freeman

Outcome Indicator	Standards	Data collected
Metabolic control		
HbA1	>8.75% ="Unsatisfactory" >10.0% ="Poor"	At diagnosis, 3 months and subsequent annual reviews
Body Mass Index	>27.5 kgm^{-2} = "Unsatisfactory"	At diagnosis and 12 months, and annual review
HDL Cholesterol	<0.9 mmol l^{-1} ="Unsatisfactory"	At annual review
Cholesterol	>6.5 mmol l^{-1} ="Unsatisfactory"	At diagnosis (3 months if high) and subsequent annual review
Complications of diabetes		
Blood Pressure	<160/90 or 95th percentiles for age <40	At diagnosis, three months and annual review
Circulatory problems	Absent - Score as abnormal pulses, claudication, ulcer or amputation.	At annual review.
Eyes - Abnormal fundi Maculopathy Proliferative retinopathy	Absent	At annual review
Neuropathy -Neuropathic symptoms -Biothesiometer	Absent - neuropathic symptoms Greater than 25 - both legs	At annual review
Kidney complications -Creatinine -Microalbumin	Absent - defined as present when >125 micromol l^{-1} >10.0 mg $mmol^{-1}$	At annual review
Visual Acuity	Better than 6/12 or 6/36 (worst eye)	At annual review
Behaviour /Knowledge		
Smoking	None	At annual review
Knowledge of diabetes	Improvement on test score	Simple questionnaire on diagnosis and after education programme
Well being		
Health status	Maintain/improve SIP score	At diagnosis, 3 months and 12 months

At the Freeman Hospital HbA1 levels have been measured for many years and are considered perhaps the most important indicator of control. Assays are typically performed at diagnosis, annual reviews and often in-between. It was possible to agree beforehand categories of HbA1 that could be deemed acceptable, unsatisfactory and poor. These values were based on international consensus values (Alberti & Gries 1988).

2. *Cholesterol Levels* are important as a risk factor in cardiovascular disease. A number of related assays are used in assessing such patient risks. Thus high density lipoprotein (HDL) cholesterol and triglycerides are all measured although the total serum cholesterol was considered the most important indicator.

There is little doubt in the literature that elevated serum cholesterol (along with hypertension and smoking) is a major risk factor in ischaemic heart disease - a condition with an increased prevalence and risk with diabetes (Abbott et al 1988; Betteridge 1989). Control of cholesterol levels is possible with a combination of diet and drugs (Winocour & Laker 1990).

Although there has been some controversy over the value of screening for hypercholesterolaemia in the general population, screening in diabetes is generally accepted. At the Freeman measurements are routinely taken on diagnosis and at annual review. If high cholesterol levels are observed intermediate readings are also recommended.

3. *Body Mass Index (BMI)* was chosen as the simplest measure to relate weight to height - it is calculated as the weight in kilogrammes divided by height in metres squared (kg m^{-2}). As such relative obesity can be assessed in a single statistic for patients of different height. Patients with values over 25 kg m^{-2} are usually considered 'overweight' whilst a BMI over 30 kg m^{-2} indicates 'obesity' (Bray 1978). In addition some data on absolute changes in weight have also been reported. The distribution of BMI values for non-insulin dependent diabetes is higher than in the general population. Changes in BMI from high to acceptable are an indication that diet regimes are being successful. Once again the standard values were taken from the literature (Alberti & Gries 1988).

4. *Incidence of hypo/hyperglycaemia* is also considered as a monitor of the degree of control. The belief is that under the correct treatment regime patients should not experience periods when blood glucose levels are sufficiently high or low to cause major symptoms. In the case of insulin dependent diabetes an episode of severe hypoglycaemia and resulting ketoacidosis can be very serious and even fatal. However more frequent episodes of lesser severity are likely to occur, thus the definition in this case required

some form of third party assistance or advice eg relative, GP help. As the study of new patients was based on non-insulin treated cases, the incidence of these episodes at follow-up was very low. As a result a valid analysis of the data could not be performed.

Indicators of complications of diabetes

These indicators cover the important health problems that exist within a diabetic population. They may only become manifest after a number of years of poor control and are not always amenable to immediate treatment. The incidence of these problems may reflect practice many years before. In many cases these complications can be inter-related and represent different manifestations of what is essentially the same underlying disease process. Nevertheless their treatment takes different forms concerning the relief of symptoms and patient problems.

These complications can be considered under a number of headings. With *symptomatic health problems* (eg the circulatory problems of claudication, ulcers and amputation, angina and neuropathic symptoms) goals are to minimise the effects of the disease on the patient. *Risk Factors* such as high blood pressure may reflect current health problems for the patients as well as increasing the risks of more serious problems later. Others are *markers of damage* (eg absent pulses) which may become serious risks to immediate health (eg leg amputation).

Blood Pressure, when high, is considered undesirable both as a problem in itself with effects on patient's health, and as a significant risk factor in other diabetic complications. (Diabetes Drafting Group 1985). In one study it was found that 40% of men and 53% of women in a newly presenting NIDDM population had hypertension by WHO criteria (Turner et al 1985). The measurement of blood pressure is subject to some unreliability; high blood pressure is confirmed only after at least two high measurements (Hope & Longmore 1984). Blood pressure measurements are routinely made at the outpatient appointments at the Freeman. For the purposes of outcome monitoring these were categorised according to WHO criteria or centile charts (Drury and Tarn 1985). Any patients over 160/90 was considered 'high', for the remainder the criteria were based on age specific 95% percentiles for a normal population. Concerted long term treatments

are available to reduce blood pressure and should when applied result in a reduction in the proportion of patients with high values.

Neuropathy is recognised as one of the long term problems affecting around 20% of diabetes patients (Nabarro 1988). As well as loss of feeling in some limbs or pain and tenderness , it is associated with foot ulceration (The Lancet 1990). Definitions of neuropathy tend to be rather vague (Laing & Williams 1989). Two approaches were adopted for this analysis, one was based on the clinical assessment of the interviewing doctor about whether neuropathic symptoms were present. The other test was based on a quantitative assessment of vibration sensation using a biothesiometer.

Retinopathy in various forms represents one of the most significant, and avoidable, complications of diabetes. One study (Foulds et al 1983) estimated that 50% of a diabetic population would eventually need laser treatment for retinopathy. Rohan and colleagues (Rohan,Frost & Wald 1989) suggested that the incidence of diabetic retinopathy will produce blindness in 184 men and 276 women in England and Wales in one year. The treatment of retinopathy is effective if the condition is caught early enough (British Multicentre Study 1984). There is also evidence to suggest that screening programmes are cost-effective (Dasbach et al 1991).

Diabetic retinopathy evolves through a series of stages from background (early) retinopathy to proliferative retinopathy or maculopathy either of which can lead to blindness. Diabetes is also associated with a high incidence of cataracts. The assessment of impaired vision was split into a number of stages. In the first instance the standard Snellen charts (routine at diagnosis and annual review) were used to score for impaired visual acuity (corrected with glasses if necessary). In this case the worst eye was selected (thus reflecting the degree of disability) and scored into one of three basic categories as follows:-

Visual Acuity:

Ok = Better than 6/12.

Impaired = 6/12 to 6/36

Seriously impaired = Worse than 6/36

In addition a routine examination using an ophthalmoscope or special camera recorded any evidence of the development of retinopathy and certain other specific features. These were summarised into three categories categorised as follows:-

- Background retinopathy
- Proliferative retinopathy
- Maculopathy

Nephropathy is a complication of diabetes that mainly affects IDDM's and is related to the duration of diabetes. It has been found to affect (in its earliest form) 6.8% of one sample of diabetes patients (Gattling, Mulle and Hill 1986) and in another study renal failure was found to be the largest single cause of death (45%) amongst IDDMs (Dorman et al 1984). It has also been observed that mortality from cardiovascular disease in diabetic patients is forty times higher in those patients who develop clinical nephropathy (Borch-Johnson & Kriener 1987).

The first sign of deteriorating kidney function is the presence of albumen in the urine which can be detected with the simple albustix test, or the more sensitive and sophisticated micro-albumin levels. In addition raised serum creatinine levels may be observed. These are all signs of impaired function and good reasons to hand over to a nephrologist for medical therapy. In this study micro-albumin levels were regarded as the preferred indicator of kidney malfunction predicting major problems 10-15 years in advance. However the assay is relatively new and is only just becoming routine, therefore the routine positive albustix and creatinine levels (>125 micromols l^{-1}) were used as indicating nephropathy.

Cardiovascular Disease can take a variety of forms and represents one of the most common complications in diabetes. Ischaemic heart disease is the most frequent cause of death in NIDDMs (Panzram 1987). In this study the presence of *angina*, as noted by the interviewing physician, was used as an indicator of outcome. This is a potentially treatable complaint and one that can be controlled by medication, angioplasty or surgery. Within the diabetic population the incidence of cardiovascular disease should be improved with control of the relevant risk factors.

In addition to coronary artery problems, peripheral vascular disease may become manifest. For the purposes of outcome measurement the development of circulatory problems in patients have been combined into an ordinal scale of increasing significance. This scale is based on a combination of clinical observations that are routine. At the lowest level is the absence of distal pulses (any limb) progressing to the presence of claudication, ulceration and then amputation. A patient will be classified according to the most severe problem present.

Indicators of general well-being

In addition to the clinical or medical considerations, the extent to which diabetes and associated problems affects (or hopefully does not affect) their daily life is important to patients. There has been relatively little work done on measures of general well-being amongst diabetic populations. Some studies have used existing instruments developed in psychiatry (Wilkinson et al 1988) to examine well-being amongst a diabetic population. Other studies have opted to look at specific symptoms/problems of the diabetic patient (Bulpitt et al 1976). The DCCT trial, looking at IDDMs, developed a composite questionnaire using instruments taken from the literature based on a number of dimensions of quality of life, satisfaction, impact, diabetes worry and social/vocational worry (DCCT Research Group 1988). Other than reporting satisfactory internal consistency and reliability for these tests there was little concluded about their validity. For NIDDMs, a series of scales to measure depression, anxiety, general well-being and treatment satisfaction have been developed (Lewis et al 1988; Lewis et al 1989; Bradley & Lewis 1990). These were purposely designed to divorce the somatic problems of diabetes (or its complications) from psychological status - as far as possible. The resulting measures are shown to be internally reliable and not to be correlated with somatic problems. The same group has also developed scales to examine health-beliefs amongst tablet-treated diabetics (Lewis et al 1990).

For our purposes the specific constructs of depression and anxiety were less important than an appreciation of overall well-being. It was also felt that interactions between a chosen scale and somatic problems of diabetes would be useful. Thus the desired scale

could be validated by the extent to which they agree with some general notions of health - typically represented by their clinical manifestations. The measure would also have to be applicable to IDDMs and NIDDMs. The relative value of different types of scales has been examined in a cross sectional study as part of this project (see Appendix 6).

At the start of the study a decision was taken to use the Sickness Impact Profile (SIP) (Bergner, Bobbitt & Carter 1981). This was chosen because it was a well-documented and widely used instrument which seemed appropriate for this type of patient - although no actual studies had been performed. It was felt that other general health status measures ran the risk of being too insensitive for a population of patients that contained few disabling health problems. The full SIP consists of 136 statements grouped into twelve categories (which can themselves be grouped into dimensions covering physical function and psychosocial function). Each item has a weight and results can be expressed as the weighted sum of responses within each category or combined into a single score on the range 0 to 100%. The profile can be administered as a self-completed questionnaire.

It was felt that the whole questionnaire would be too long and too impractical to use. Therefore four independent categories of the Sickness Impact Profile (SIP) were selected based on a pilot study of 50 patients. The categories were selected because they were considered representative of the whole instrument (ie one physical,one psychosocial and two independent categories) and because they were found to elicit enough positive responses to make analysis possible. The sensitivity of this instrument to changes within the population under review was unknown at the start.

Indicators of education/behaviour/compliance

eg Giving up smoking, improved knowledge of diabetes

There are a number of areas in which the knowledge and behaviour of the patient can help them in their daily lives and can help forestall future health problems. Enhanced knowledge of diabetes and limited changes in behaviour can be considered as legitimate goals of the health service. A number of measures were considered as possible candidates. In the end the indicators used were based on smoking and knowledge of

diabetes. Work has been done at Charing Cross Hospital to develop knowledge questionnaires for Type I diabetes (Meadows et al 1988). This was used as the basis for the knowledge test in this present study.

Timing of observations

For newly diagnosed/referred patients the critical time periods were considered to be three months and one year after diagnosis (or first appointment). During this period one expects stabilisation of the patient and minimal complications to emerge. After a year the treatment regime will have been identified. Subsequent analysis is based on annual review of the patient thus looking at year on year changes. After one year newly diagnosed patients effectively join the larger group of patients representing the historical caseload of the clinic.

Other data items

In addition to the data required to examine outcome indicators some basic information was required to classify the patient. Most importantly age, sex, time of diagnosis (and hence age at diagnosis) and treatment regimes were collected. The dataset for the study of changes in the historical caseload of patients was based on a subset of data collected during the annual review process. The details of the micro-computer systems are shown in Appendix 6.

Data collection

The collection of data was divided into two main parts, with data for newly diagnosed cases requiring new collection systems while much of the historic data exploited existing information systems.

Newly diagnosed/referred patients

At the Freeman Hospital, all new patients are concentrated into one out-patient session which made identification of patients relatively easy. A quick scan of the list usually

identified the few patients who were not eligible, say for example if they were already insulin-treated patients. The eligible patients were given questionnaires (by clinic staff, with project staff 'standing in' when necessary) whilst waiting in outpatients. At the end of the clinic these forms were collected for data input. In addition the clinical features of the patient were taken from the notes and transcribed onto a basic pro-forma. This happened about 1-2 weeks after the outpatient session, allowing time for laboratory results to return. If the notes from one clinic were still together this process was fairly straightforward (approximately 3-4 minutes per patient). If the notes had been 'dispersed' the process took considerably longer. Throughout the project this transcription has been undertaken by a member of the clinical staff.

The three month follow-up was more problematic as the study patients were mixed into a larger clinic population (3-4 study patients in up to 90-100 total appointments per clinic). The process was further complicated as patients were quite likely to have been seen before the three month time period, or after. This meant that the follow-up had to fall within a window of between 2 to 4 months after diagnosis/first appointment. With a non-attendance rate of up to 15-20% this presented problems. The tracking of patients, knowing which clinic they were to re-appear in, was considerably eased by using the outpatient PAS (appointments) computer system which was recently introduced. Once the patient had been identified the forms were left in the clinic in a personally addressed envelope for the clinic staff to deliver and collect. Clinical details were again taken as before 1-2 weeks after the appointment. Given the resources required to carry out this interim follow-up the value of the information was being questioned. The 12 month follow-up of patients was far easier as they fell into the routine clinic system of the annual review process.

Historic Caseload (Annual Review Data)

Data collection for the historic caseload was based on the existing micro-computer information system which had been in place at the hospital for a number of years. All patients were given an annual review which checked a number of basic clinical and laboratory features. This process is recommended as good practice in the care of people with diabetes. The data was recorded on a simple pro-forma (see Appendix 6), some of

which was completed by the doctor when seeing the patient, some completed by specialist nurses, and laboratory results added when they arrived. This data was input into the Metabase system, taking approximately 3 minutes per patients to input.

For this study most of the basic data required to plot the patient's health was collected, but the analysis offered by the Metabase system was too limited. Therefore additional software was developed to extract the collected data and analyse and present the results in a suitable form. This extraction process is now being carried out relatively easily and at regular intervals.

Knowledge questionnaire

The knowledge questionnaires were given out at the first appointment for newly presenting patients. A subset of these patients attended and completed the education programmes organised by the diabetes specialist nurse. On completion the same test was repeated. The test has also been given to patients to complete at their first annual review appointment. The conditions for the completion of these forms was of necessity not ideal - whilst waiting for their appointment. It is possible that this added some variability to the results.

Validation

The bulk of the data collected was a duplicate of existing records. Software for analyzing the historical data base did enable checks to be made on the internal consistency of the data. A number of areas of inconsistency were identified:

- once a patient has an amputation, all subsequent reviews should record that fact. A handful of cases were found where past amputations had disappeared. These were traced to the failure to note 'not observed', as distinct from 'not present' at the time of data collection. These results were useful in encouraging staff to use the forms properly.

- in some cases the recording of 'not assessed', which should have been represented by a nine was not consistent. For example missing HbA1 values

were sometime represented as blanks (which became equivalent to zero), sometimes as 99 or 99.9. To overcome this acceptable ranges were used as a screen to exclude extreme values from the analysis. Where possible the rates of not-recording are also reported. Changes in these values gave a useful indication of how thorough data collection had been over time.

- the data base contained a handful of typographical errors, HbA1 of 8.5 entered as 85. Once again acceptable ranges eliminated these problems, if correction was not possible. Given the large amounts of data there were relatively few problems of this type.

- the calibration of certain tests and assays have changed over the years The outcome reports helped identify these problems. This will be discussed later in the data analysis section.

- for newly diagnosed patients, the tracking necessary for three month and twelve month follow-ups helped reduce the incidence of cases who 'fell out' of the routine appointment system. For example it ensured that all patients who were due for an annual review were given an appointment etc.

Other studies in diabetes

In addition to these two main studies on newly presenting patients and the historical caseload, there have been a number of related issues which have been explored by Freeman staff during this period with either direct or indirect input from project staff.

GP survey.

A survey of diabetic care in a number of GP practices has been carried out (Tunbridge 1991). The data set for the evaluation of care amongst the GP's was based on that agreed for the main hospital studies. Data was specifically collected retrospectively and prospectively on a range of mainly comparable indicators. Results to date were fed back to the GP's concerned with a positive

response. When compared to the hospital it appeared that control of diabetes by GPs was as good if not better than at the Freeman. However it is also clear that the nature of the patients varied between hospital and general practice.

A study of deaths amongst the clinic population

An issue of local concern has been the identification of when patients have died: the hospital did not always know of the fact of death or of the reason. A diabetic research nurse was therefore made responsible for identifying deaths and updating the Freeman database. Though an analysis of the reasons for deaths has not yet been possible the updated database is important in avoiding distressing and unnecessary annual review appointments being sent out.

A study of non-attenders at annual reviews

Once again an issue of local concern has been patients who have not received an annual review appointment. Apart from improvements to the administrative mechanism for identifying when cases are not receiving annual reviews, it was also possible to examine, for a subset of patients, whether basic checks on their health had been conducted during the year and the degree of metabolic control. The results are included in the report shown in Appendix 6.

Cross-sectional study of health status/treatment satisfaction measures.

A fairly large cross-sectional study of three different health status measures was undertaken. The results of this study formed part of the exploratory work undertaken in this area and so have not been presented in the data analysis section. Details are given in a draft paper for publication (Bardsley et al 1991).

Health changes in insulin starters

Towards the end of the study the question of whether transferring patients to insulin treatment yielded the expected benefits in well-being and control was raised. Clinical staff expressed genuine uncertainty over whether sufficient benefits were obtained and a study to examine the issue is being actively pursued.

In-patient study

The diabetes department typically includes a number of inpatient admissions. A study to examine the short term outcome of these patients has begun though as yet the rate of recruitment is small and forms have barely reached re-appraisal after piloting (to date less than 20 patients in four categories have been included). The intention is to continue with this work and provide a mechanism for outcomes monitoring in this area of inpatient care.

C. Data Analysis

This section is divided into two parts - the first deals with the newly referred patients, the second with the historical caseload attending the annual review clinics.
For each section there are three basic questions:

1. Do the outcome indicators change as expected?
2. Are the outcome indicators independent or are they related to each other?
3. Are the outcome indicators sensitive to variations in the presenting characteristics of patients?

Finally a summary section brings together the observations from both patient populations.

Part I - Newly referred patients

Change in Indicators - Newly referred patients

Table 6.2 shows the numbers of cases accrued to August 1990 for which three month follow-up data should be available. Considering first the response rates to three months there is a fairly high proportion of missing cases. The reasons why cases are missing are varied but include:

- deaths
- patients transferred to GP care
- patients not diagnosed as diabetic
- patient not attending out-patients

- patients moving out of the district

The follow-ups at twelve months suffer from similar problems. The relatively small number who have received an annual review is related to the time delays first in inputting data into the annual review system, and then transferring to software developed by tgis study. The results therefore cover only the first few months of patients recruited.

Table 6.2 Response Rates Newly diagnosed Non-insulin Treated diabetes.

No. New patients identified to Aug 1990	= 208
3 month follow-up	= 146 (70.2%)
12 month follow-up	= 62
Linked Annual Review Data	= 35

Changes in metabolic indicators.

The three key indicators of control are shown in table 6.3. For HbA1, cholesterol and BMI there are significant reductions (using paired t-tests) between diagnosis and three months. Further details are given in Tables 6.4(a)-(c) which show the changing number of cases within pre-defined categories of these variables. With BMI and cholesterol levels there is a problem in that observations are not always taken - particularly at the three month follow-up. In the case of cholesterol the policy is that if the reading at diagnosis is normal (acceptable) then repeating the test at three months is not necessary. By and large this appears to be the case as can be seen in Table 6.4(b). This data can also provide a rough check on when tests were performed unnecessarily or not done when they should have been. There appear to be a substantial number of cases with high values at diagnosis yet no follow-up reading.

For HbA1, movement between categories is more common than in the other indicators. It is clear that quite large changes can be observed in certain patients for example 13 patients have gone from one extreme (HbA1 $>10\%$) to the other (HbA1 $<7.5\%$) at three months. Improvements in HbA1 are not restricted to those cases with high values at diagnosis though there is the natural tendency for higher values at the first appointment

to show the greatest fall.

For cholesterol and BMI the results when viewed in these broad categories are less mobile. It is likely that many of the changes explaining the statistical significance in the mean values are occurring in patients initially in the highest category.

Table 6.3 Mean values of key indicators and categories of the Sickness Impact Profile from first appointment to three months and 12 months after first appointment (diagnosis). Figures in brackets are standard deviations where appropriate.

	First appoint. to 3 months			First appt. to 3 and 12 months			
Indicator	No.	Mean First Visit	Mean +3mnths	No.	Mean diag	Mean +3m	Mean +12m Sig vs 3 mnths
HbA1	123	10.5 (2.49)	8.7 (1.81) p=.000	26	10.9 (2.59)	9.6 (2.27) p=.009	8.8 (1.50) ns
Chol	68	7.0 (1.55)	6.7 (1.48) p=.024	14	6.3 (1.21)	6.3 (1.37) ns	6.4 (1.24) (ns)
BMI	123	29.4 (5.39)	28.7 (4.96) p=.000	23	29.8 (6.4)	29.6 (6.03) ns	29.3 (6.23) ns
SIP Ambul	146	17.4	13.9 p=.001	62	17.5	16.6	17.6
Soc Isol		11.8	11.8 ns		11.2	14.5	15.9
RecrPa		19.0	16.9 ns		14.1	21.2	18.6
Hswork		20.2	18.3 ns		16.3	19.6	20.0
Combined		16.0	14.4 p=.023		14.4	17.0	17.6

Significance test for changes from first visit in HbA1, BMI and cholesterol are based on t-test for SIP categories. Wilcoxon rank sums. (ns= p>.05)

Changes to 12 months

For the few patients who have data 12 months after diagnosis, there appeared to be no significant change in the indicators (for better or worse) after three months (Table 6.3). Though there was a change in the means of HbA1 and BMI, in the right direction, these were not significant and the scale of change was less than that observed in the first three months.

Table 6.4 Changes in indicators to 3 months by category

N/R=Not recorded

a. HbA1

First Visit	Count	After three months: N/R	Good <7.5	Acceptable 7.5-8.75	Poor 8.75-10	V.Poor >10	Row Total
N/R			2	1	1	1	5
Good	<7.5	4	12	3	4	1	24
Acceptable	7.5-8.75	6	7	6	1	1	21
Poor	8.75-10		6	7	4	3	20
V. Poor	>10	6	13	13	20	23	75
	Total	16	40	30	30	29	145
	%	11.0	27.6	20.7	20.7	20.0	100.0

b.) Cholesterol

First Visit	Count	After three months: N/R	Good <5.25	Poor 5.25-6.5	V Poor >6.5	Row Total
N/R		5		2	1	7
Good	<5.25	21	4			25
Poor	5.25-6.5	29	4	17	6	56
V Poor	>6.5	19	1	8	28	56
	Total	74	9	27	35	145
	%	51.0	6.2	18.6	24.1	100.0

c.) Body Mass Index

First Visit	Count	After three months: N/R	Good <25	Ok 25-27.5	Poor 27.5-30	V Poor >30	Row Total
N/R		13					13
Good	<25	3	20	2			25
Ok	25-27.5	1	3	18	3		25
Poor	27.5-30	1		8	17	2	28
V Poor	>30	4	1	1	9	39	54
	Total	22	24	29	29	41	145
	%	15.2	16.6	20.0	20.0	28.3	100.0

Changes in Health Status

Fig 6.1 shows the distributions of the four categories of the Sickness Impact Profile at diagnosis together with a combined score. Each category is scored from 0 to 100% - the higher the score the 'sicker' the patient. Summary statistics are shown in Table 6.3. The distributions of all the categories are highly skewed with a high proportion of cases scoring zero. Therefore non-parametric tests have been used to test for statistical

significance. The score for 'Ambulation' shows a significant improvement at three months for the main sample (though not for the subset who are included in the 12 month analysis). All the other categories score slightly lower at three months but these differences are small and not significant. Why the 'Ambulation' category should change and not the others is a little puzzling and requires further exploration. The combined score also shows some differences to three months - though the relationship is not as strong.

This picture is typical of the one that has emerged during the course of the project. The key question is whether the absence of change represents genuine stability in the population - or merely reflects the insensitivity of the instrument. Further evidence on the behaviour of this scale is discussed later. As mentioned earlier a study was undertaken to compare the SIP with two other instruments and the results reported in Appendix 6.

Clinical Problems

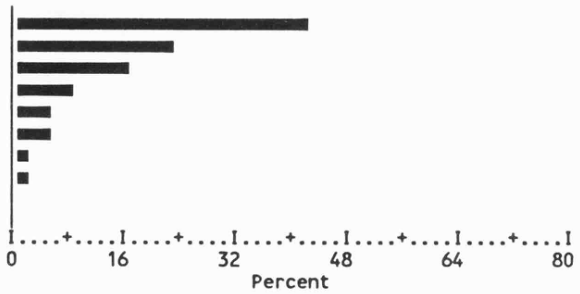
Table 6.5 summarises the prevalence of clinical problems in patients at their first visit. The corresponding values for the cohort of patients who have been linked to annual review data is also shown.

It is clear that by the time of their referral these patients have a number of serious clinical problems. There are some problems with missing data - particularly for the visual acuity but by and large this group exhibits similar prevalence rates as those seen in the full population at annual review as discussed later.

Fig 6.1 Distributions of SIP score categories (n=208)

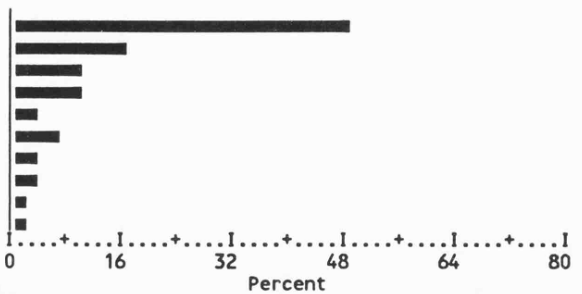
Ambulation

Count	Midpoint
88	5.00
45	15.00
33	25.00
18	35.00
9	45.00
10	55.00
2	65.00
2	75.00
0	85.00
1	95.00



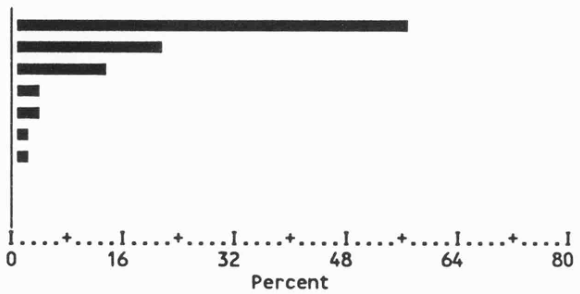
Housework & Hobbies

Count	Midpoint
101	5.00
32	15.00
19	25.00
19	35.00
7	45.00
12	55.00
6	65.00
6	75.00
2	85.00
4	95.00



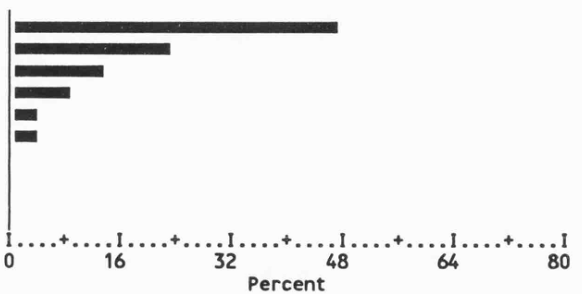
Social Interaction

Count	Midpoint
117	5.00
43	15.00
27	25.00
5	35.00
7	45.00
4	55.00
4	65.00
0	75.00
0	85.00
1	95.00



SIP Combined categories

Count	Midpoint
98	5.00
47	15.00
28	25.00
16	35.00
8	45.00
8	55.00
1	65.00
1	75.00
0	85.00
1	95.00



Recreations & Pastimes

Count	Midpoint
102	5.00
24	15.00
30	25.00
13	35.00
13	45.00
9	55.00
7	65.00
4	75.00
4	85.00
2	95.00

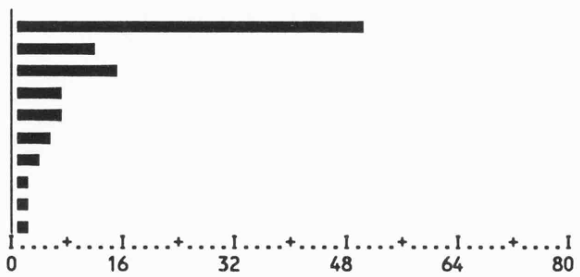


Table 6.5 Clinical Problems of new patients. Percentage of cases with particular clinical problems at first visit for all cases and at first visit and first annual review for a subset of these patients.

N/R=Not Recorded

% cases	At Diagnosis All cases n=207	Cohort At diagnosis n=35	Cohort at Annual Review n=35
BP Ok	58.2	57.1	51.4
High Systol	21.6	28.6	17.1
Diastol	2.4	0	11.4
Both	12.0	8.6	20.0
N/R	5.8	8.6	0
Circulation Ok	69.2	51.4	68.6
Pulses	14.9	37.1	14.3
Claud	12.5	11.4	8.6
Ulcer	2.9	0	5.7
Amput	0.5	0	0
Angina	12.0	8.6	14.3
Creatinine>125	8.7	5.7	8.6
Neuropathic Symp.	26.9	25.7	20.0
Background Ret.	14.9	14.3	25.7
Maculopathy	6.3	11.4	11.4
Prolif. Ret	1.4	2.9	8.6
Any	16.8		
Vis Acuity Ok	63.0	34.3	97.1
6/12 to 6/36	7.7	5.7	2.9
Worse 6/36	3.8	0	0
N/R	25.5	60.0	0
Smoking	27.4	37.1	34.3
Impotence/Amen.	9.6	2.9	14.3

The small number of these new patients where annual review data became available later showed few differences from the overall population although with small numbers it is difficult to draw firm conclusions. The cohort at diagnosis reported rather a higher rate of circulatory problems and smoking, although there was less reported angina and impotence/amenorrhoea. By annual reveiw this cohort exhibits properties closer to the overall new patient population though the incidence of background retinopathy was higher and the rates of impotence/amenorrhoea had increased substantially.

Changes in Education Scores

The education test was given at three separate points, at diagnosis, after the education (programme for those patients who have taken part), and again at the annual review. This enabled the monitoring of changes resulting from the education programme and a comparison with those patients who had not attended the programme. The test is based

on multiple choice questions with multiple correct answers (Meadows et al 1988). Scores are calculated as an adjusted percentage - the proportion of correct responses minus the proportion of incorrect responses. Random completion of the test should give an answer of zero, in some cases the scores were negative ie worse than chance.

Fig 6.2 Changes in education score before and after education programme (n=44)

Scores are expressed as adjusted percentages (% correct responses - % incorrect)

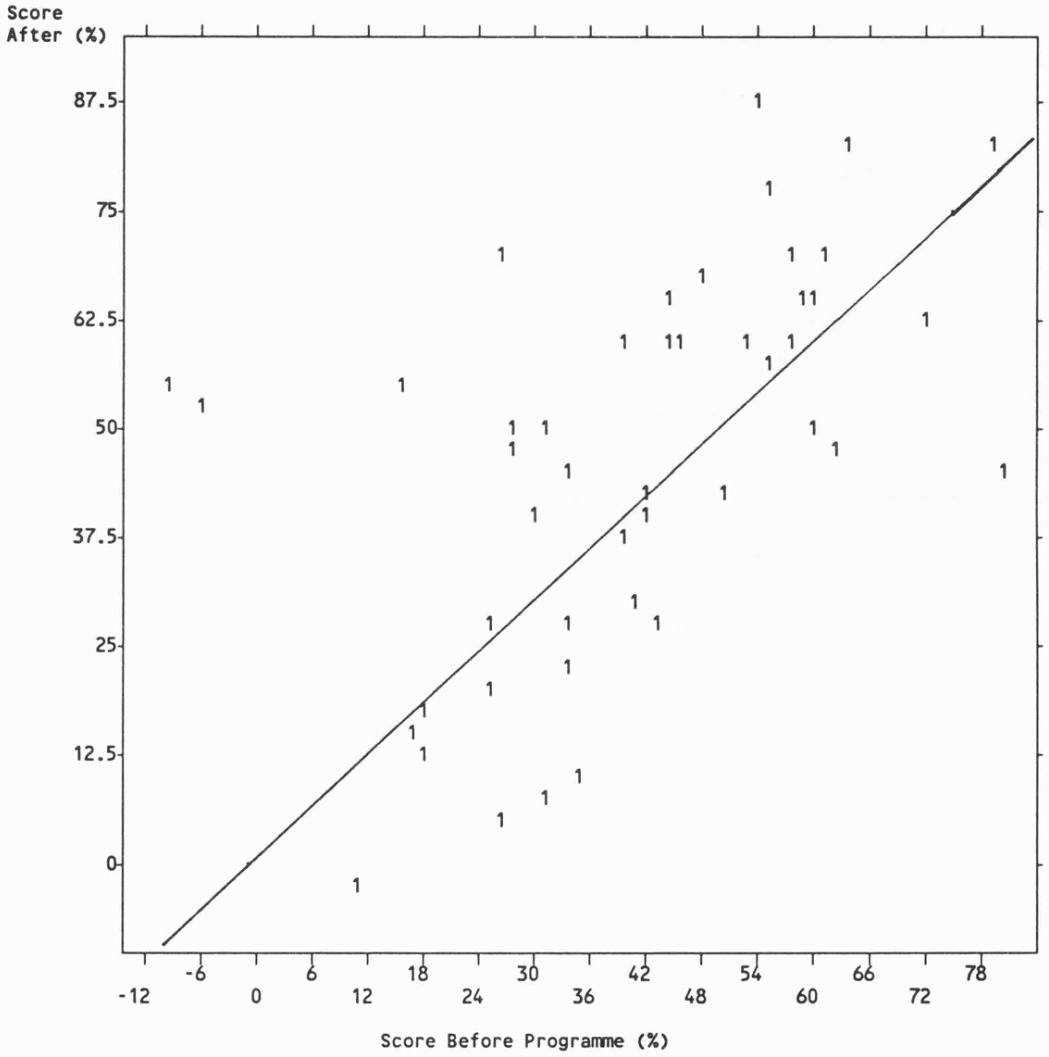


Fig 6.2 plots the scores before and after the educational programme for those patients who have completed the programme. There appears to be a significant improvement in score after the programme though the differences are not large: the mean before was 40.4% and afterwards 46.6% (p=.034). A handful of patients showed large increases in score (over 20%), though typically the increase was of the order of about 10%.

Earlier results had indicated greater improvements amongst those patients who had low scores at the start, however this later analysis does not support this - if anything a number of patients who had low scores at the start appear to score slightly worse after the programme.

For the subsets of patients where data are available one year after their first visit it appears that whilst patients who have attended the education programme show an improvement in mean scores, those cases who did not attend show no improvement (Table 6.6). The differences in change in score between these two samples is statistically signficant.

Table 6.6 Effects of education programme on education score

Mean Score % (Std Dev)	No.	Mean at start	Mean at finish
Without education programme	17	35.5 (21.0)	34.0 (17.5)
With education programme	25	38.9 (18.6)	41.6 (20.9)

The education scores themselves have been compared to a number of other variables present at the first visit. In general there was little relationship between the clinical problems or levels of metabolic indicators and the education score. One exception to this was that the few cases of patients who have abnormal fundi score slightly better. The explanation for this is probably that these patients have been diagnosed for a longer time.

Overall it does appear the newly diagnosed patients scored worse than those patients who have been diagnosed a number of years earlier which one might expect. In fact the difference between the scores was quite large and significant with the newly diagnosed cases scoring 38.3% against 47.9% ($p=.006$) for patients who had been diagnosed for at least one year. This gives an indication of the ranges of score that are likely using this test. It appears that based on this sample the differences between the mean score on this test will never be very large.

Are indicators related to each other - New patients

Correlations between metabolic indicators - New patients

Table 6.7 shows the correlation between the three metabolic indicators and changes in those indicators at three months. It is clear that the only significant correlations are between the indicators and subsequent change in that indicator. These correlations are very strong and represent the 'regression to the mean' effect that is likely to occur. That is the larger the initial value the greater the amount of change. Beyond this there is little evidence that high values on one of these metabolic indicators are necessarily associated with high values on the others or that changes in one indicator are linked to changes in another. It should be noted that the numbers are relatively small but for this population of new patients it appears these three indicators represent quite distinct and independant changes in patients. A similar analysis (later) for annual review cases reported later in this chapter reveals a rather different picture.

Table 6.7 Pearson correlation coefficients for metabolic indicators at diagnosis and subsequent changes in those indicators at three months later (n=68).

Correlations:	HBA1	Change HbA1	BMI	Change BMI	Chol-esterol	Change Chol.
HBA1	1.0000	.5678**	-.1680	-.3124	.2262	.0828
Change HbA1	.5678**	1.0000	-.0844	-.1127	.0783	.1242
BMI	-.1680	-.0844	1.0000	.6284**	.0913	-.0095
Change BMI	-.3124	-.1127	.6284**	1.0000	-.0330	.0056
CHOL	.2262	.0783	.0913	-.0330	1.0000	.4279**
Change Chol	.0828	.1242	-.0095	.0056	.4279**	1.0000

1-tailed significance test. * = (p<.01) ** = (p<.001)

Links between metabolic indicators and clinical problems

The significance of relationships between the three metabolic indicators and clinical problems of presenting patients are later contrasted with results from annual review cases in Table 6.16(a). In the new patient population there appeared to be only one strong positive link - between BMI and high blood pressure, though there were a few other possible links especially between raised blood pressure and the metabolic indicators. Once again it would appear that in this population patients with high values on the metabolic indicators do not necessarily exhibit clinical problems as well. Although the metabolic indicators may be predictive of longer term problems in the future, they were

not necessarily related to problems in the present.

Links between SIP scores and other indicators

Table 6.8 summarises the relationships between observed SIP scores and changes in total SIP (to three months) and the clinical indicators. The results from this sample can be contrasted with the separate study reported in Appendix 6 where a larger cross sectional sample gives slightly more conclusive results.

Table 6.8 Links between clinical indicators and SIP. Mean scores of SIP categories for patients with particular clincal problems (n=208).
Kruskal-Wallis one way analysis of variance with ranks. Where p is not shown then p>.05.

Clin problem		Ambulation	Soc Inter	Housewk	Rec+Past	Tot	Change SIP
BP high	Ok	16.8	13.0	19.8	21.3	16.5	1.7
	Sys	16.2	12.6	20.9	16.1	15.8	3.0
	Sys+Dias	20.5	11.6	18.9	18.1	16.4	2.4
Circul	Ok	14.2	13.3	18.9	19.4	15.5	1.7
	Pulses	22.9	11.7	23.0	22.3	18.5	0.1
	Claud	23.3 p=.029	11.8	20.5	17.2	17.5	3.5
Retinop	None	16.4	12.5	19.6	19.6	15.9	1.6
	Any	17.9	13.1	20.4	18.6	16.6	1.8
Visual Acuity	Ok	18.0	13.5	21.3	19.6	17.1	1.9
	6/12-6/36	14.7	6.8	23.2	18.8	13.9	2.8
	Worse 6/36	25.0	11.8	19.4	18.8	18.0	4.6
Angina	No	15.4	11.9	18.2	19.1	15.1	1.6
	Yes	26.0 p=.006	17.3	30.6 p=.017	22.4	23.2 p=.022	1.6
Creatinine	<125	16.0	12.1	18.7	19.1	15.4	1.8
	>125	23.4	18.0	29.7 p=.071	22.0	22.4	0.9
Neurop. symptoms	No	16.0	11.5	19.1	20.2	15.4	2.5
	Yes	18.4	15.6	21.3	17.5	17.8	-0.7

The picture for new patients was that scores for 'Ambulation' and the combined score showed up the types of relationship that one might expect. The greater the clinical problems of the patient, the higher the SIP scores. Circulatory problems and angina were associated with higher scores. The same pattern of mean scores was also seen in patients with neuropathy, raised creatinine and abnormal fundi though the differences were small and not significant. There were no significant differences between the change in score (to 3 months) and any of these clinical problems.

It should be noted that a cross sectional study of patients showed a much clearer link

between SIP scores and particular clinical problems (Bardsley et al 1991). The differences from this sample may be due to smaller numbers or to differences related to the population of newly presenting patients. Further analyses of these scores are being considered.

Effects of presenting characteristics - New patients.

The main differences in the presenting characteristics of patients that were examined were age, sex, time since original diagnosis and source of referral.

Age, Sex and time since diagnosis on SIP and metabolic indicators

Table 6.9 shows the mean scores of the SIP categories as well as HbA1, BMI and cholesterol according to sex and age category and time since original diagnosis.

Table 6.9(a) Mean values for SIP categories, HbA1, BNI and Cholesterol according to age, sex and time since original diagnosis in newly referred patients (n=208).

Significance for SIP is based on Kruskal-Wallis, for others ANOVAR.

	Ambulation	Soc Isol	Housew k	Rec+P ast	Tot	HbA1	BMI	Chol
Sex Male	14.2	13.0	15.6	17.6	14.4	10.1	28.2	6.1
Female	19.8	12.1	24.5	21.3	17.9	10.7	29.7	6.9
	p=.0012	ns	p=.004	ns	p=.041	ns	ns	p=.001
Age <40	11.6	13.4	22.6	20.9	15.8	8.8	33.0	6.2
50	10.9	11.8	14.0	18.0	12.8	9.7	29.7	6.4
60	21.5	14.6	23.7	18.2	18.9	11.0	29.6	6.9
70	21.3	12.1	23.1	20.1	18.0	10.1	28.0	6.4
>80	26.2	10.7	25.8	30.1	20.6	11.6	27.3	5.8
	p=.0017	ns	ns	ns	ns	p=.005	ns	ns
PrevDx >10yrs	17.4	12.5	22.6	19.5	16.8	10.1	6.1	28.1
<10yrs	17.1	15.2	19.3	21.0	17.3	9.8	6.4	29.7
New	16.8	12.3	20.1	19.4	16.0	10.5	6.6	28.8
	ns	ns	ns	ns	ns	ns	ns	ns

Table 6.9(b) Changes in HbA1 and Cholesterol according to time since original diagnosis for newly referred patients.
Differences between HbA1 are significant (p=.01), differences in cholesterol are not.

	Number cases	Change HbA1	Change Cholesterol
Diagnosed>10yrs	14	0.88	-.52
Diagnosed<10yrs	16	1.09	.27
Newly diagnosed	89	1.94	.39

Considering first the **SIP scores**, 'Ambulation', 'Housework & hobbies' and the combined score showed significant differences between men and women, with women scoring significantly higher. The tendency for women to score higher has been observed in other health status measures and in other conditions. Only 'Ambulation' showed a significant relationship with age. In this case it was expected that there would be stronger links between the SIP scores and age. The time since original diagnosis showed no consistent pattern with new patients scoring higher on some SIP categories and lower on others. It would therefore appear that the sex of the patient was the most important variable to consider when comparing absolute levels of these SIP scores. There were however no significant relationships between the changes in SIP score and any of these variables.

HbA1 did show differences according to age categories. In particular the older patients tended to have higher initial HbA1 values. There was also observed differences between the mean scores by sex. When aggregate scores for men and women were compared or scores within age categories (not shown), women tended to score higher than men though the differences were not signficant. A similar pattern of results was found in the larger annual review sample and has been reported in the literature (Stickland, Paton & Wales 1984). These results suggest that when age distributions or the balance between the sexes varies then some form of standardisation of HbA1 scores may be necessary in order to make valid comparisons.

Considering the effect of time since diagnosis on HbA1, it appeared that there was no observed difference between the mean values of newly diagnosed cases (HbA1 is 10.5% in Table 6.9(a) against those that have been diagnosed for some time (HbA1 is 10.1% for those cases originally diagnosed over 10 years ago). However it is possible that differences in age between these populations may be masking some effect, those who have been diagnosed for a long time tended to be older.

When the *change* in HbA1 to three months was compared there were marked differences between the newly diagnosed cases and those diagnosed over 10 years ago (Table 6.9(b)). It appeared that the fall in HbAl levels following first attendance at the new patient clinic

was far greater in the cases who are newly diagnosed.

BMI showed no significant differences across age or sex categories though women had slightly higher values than men as did the younger patients.

Cholesterol levels showed quite large and significant differences between men and women - though no differences were related to age. The scale of this difference suggests that in developing comparable data sets some standardisation for sex may be necessary. The pattern of results for newly diagnosed versus previously diagnosed patients (Table 6.9b) is similar to HbA1, though the differences in this small population were not significant, the mean improvement in cholesterol levels was largest in the newly diagnosed cases.

Effects of Age, Sex and time since diagnosis on clinical problems

Table 6.10 looks for relationships between these presenting variables and the various clinical problems reported at the first visit. There were no significant differences between the sexes in the incidence with which these various problems were observed. However the age of the patients did appear to be significantly related - typically the average age of patients with a given clinical problem was higher than those without. These effects were strongest for raised systolic blood pressure, circulatory problems, angina, any retinopathy, and raised creatinine levels. This observation is hardly surprising - older patients will tend to have more clinical problems. It would be interesting to compare the observed increase in the incidence of clinical problems amongst these diabetic patients against a non-diabetic population in an attempt to see if the measures are sufficiently sensitive to show the excess morbidity in these cases that can be related to diabetes.

The comparisons of reported problems in the newly diagnosed patients against those previously diagnosed showed the opposite pattern to that seen for age. Newly diagnosed patients had a significantly lower incidence of retinopathy, neuropathy and impaired visual acuity. In the case of retinopathy the differences between these groups were quite

large and an observed incidence of some form of retinal damage in 60% of those cases diagnosed over 10 years ago, yet new to the Freeman clinic is alarming.

Table 6.10 Clinical problems vs age, sex and time since first diagnosed (n=208)

Difference of age test by ANOVAR or t-test; of sex and previous diagnosis by Chi-squared. (ns = p>.05)

Clinical Problem	%Male	Mean Age	Previous Diagnosis >10yrs n=21	<10yrs n=35	New n=142
BP Ok	58%	61.7	63%	73%	60%
Sys	47%	66.1	18%	18%	24%
Sys+D	44%	60.6	6%	6%	13%
	ns	p=.029			(ns)
Circul Ok	56%	59.5	67%	66%	70%
Pulses	48%	65.3	10%	17%	15%
Calud	50%	66.6	20%	11%	13%
	ns	p=.022			(ns)
Abnormal Ok	55%	60.0	38%	83%	89%
fundi Any	43%	67.8	62%	17%	11%
		p=.009			P=.000
Vis Acuity Ok	50%	60.5	93%	93%	83%
6/12 -6/36	50%	68.2	7%	8%	12%
>6/36	63%	60.6			p=.003
		ns			
Angina No	54%	60.0	81%	83%	90%
Yes	52%	69.5	19%	17%	10%
		p=.003			ns
Creat <=125	54%	60.7	95%	88%	92%
>125	50%	70.5	5%	12%	8%
		p=.0112			ns
Neurop No	54%	60.4	52%	57%	80%
symps Yes	52%	63.9	48%	43%	20%
		ns			p=.006

Differences by source of referral

The various routes by which patients can arrive at the Freeman new patient clinics were classified into three groups. The two largest groups were direct referrals from GPs and referrals within the Freeman Hospitals. Particular specialties dominate internal referrals eg cardiology, urology. Within hospital referral may be because complications of diabetes have become manifest and treatment of these problems is underway in other specialties.

Alternatively it may be that routine screening of the patient has unearthed a suspicion of diabetes. Ideally these two ought to be separated. Finally there was a group of patients

transferred from care in other hospitals - either because the Freeman was local or because of special requirements of the patient. This data has been examined during the course of the project. In this case the question was whether there were differences between patients referred from within the hospital to those sent by GPs. Table 6.11 shows that mean HbA1 levels were higher in those patients who had come from GPs which suggests that the imbalance in blood glucose control is greater in these cases.

Table 6.11 Mean indicator values and proportion of clinical problems reported by Source of referral.

(ns = p>.05)

	In Freeman n=79	GP n=108	Other n=21
Mean HbA1 (Dx)p=.001	9.4	11.1	10.2
Change p=.024	1.1	2.4	1.6
Chol Dx (ns)	6.4	6.5	6.5
Change (ns)	0.38	0.42	-.37
BMI Dx (ns)	29.3	28.7	28.5
Change			
SIP Ambul	19.6	14.8	15.6 (ns)
Soc Inter	12.5	12.6	12.5 (ns)
Housework	23.0	18.5	13.7 (ns)
Rec+Past	21.1	18.5	18.2 (ns)
Total	17.7	15.2	14.3 (ns)
Change in total	0.0	2.5	2.8 (ns)
Sex % Male (ns)	58.2%	47.2%	66.7%
Mean Age (ns)	60.6	61.6	63.2
Prev. Dx. >10 yrs	10.1%	9.3%	14.3%
(ns) <10 yrs	20.3%	12.0%	28.6%
New	67.1%	73.1%	47.6%
BP Ok (ns)	68.0%	57.0%	61.1%
Circul Ok	69.6%	72.2%	52.4%
(ns) Pulses	15.2%	14.8%	14.3%
Claud	11.4%	9.3%	33.3%
Ulcer	2.5%	3.7%	-
Amput	1.3%	-	-
Eyes Fundi	16.5%	11.1%	28.6%
(ns) Maculop	3.8%	7.4%	9.5%
Prolif	2.5%	0.9%	-
Any (ns)	19.0%	13.0%	28.6%
Angina (ns)	16.5%	9.3%	9.5%
Creatinine >125 (ns)	9.1%	10.0%	5.3%
Neurop Symps (ns)	29.1	22.2%	42.9%
Smoking (ns)	22.8%	30.1%	28.6%
Impotence (ns)	13.9%	5.6%	14.3%

The subsequent changes in HbA1 over the next three months were also slightly better in

the GP group, probably due to their initially higher values.

Otherwise there were no significant differences between the two groups with regard to all the other indicators. Perhaps the most worrying feature of the table was the high proportion of cases referred from GP's with significant clinical problems. The higher rates of angina observed in referrals within the hospital was related to referrals from cardiology. It is interesting to note the slightly higher SIP scores from the in-hospital referrals which may be a reflection of the other problems in these patients.

Part II - Annual Review Cases

Do the indicators change as expected - Annual review cases

As well as information on newly recruited patients, the diabetes project was also able to call upon the existing clinical data base of annual review information stretching back to 1984. This data was analysed and presented in a number of ways that were not possible using the existing system. In this report the main interest is in the longitudinal changes that are observed between successive annual reviews - the time period chosen for this was the latest complete year ie 1988 to 1989. More recent data has been used in reports to clinicians including the first six months of the 1989-1990 cycle (Appendix 6). The data covers all four patient types and results for these are shown seperately. Later analyses tended to concentrate on the largest group of patients, those with Type II diabetes - and in some cases excluded insulin takers.

Changes in Laboratory Indicators

Tables 6.12(a) and (b) show the changes in patients between successive annual reviews. The tables are split into five subsections. Table 6.12(a) summarises all cases, while Table 6.12(b) has four sections giving results for the individual treatment types.

Each table shows, for a range of indicators, the number of valid cases available for analysis followed by the number of cases classified as unsatisfactory (Hi) in the first year

(1988) and then the second year (1989). The numbers who have changed between these years are then shown for each of the possible combinations. Beneath each figure is shown the percentage of cases. For example in Table 6.12(a) a total of 483 cases came to both annual reviews in 1988 and 1989. 443 of these (92%) had two valid HbA1 values. Of these 157 (35%) were 'unsatisfactory' in 1988, whilst 213 (48%) were unsatisfactory in 1989. 195 were low at the start and stayed low, 122 were high and stayed high. 35 got 'better' moving from high to low while 91 got 'worse' moving from low to high.

Table 6.12(a) Changes In Lab indicators 1988-1989

Total of all cases = 483

NB: 'Hi'= Unsatisfactory (not always high value)

	Total	No. unsatisf. '88	No. unsatisf. '89	Changes '88 to '89 LoLo	HiLo	LoHi	HiHi
HbA1	443	157	213	195	35	91	122
	92%	35%	48%	44%	8%	21%	28%
Chol HDL	360	138	142	107	35	31	187
	75%	38%	39%	52%	9%	10%	30%
Body Mas	461	175	190	252	19	34	156
	95%	38%	41%	55%	4%	7%	34%
Choleste	441	148	129	256	56	37	92
	91%	34%	29%	58%	13%	8%	21%
Creatini	425	39	50	364	11	22	28
	88%	9%	12%	86%	3%	5%	7%
Triglyce	435	184	164	216	55	35	129
	90%	42%	38%	50%	13%	8%	30%
Microalb	60	1	12	47	1	12	0
	12%	2%	20%	78%	2%	20%	0%

The range of indicators used here is larger than for the newly presenting patients and includes triglycerides, high density lipoprotein (HDL) cholesterol. Though these latter indicators have been presented, they have not formed an important part of discussions and considered as a secondary indicator. The following analyses therefore concentrate on the three main metabolic indicators - HbA1, BMI, and cholesterol. Creatinine is included in this section yet it should more appropriately be grouped with the later indicators of clinical complications.

Table 6.12 (b) Changes in lab indicators 1988-1989

Insulin treated (Age at diagnosis >35 years) Total cases = 69

NB: 'Hi' = Unsatisfactory (not always high value)

	No. unsatisf.			Changes '88 to '89			
	Total	'88	'89	LoLo	HiLo	LoHi	HiHi
HbA1	65	32	40	20	5	13	27
	94%	49%	62%	31%	8%	20%	42%
Chol HDL	45	15	14	12	2	3	28
	65%	33%	31%	62%	7%	4%	27%
Body Mas	66	20	26	38	2	8	18
	96%	30%	39%	58%	3%	12%	27%
Choleste	62	23	19	35	8	4	15
	90%	37%	31%	56%	13%	6%	24%
Creatini	63	7	8	52	3	4	4
	91%	11%	13%	83%	5%	6%	6%
Triglyce	60	27	20	29	11	4	16
	87%	45%	33%	48%	18%	7%	27%
Microalb	9	0	2	7	0	2	0
	13%	0%	22%	78%	0%	22%	0%

Insulin treated (Age at diagnosis <35 years) Total cases = 92

NB: 'Hi' = Unsatisfactory (not always high value)

	No. unsatisf.			Changes '88 to '89			
	Total	'88	'89	LoLo	HiLo	LoHi	HiHi
HbA1	85	44	55	23	7	18	37
	92%	52%	65%	27%	8%	21%	44%
Chol HDL	71	12	11	5	6	7	53
	77%	17%	15%	75%	10%	8%	7%
Body Mas	91	16	16	72	3	3	13
	99%	18%	18%	79%	3%	3%	14%
Choleste	84	16	13	64	7	4	9
	91%	19%	15%	76%	8%	5%	11%
Creatini	80	3	5	73	2	4	1
	87%	4%	6%	91%	3%	5%	1%
Triglyce	83	8	8	69	6	6	2
	90%	10%	10%	83%	7%	7%	2%
Microalb	12	0	1	11	0	1	0
	13%	0%	8%	92%	0%	8%	0%

Table 6.12(b) continued

Treated by tablet **Total cases = 171**

NB. 'Hi' = Unsatisfactory (not always high value)

	No. unsatisf.			Changes '88 to '89			
	Total	'88	'89	LoLo	HiLo	LoHi	HiHi
HbA1	158	60	77	65	16	33	44
	92%	38%	49%	41%	10%	21%	28%
Chol HDL	131	63	66	52	14	11	54
	77%	48%	50%	41%	8%	11%	40%
Body Mas	159	71	71	80	8	8	63
	93%	45%	45%	50%	5%	5%	40%
Choleste	156	59	60	77	19	20	40
	91%	38%	38%	49%	12%	13%	26%
Creatini	150	16	25	124	1	10	15
	88%	11%	17%	83%	1%	7%	10%
Triglyce	153	89	80	52	21	12	68
	89%	58%	52%	34%	14%	8%	44%
Microalb	20	1	4	15	1	4	0
	12%	5%	20%	75%	5%	20%	0%

Treated by diet only **Total cases=151**

NB: 'Hi' = Unsatisfactory (not always high value)

	No. unsatisf.			Changes '88 to '89			
	Total	'88	'89	LoLo	HiLo	LoHi	HiHi
HbA1	135	21	41	87	7	27	14
	89%	16%	30%	64%	5%	20%	10%
Chol HDL	113	48	51	38	13	10	52
	75%	42%	45%	46%	9%	12%	34%
Body Mas	145	68	77	62	6	15	62
	96%	47%	53%	43%	4%	10%	43%
Choleste	139	50	37	80	22	9	28
	92%	36%	27%	58%	16%	6%	20%
Creatini	132	13	12	115	5	4	8
	87%	10%	9%	87%	4%	3%	6%
Triglyce	139	60	56	66	17	13	43
	92%	43%	40%	47%	12%	9%	31%
Microalb	19	0	5	14	0	5	0
	13%	0%	26%	74%	0%	26%	0%

HbA1

The results observed for 1988 to 1989 were rather different from earlier years. Table 6.12(a) illustrates a significant increase in the proportion of patients with HbA1 levels which have been classified as 'unacceptable' ie >8.75. The goals of care are to maintain low levels in those cases low at the start and to reduce levels for patients who were previously high. These figures suggest a net shift for the worse during the year. The shift is observed across all patient types.

When these figures were presented, they surprised the clinicians and caused some concern - and raised questions over whether the assay had remained constant. Further investigation revealed (Home et al 1991) that though laboratory quality controls had been adhered to there had been problems with the calibration of the assay and that an increase of 1% had been observed across all cases in three hospitals served by the same laboratory. It is unlikely that this discovery would have been made otherwise and could have potentially affected the treatment of many cases in the area - ie patients could have been given more aggressive treatments. The assay is currently being changed.

Results from earlier years had presented a more realistic picture. Typically the pattern of HbA1 levels revealed large differences between the various treatment groups as would

Table 6.13 Changes in mean HbA1. Year on year change from 1985 to 1988 as means or percentage classified as Lo (<=8.75%) or Hi (>8.75%)

(ns = p>.05)

	1985-86	1986-87	1987-88
Number	107	170	222
Mean HbA1 Start Finish	8.52 8.68 ns	8.44 8.50 ns	8.44 8.20 p<.05
% Lo to Lo Lo to Hi Hi to Lo Hi to Hi	44 22 11 23	53 13 14 20	52 12 15 22

be expected. Table 6.13 shows that the year on year changes in mean values were typically small and no large changes in the proportion of cases classified as 'acceptable' or 'unacceptable'. Perhaps the most important feature of these earlier results had been the overall stability in this population even in those cases with high HbA1 levels. This

had been a subject for discussion amongst the clinical staff.

Body Mass Index (BMI)

The picture with regard to BMI was of even greater stability between years in these patients. Thus in total only 4% of cases moved from above 27.5 kg m^{-2} to below whilst 7% moved in the opposite direction. In total 34% were classified as high in both years. Once again there were large differences between the patient types - high BMI values being much more of a problem in non-insulin treated cases.

Cholesterol

Total serum cholesterol (as opposed to HDL - high density lipoprotein) showed fewer patients overall in the 'unsatisfactory' category and there was a net improvement - more patients going from high to low than vice versa - particularly in diet treated patients where the net change is almost 10%. Type I patients (Insulin treated and dx < 35 yrs old) showed lower levels than the other groups - as was expected.

Changes in clinical problems

Table 6.14 reports on the observed incidence of clinical problems in 1989 and the changes from 1988. It was originally hoped that it might be possible to determine the rate at which particular problems are likely to emerge during the year. For such rates to be calculated a reasonable number of emerging problems relative to the uncertainties of the measurement and recording system would be needed.

Non-recording of data

The proportion of cases with no observation is noted as has been the case throughout the study. It is recognised that there can be a problem either when basic checks that should be made at annual reviw are not made or, more commonly, if the check has been made it is not reported correctly. The actual level of non-recording in 1989 data was better than that observed in previous years. The exception to this was the biothesiometer measurements where the machine was not available for some time during the year.

Changes in calibration of tests

The observed changes in the proportion of patients recording high biothesiometer readings or positive albustix are far greater than would be expected either by chance or by disease progression amongst the population. In fact both these measurements ran into problems of calibration changes. The effect on observed incidence of clinical problems is quite dramatic and such changes can be expected to present problems in longer term longtitudinal studies such as this. It is interesting to observe that though the unreliability

Table 6.14 Incidence of clinical complications 1988-1989

% incidence in 1989; [% change 1988 to 1989, -ve figures show fall]

Variable	Total	Ins Dx<35	Ins Dx>35	Tablet	Diet	
Number cases =	483	98	83	177	129	
BP ok	67% [3]	85% [1]	66% [8]	64% [0]	57% [7]	
BP Sys	17% [-1]	8% [-5]	19% [-11]	18% [4]	19% [1]	-
BP Dias	5% [0]	3% [3]	4% [5]	5% [0]	9% [-5]	-
BP Both	10% [-2]	1% [3]	11% [-5]	10% [-3]	14% [-4]	-
BP NotRec	1% [1]	3% [-2]	0% [2]	0% [1]	1% [1]	
V Acuity ok	88% [4]	96% [-1]	87% [2]	84% [7]	86% [8]	
6/12 to 6/36	9% [-5]	3% [-1]	11% [-6]	9% [-3]	11% [-9]	-
Worse 6/36	2% [-1]	0% [1]	1% [-1]	4% [-2]	2% [-2]	-
Not Rec	1% [2]	1% [1]	1% [5]	1% [1]	1% [2]	
Circul Ok	70% [4]	92% [0]	70% [2]	63% [1]	60% [11]	
Pulses	15% [-5]	7% [-3]	11% [-2]	19% [-3]	17% [-9]	-
Claudic	11% [0]	0% [2]	12% [-2]	12% [2]	18% [-2]	-
Ulcer	1% [0]	0% [1]	1% [2]	1% [0]	2% [-2]	-
Amputation	3% [0]	1% [0]	6% [0]	3% [0]	3% [2]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Fundi Ok	74% [-1]	62% [5]	75% [-7]	72% [-2]	84% [0]	
Abnormal	26% [1]	38% [-5]	25% [7]	26% [2]	16% [0]	-
Not rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Maculopathy No	92% [1]	94% [2]	89% [4]	88% [-1]	96% [1]	
Yes	8% [-1]	6% [-2]	11% [-4]	10% [1]	4% [-1]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Prolif Retinop No	94% [-1]	95% [-5]	94% [0]	91% [-1]	96% [2]	
Yes	6% [1]	5% [5]	6% [0]	7% [1]	4% [-2]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Angina No	80% [5]	98% [0]	80% [5]	75% [8]	73% [4]	
Yes	20% [-5]	2% [0]	20% [-5]	23% [-8]	27% [-4]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Neur Symps No	76% [7]	88% [2]	66% [7]	74% [6]	74% [12]	
Yes	24% [-7]	12% [-2]	34% [-8]	23% [-6]	26% [-12]	-
Not Rec	0% [0]	0% [0]	0% [1]	1% [0]	0% [1]	
Biothesiom <25	26% [27]	6% [10]	33% [27]	31% [57]	29% [35]	
>25	45% [-8]	68% [4]	37% [-8]	38% [7]	39% [-12]	-
Not Rec	30% [-19]	26% [-14]	30% [-18]	29% [-12]	33% [-23]	
Albustix -ve	93% [-8]	94% [-12]	95% [-17]	90% [-6]	93% [-1]	
+ve	6% [7]	4% [11]	5% [12]	7% [6]	7% [0]	-
Not Rec	0% [2]	1% [2]	0% [5]	0% [1]	0% [1]	
Smoking No	83% [-2]	77% [-3]	86% [-6]	83% [-3]	84% [2]	
Yes	17% [2]	23% [3]	14% [6]	15% [3]	16% [-2]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	

in biothesiometer readings was suspected beforehand, the change in albustix was not. This may potentially lead to different treatments being offered for some patients with important implications for their health and also for resource utilisation.

Significance of changes

Table 6.15 shows the significance of the year on year changes for a selection of indicators. In all cases, for the cohort of patients where two readings were available there was a higher proportion of reported problems at the end of the year. However in all cases there was also a proportion of patients where the problems had 'disappeared'. The scale of observed change during the year (ie the percentage of cases where problems 'appeared' or 'disappeared') when compared to the overall percentage reporting the problem gives some idea of the consistency with which a problem may be present.

Table 6.15 Significance of changes in the prevalence of clinical problems between successive annual review - Non-insulin treated patients

(1988-89; n = 391)

	Prevalence Year1	Prevalence Year2	Change - to+	Change + to -	Signif.test McNemar Chi-sq
Circulation Ok	69.1	64.7	--	--	p=.02
Pulses	11.5	16.6	7.4	3.1	
Claud	13.8	13.6	7.2	6.9	
Ulcer	1.6	1.5	1.1	1.4	
Amput	4.1	3.6	---	1.4	
BPressure High	32.3	36.2	16.7	12.8	ns
Eyes Background	25.8	23.3	6.9	9.5	ns
Maculop	7.7	8.4	4.6	3.8	ns
Prolif	6.1	5.9	3.3	3.6	ns
Angina	17.6	23.8	9.2	3.1	.0009
Creatinine	9.9	12.8	5.2	2.3	.0776
Neuropthaic Symps	18.1	26.1	15.5	7.5	.002
Visual Acuity	5.1	12.0	9.3	2.4	.0002

In the cases of angina, neuropathic symptoms and impaired visual acuity there were significant changes for the worse. Changes in the incidence of retinopathy were not large and it appeared that one year may be too short a time scale to see an effect with this volume of cases. There were relatively large changes in the proportion of cases with high blood pressure yet no net shift for better or worse. It is not possible to tell from the data whether these results reflect genuine changes in the patients or are merely due

to the unreliability of blood pressure measurements.

Are indicators related to each other - Annual Review Cases

Correlations amongst metabolic indicators and changes in indicators

Table 6.16(a) shows the correlation between the metabolic indicators and changes in those indicators between successive annual reviews. The strongest relationships were between change in score and the original scores for cholesterol and HbA1 a feature in common with that seen earlier in the newly diagnosed patients (Table 6.7). Thus those patients with high values in the first year tended to show the greatest change by the subsequent year. This regression to the mean effect is to be expected and it is perhaps surprising that in this case the same effects were not observed for BMI.

Table 6.16(a) Links between metabolic indicators. Pearson correlation coefficients for metabolic indicators and changes in metabolic indicators.

Non-insulin treated cases only, both sexes. n = 302

Correlations:	HBA1	BMI	CHOL	Change HbA1	Change CHOL	Change BMI
HbA1	1.0000	.0903	.1360*	.3932**	.0632	.1248
BMI	.0903	1.0000	.1150	.0129	.0198	.1052
CHOL	.1360*	.1150	1.0000	.0163	.5174**	.0505
Change HbA1	.3932**	.0129	.0163	1.0000	.1049	.0725
Change CHOL	.0632	.0198	.5174**	.1049	1.0000	.1111
Change BMI	.1248	.1052	.0505	.0725	.1111	1.0000

1-tailed significance test. * = p<.01, ** = p<.001

Links between metabolic indicators and clinical problems

Table 6.16(b) shows any links there may be between the clinical problems the patient faced and the absolute values of the metabolic indicators. The clinical rationale behind control of the metabolic indicators is that they will predict future complications. The table shows the significance of the relationship drawn from analysis of variance. Only one link held true for both new patients and annual review cases - that being the positive relationship between high blood pressure and BMI where the correlations were highly significant.

Apart from this there appeared little pattern to the results and little consistency between the two samples. For the annual review sample higher cholesterol levels seem to be

associated with a number of other problems (circulatory, angina, retinopathy).

Table 6.16(b) Significance of links between clinical indicators and laboratory indicators
New = newly recruited patients (n=208); AR = Type II patients in annual review clinic (n=391)

One way ANOVAR. (ns = p>.05)

	HbA1		Chol		BMI	
	New	AR	New	AR	New	AR
Circulation	ns	ns	ns	.0505	ns	ns
Blood Pressure	.039	ns	.043	ns	.000	.0002
Angina	ns	ns	ns	.0093	ns	.0051
Back Retinop	ns	.04	ns	ns	ns	ns
Maculop	ns	ns	ns	ns	ns	ns
Prolif	ns	ns	ns	.0037	ns	ns
Vis Ac	ns	ns	ns	.0525	ns	ns
Creat>125	.009	ns	ns	ns	ns	ns
Neurop symps	ns	ns	ns	ns	ns	.0291

The BMI values were also associated with the presence of angina and neuropathic symptoms. However there was no association between the levels of HbA1 and any of the clinical problems.

There was little evidence of any strong relationships between changes in the clinical indicators and the absolute levels or changes in the metabolic indicators as seen in Table 6.16(c). The only relationships with a probability less than five percent were between patients who develop circulatory problems or high blood pressure during the year having slightly higher average cholesterol levels. Otherwise the evidence suggests a deteriorating clinical condition was largely independent of these metabolic indicators. Or the other way round that short term changes in the metabolic indicators were not linked to the appearance of the various clinical problems.

The absence of relationships between these sets of indicators does not lead to the conclusion that short-term metabolic indicators are not useful predictors of complications in the longer term. The speed at which the various complications develop suggest that complications arise due to poor control 10-15 years earlier and so there is a substantial time lag between high metabolic indicators and the development of these clinical

problems.

Table 6.16(c) Relationships between changes in indicators between annual reviews (n=391).

(ns = p>.05)

		% cases	Mean HbA1	Mean change HbA1	Mean Chole	Mean change Chole	Mean BMI	Mean change BMI
Circulation	Same	82.8%	8.1	-.60	6.0	.07	27.3	-.44
	Worse	17.1%	8.3	-.52	6.4	.13	27.3	-.07
			ns	ns	p=.03	ns		ns
BP	Same	83.4%	7.9	-.79	6.0	.09	26.4	-.45
	Worse	16.6%	7.9	-.57	6.4	.27	27.2	-.53
			ns	ns	p=.04	ns	ns	ns
Eyes	Same	84.8%	8.0	-.75	6.2	.15	27.3	-.29
	Worse	15.2%	8.5	-.39	6.1	.18	27.6	-.29
			ns	ns	ns	ns	ns	ns
Angina	Same	88.5%	8.1	-.65	6.1	.17	26.9	-.25
	Worse	11.5%	8.3	-.52	6.5	.25	28.5	-.44
				ns	ns	ns	p=.09	ns
Neuropathy	Same	82.1%	8.1	-.63	6.2	.13	27.1	-.30
	Worse	17.9%	8.0	-.49	6.1	.12	28.2	-.36
			ns	ns	ns	ns	ns	ns
Creatinine	Same	94.3%	8.1	-.66	6.1	.10	27.1	-.37
	Worse	5.7%	7.5	-.79	6.5	.26	28.9	+.40
			p=.09	ns	ns	ns	ns	ns
Vis Acuity	Same	90.3%	8.1	-.65	6.1	.10	27.2	-.36
	Worse	9.7%	8.2	-.37	6.5	.20	28.5	-.05
			ns	ns	ns	ns	ns	ns

Links between clinical problems

Table 6.17 shows the proportion of cases that attended 1989 annual review with two different clinical problems and the significance of the association between these problems (using Chi-sqaured test). There appeared to be a number of significant pairings of complications where the reported frequency of both problems being present exceeds that expected by chance. Thus for example circulatory problems were associated with angina, background retinopathy, impaired visual acuity, neuropathic symptoms and raised creatinine levels.

The table suggests that there was a tendency for clinical problems to cluster in patients rather than being completely independent. This is generally the picture one might expect in any population. It is also possible that another variable could be explaining these observed correlations, the most likely being age or that they are indicative of a common underlying disease process? Further analysis is needed to address these questions fully.

Table 6.17 Significance of association between pairs of clinical problems (n=391)

(ns = p>.05)

	BP High	Angina	Backgr retino p	Maculo p	Prolif retinop	Visual Acuity	Neurop. symps	Creatin. >125
Prevalence	30.3%	21.7%	22.6%	6.2%	5.2%	6.5%	25.6%	11.3%
Circulation P 15.4%	5.3	4.4	5.0	1.5	1.0	1.6	4.5	3.7
C 14.1%	4.7	5.8	3.5	0.6	1.5	1.4	6.4	2.0
U 1.4%	0.6	0	1.0	0.2	0.2	0.4	0.4	0
A 2.9%	1.2 (ns)	1.0 p=.000	1.0 p=.001	0.4 ns	0.2 ns	0.4 p=.005	1.0 p=.000	0.4 p=.002
Blood Pressure	--	6.8 ns	6.6 ns	2.0 ns	2.1 ns	2.2 ns	10.0 p=.012	3.9 ns
Angina	--	--	5.2 ns	1.0 ns	1.7 ns	1.4 ns	9.9 p=.000	4.3 p=.004
Back Retinop	--	--	--	6.2 --	5.2 --	1.8 ns	7.0 ns	3.5 ns
Maculop	--	--	--	--	1.4 p=.0001	0.6 ns	1.9 ns	1.8 p=.003
Prolif	--	--	--	--	--	0.6 ns	2.1 ns	1.2 ns
Visual Acuity	--	--	--	--	--	--	2.4 ns	1.7 p=.022

Effects of presenting characteristics - Annual review cases

Effects of age and sex on BMI, HbA1, Cholesterol

Table 6.18 summarises the relationships between the three metabolic indicators and age, sex and time since diagnosis. In addition to showing the mean values on these indicators by age and sex categories, the table also shows correlation coefficients between mean variables and the age, the time since diagnosis and the age at diagnosis.

For HbA1 there was a very strong difference between the sexes yet with little effect due to age. Analysis of variance using age as a covariate suggested that there were highly significant ($F=16.2$, $p=.000$) differences between mean HbA1 levels across the sexes and no significant relationship due to age. Women typically scored higher than men.

With regard to BMI levels, there appeared to be no direct relationship between mean BMI and either age or sex or the age at diagnosis. There was an observed negative correlation between time since diagnosis and BMI in men, ie men who have been

diagnosed for a long time tended to have slightly lower average BMI. The mean cholesterol levels for women tended to be slightly higher than for men and there appeared to be no consistent relationship with age. or time since diagnosis. However the age at diagnosis was postively correlated with cholesterol levels in women.

These results suggest that sex is probably the most important variable to adjust for when making comparisons of these metabolic indicators and appears more influential than age, the time since diagnosis or the age at diagnosis. Though it is recognised that age is not necessarily linearly related to these indicators the lack of large observed differences between age groups is a little surprising.

Table 6.18 Effects of age and sex on metabolic indicators (Non-insulin treated patients only) and Pearson correlation coefficients against age and time since diagnosis.

(n = 236 male, 149 female)

Age Group	Mean BMI		Mean Cholesterol		Mean HbA1	
	Male	Female	Male	Female	Male	Female
41-50	27.1	27.1	5.9	5.9	7.8	9.5
51-60	27.4	26.5	6.0	6.7	7.6	8.4
61-70	26.2	29.0	5.8	6.4	8.2	8.5
71-80	26.5	28.2	6.3	6.7	7.2	8.4
>80	27.1	26.6	5.3	6.8	7.1	8.5
Correlation coefficient: vs age vs time since diagnosis vs age at diagnosis	 -.001 -.211 ** -.11	 .110 -.137 .046	 .067 .009 .047	 .193 .035 .199*	 .009 .273 ** -.028	 -.04 .074 -.233**

* = p<.01, ** = p<.001

Effects of age, sex and time since diagnosis on clinical problems.

Table 6.19 shows the relationships between the clinical complications of diabetes and age, sex and time since diagnosis. With regard to sex, it appeared that the proportion of women with proliferative retinopathy, impaired visual acuity and raised creatinine levels is greater than men.

A more consistent picture emerged with respect to age where patients with clinical problems who tend to be older than those without them. These differences are highly

Table 6.19 Clinical problems vs age,sex and time since first diagnosed (n=482)

Differences in age tested by ANOVAR or t-test; in sex and previous diagnosis by Chi-square. (ns = p>.05)

Clinical Problem	%Male	Mean Age	Mean Dx Time
BP Ok	61%	61.5	7.9
Sys	56%	64.2	7.2
Sys+D	ns	p=.04	ns
Circul Ok	59%	59.9 yrs	7.1
Pulses	50%	69.9	9.6
Claud	69%	65.7	7.6
Ulcer	43%	59.9	6.0
Amput	60%	62.1	9.1
	ns	p=.000	p=.009
Back Ab No	58%	61.9	7.3
Yes	63%	64.0	8.8
	ns	ns	p=.014
Macul	58%	62.2	7.6
	75%	64.6	8.6
	ns	ns	ns
Prolif	59%	62.5	7.6
	56%	60.4	8.0
	p=.017		ns
Vis Acuity Ok	60%	61.5	7.6
6/12 -6/36	39%	75.9	7.5
>6/36	p=.005	p=.000	ns
Angina No	59%	61.5	7.7
Yes	57%	65.2	7.5
	ns	p=.012	ns
Creat <=125	59%	61.8	7.6
>125	58%	66.8	8.6
	p=.0004	p=.012	ns
Neurop No	58%	61.8	7.6
symps Yes	62%	63.7	7.8
	ns	ns	

significant for circulatory problems, visual acuity, angina, raised creatinine levels and blood pressure. The increased incidence of these problems with increasing age suggests that some form of standardisation of these indicators would be essential for comparisons.

Similarly the related effects of time since diagnosis showed significant associations with the presence of circulatory problems and background retinopathy. The correlation between time since diagnosis and age was not as strong as one might expect and it is clear in Table 6.19 that time since diagnosis did not behave in the same way as age. For example the mean age of patients with impaired visual acuity was much greater than patients with satisfactory vision, yet there was no differences between the mean time since diagnosis. It may be that impairment in vision was not related to the diabetes.

Table 6.20 examines these relationships in more detail by showing the observed proportion of cases with a particular problem by age category and time since diagnosis category. Though the numbers in the various cells in this table are uneven - for example very few patients are under 40 (or over 80) and diagnosed over 15 years the patterns are interesting. The table illustrates the dramatically increased prevalence of these clinical complications in the older age groups. These gradients are seen particularly clearly in the largest sample of cases who have been diagnosed less than five years where for example circulatory problems rise from 3.4% in the under 40 age band to 73% in the over 80 age band. These age related effects were far greater than the range in prevalance associated with time since diagnosis ie 34% for those diagnosed for under 5 years to 46% in those diagnosed over 15 years ago. The prevalence of high creatinine levels showed a similar pattern with a steeper gradient with respect to age than with increasing time since diagnosis.

For impaired visual acuity, neuropathic symptoms and angina, age appeared to be an important factor but the time since diagnosis less relevant. This may be because the care provided is minimising the incidence of complications from the time of diagnosis. Alternatively it may be related to the relative survival rates of patients with and without clinical problems.

When changes in the under 40 age band were examined there appeared to be slightly steeper increases in the reported prevalence of the various clinical problems over time since diagnosis. For most indicators the reported incidence of problems in those diagnosed within five years is less than that for those diagnosed over five years ago. This can be contrasted with the 50-60 and 60-70 year age bands where the opposite tends to be true, namely that the observed incidence of clinical problems was lower in those patients who have been diagnosed over 5 years than in the newly diagnosed cases. It is possible that in these age groups this differential is a reflection of the effects of the services offered after diagnosis.

Table 6.20 Proportion of cases with observed clinical problems by category of age and time since diagnosis (n=482).
Chi-squared. (ns = p>.05)

	Time since diagnosis			
	< 5 years n=216	5 - 10 years n=181	10 - 15 years n=37	> 15 years n=48
BP				
Age <40 n=56	17.2	18.8	0	0
50-60 n=118	29.3	14.3	42.9	18.2
60-70 n=158	38.5	34.9	15.4	17.6
70-80 n=122	35.8	32.0	14.3	50.0
> 80 n=28	63.6	50.0	0	25.0
Sig. (Chi-Sq)	p=.057	p=.069	ns	ns
All ages	33.8%	28.7%	16.2%	25.0%
Circul				
Age <40 n=56	3.4	18.8	28.6	0
50-60 n=118	22.4	19.0	14.3	27.3
60-70 n=158	32.3	27.0	23.1	35.3
70-80 n=122	50.8	48.0	42.9	75.0
> 80 n=28	72.7	70.0	66.7	100
Sig. (Chi-Sq)	p=.000	p=.0015	ns	p=.006
All ages	32.4%	32.6%	29.7%	45.8%
Eyes				
Age <40 n=56	6.9	25.0	28.6	0
50-60 n=118	24.1	16.7	0	18.2
60-70 n=158	21.5	15.9	23.1	41.2
70-80 n=122	22.6	24.0	57.1	58.3
> 80 n=28	18.2	30.0	33.3	0
Sig. (Chi-Sq)	ns	ns	ns	p=.062
All ages	20.4%	19.9%	27.0%	33.3%
Vis Acuity				
Age <40 n=56	0	0	0	0
50-60 n=118	0	0	0	0
60-70 n=158	4.6	1.6	0	5.9
70-80 n=122	17.0	8.0	14.3	16.7
> 80 n=28	45.5	10.0	33.3	25.0
Sig. (Chi-Sq)	p=.0000	ns	ns	ns
All ages	7.9%	3.3%	5.4%	8.3%
Angina				
Age <40 n=56	0	12.5	42.9	0
50-60 n=118	19.0	23.8	0	18.2
60-70 n=158	26.2	20.6	23.1	17.6
70-80 n=122	30.2	38.0	0	25.0
> 80 n=28	27.3	10.0	0	0
Sig. (Chi-Sq)	p=.022	ns	ns	ns
All ages	21.8%	24.9%	16.2%	16.7%
Neurop. Symps				
Age <40 n=56	6.9	25.0	28.6	25.0
50-60 n=118	20.7	23.8	0	27.3
60-70 n=158	38.5	25.4	30.8	35.3
70-80 n=122	28.8	24.0	42.9	16.7
> 80 n=28	0	0	66.7	25.0
Sig. (Chi-Sq)	p=.003	ns	ns	ns
All ages	25.1%	23.2%	29.7%	27.1%
Creatinine>125				
Age <40 n=56	3.6	0	0	0
50-60 n=118	1.9	12.5	20.0	18.2
60-70 n=158	5.0	10.7	15.4	6.3
70-80 n=122	19.2	23.4	14.3	33.3
> 80 n=28	18.2	0	0	0
Sig. (Chi-Sq)	p=.007	p=.071	ns	ns
All ages	8.3%	13.0%	11.4%	14.9%

Summary - Data Analysis

The choice of outcome indicators

The analysis of outcomes in diabetes has involved a greater number of different indicators than in other specialties. Indeed the volumes of data are so great that some indicators have received only a cursory examination or have been dropped from the analysis either because:-

- they replicated other indicators,
- recording was unreliable,
- the observed incidence was too low, or
- they had less clinical significance.

These indicators included the observed frequency of smoking, impotence or amenorrhoea, lipid measures of triglycerides and HDL cholesterol, frequency of hypoglycaemia and inpatient admissions among newly referred type II patients.

Two other indicators, albustix protein and biothesiometer measures have proved too unreliable in this setting as the calibration of these tests has changed over time and between suppliers of the test - alternatives are therefore used. While not central to this study the questions raised over the reliability of some common diagnostic tests could have very important implications particularly when decisions about treatment critically depend on such results. It is also a waste of resources to collect data which is of no value or unreliable.

The preferred indicators therefore fell into the following categories (as noted in Table 6.1):-

1. Indicators of metabolic control
    a. HbA1
    b. Body Mass Index
    c. Serum cholesterol

2. Indicators of complications of diabetes or risk factors
    a. Raised blood pressure
    b. Impaired visual acuity
    c. Retinopthy - background, proliferative or maculopathy
    d. Circulatory problems, worst of;
        - absent pulses
        - claudication
        - ulcers
        - amputation
    e. Angina
    f. Marker of kidney disease - raised creatinine levels (though micro-albumin would be preferred if available)
    g. Marker of neuropathy - symptoms

3. Improvement/maintenance of general health status

4. Improvement in knowledge of diabetes

All these categories have been applied to examine changes in newly referred patients - and in some form they are all applicable for all diabetes cases.

It is these indicators that form the basis of the following review. Returning to the original questions at the start of the data analysis section therefore:

1. Do the indicators change as expected?
2. Are the indicators related to each other?
3. Are the indicators sensitive to patient characteristics?

1. Metabolic indicators - HbA1, BMI, Cholesterol

Newly referred patients

HbA1, BMI and serum cholesterol showed marked and significant changes in mean

values in the three months after the first visit to the new patient clinic. Though some change for the better is also observed to a year it is of a lesser scale. The effect of the early stages of treatment can be clearly seen in these indicators at three months and one year. Despite this improvement in the population, not all patients fall into the 'satisfactory' category at follow-up.

It appears that the absolute levels of these three indicators are independant of each other and that a patient classified as 'unsatisfactory' on one will not necessarily be 'unsatisfactory' on the others. The observed changes over the first three months in any one indicator are related to the initial value of the indicators - higher initial readings are associated with greater subsequent change. However the scale of such changes are not related *between* indicators.

The values of these variables are largely independent of indicators of complications and knowledge of general health status. The exception to this was the significant link between high BMI and raised blood pressure.

Though the relationships are not strong, the values of these indicators may be influenced by the age and sex distributions of the new patient population - sex seems more important as a standardising variable if comparisons are to be made to other data sets.

Similarly the scale of change in HbA1 and cholesterol is greater for patients who are newly diagnosed when compared to those cases who have been diagnosed for some time. Once again for comparative analyses this variable may be important.

Annual Review Patients

In contrast to the newly referred patients these indicators show little change in mean values between annual reviews (with the exception of spurious changes in the assay). Though there is little net change a proportion of patients (of the order of 10-20%) will move between the classification of 'satisfactory' and 'unsatisfactory' and vice versa. Though the apparent stability of these indicators is seen as disappointing, it has been generally accepted that this reflects practice in the clinic and is not a reflection of the

measures being too insensitive.

As with the newly referred population the indicators are largely independent, though a weak correlation between high HbA1 and high cholesterol is observed. The scale of change between annual reviews is related to the initial values.

With the exception of the BMI-Blood pressure link, the values of these indicators is independent of the clinical complications of diabetes.

Similarly there are intermittent relationships between the values of these indicators and the sex of the patient and the time since diagnosis. Women typically score higher than men, and higher values are found in those who have been diagnosed longer.

Finally there are the expected large differences between the different categories of patients -especially between type I and type II diabetes.

2. Indicators of complications of diabetes or risk factors

The recording of the variety of clinical complications can be beset by a number of problems including:-

- missing data
- problems with the reliability of measures eg blood pressure
- changes in the calibration of tests
- changes/differences in the way problems are defined (for example a large rise in the incidence of retinopathy followed the introduction of a new camera)
- problems of small numbers of affected patients

When used as cross sectional descriptions of a population - even newly referred cases, most of the clinical problems are present in at least five percent of the sample. The exceptions being amputations, ulceration and proliferative retinopathy. The small numbers of these complications make interpretation even more difficult.

In the newly referred population, there were insufficient cases with two data points to draw any firm conclusions about change in the first year following referral. However in the annual review population the numbers are larger and change can be seen between successive years. The largest changes are seen in the increasing incidence of impaired visual acuity, neuropathic symptoms and angina, with lesser changes noted in the circulatory problems and raised creatinine levels. No significant change is observed for blood pressure and retinopathy. It is clear that not all of the apparent change between years is for the worse.

It was hoped that it would be possible to identify an expected annual increase in the incidence of particular complications. Analysis to date has suggested that the changes in successive years are not predictable with this data and further analysis over longer time periods is required.

The indicators when used to describe a population in cross section appeared to behave largely as expected. They tended to be related to age and to general health status. There was also some relationship between the different indicators (possibly linked via age) which would suggest that their appearance in the population was not wholly independent.

3. Improvement/maintenance of general health status

The Sickness Impact Profile has been used to monitor longitudinal change in newly referred patients though the results have been generally disappointing. The distributions of scores were highly skewed, as was feared at the outset, with a high proportion of patients scoring zero. It was therefore not surprising that it was difficult to observe any change in the score. The change that did occur was in the category dealing with 'Ambulation'.

The SIP score was found to be linked to certain clinical problems and in a larger cross sectional study to behave largely as expected with older people and women scoring higher. In practice the measure seemed to work better when identifying the minority of

patients who had severe health problems rather than the day to day effects of diabetes. As such it was more akin to a summary measure dealing with the variety of complications of diabetes and could be expected to change along similar time-scales.

4. Improvement in knowledge of diabetes

The knowledge test has shown itself to be useful in monitoring the subset of patients who attend and complete the educational programme. Though the improvement in scores was not large there was a significant increase in score after the education programme. Moreover there was also a greater increase in scores observed in these patients than in those who did not attend the course.

Scores on the simple test were largely independant of other outcome indicators but they were related to the time since diagnosis as one would hope. Patients who have been diagnosed for a long time score signficantly higher than those who have just been diagnosed. The difference in scores is of the order of 10%. Knowledge scores were not significantly related to age and sex.

D. Review Process

The review of data in diabetes has taken two distinct cycles. At one level, meetings were held with medical colleagues invited from other hospitals as well as a local GP to examine the results. Appendix 6 includes the report circulated to the latter of these meetings.

In addition meetings were held approximately every two months within the Freeman Hospital to discuss progress and the results of the project. The membership of this group included consultants, senior registrars and nurse specialists. The study of patient outcomes itself has proved useful in a number of respects.

1. Data collection/recording

As a result of the study and the ability to actually look at results longitudinally, data that had been collected for some time is now being analysed and used. This made problems in data collection easier to identify and gave a greater incentive to improve data collection and inputting. The envisaged review system in which the annual review data base would be examined every 6 month is now in place. In addition the data helped reinforce some of the basic checks that should be carried out, for example cholesterol levels on patients whose last reading was unacceptably high.

2. Identify patients/booking appointments

The effort put into looking at the exisitng information systems and following patients has led to improvement in the identification of patients and the follow-up of non-attenders. In particular it has been found that during 1989 a high proportion of patients did not receive an annual review. The changes in some key outcome indicators for these patients have been specifically collected and analysed. Similarly problems were found in identifying when patients had died and a specific search was needed to bring the basic data-base up to date. The identification of patients who have died has now become more rigorous.

3. Application of guidelines for junior doctors.

The results from the annual review clinic were worse than expected - the pattern of little change among poorly controlled patients was particularly alarming. As a result an attempt was made to help junior doctors by introducing a system of clinic guidelines.

4. Specific checks for hypercholesterolaemia

The high cholesterol rates observed among patients also caused concern and have led to greater efforts to monitor and control cholesterol levels in at risk patients. As yet no beneficial effects of this policy have been observed - hopefully the next year will show some improvement in the general cholesterol levels observed in the clinic population.

5. Refinement of the education programme

The study of patient knowledge of diabetes revealed a number of areas where

improvements to the education programme could be made. In particular an analysis of the responses that patients were making to individual questions, those they got right and those they got wrong proved useful in examining the content of the course. For example the ability to differentiate between hyper and hypoglycaemia was deemed to be poor and the course was changed accordingly.

6. Examination of non-responding patients

The data raised some important issues concerning the nature of patients who had relatively poorly controlled blood glucose. For example these patients included a higher proportion of women, and a high proportion of patients that had been diagnosed for some time.

In fact the general picture of blood glucose control against time since diagnosis revealed a depressing though realistic picture of what should be expected from the clinic. Some analyses examined specific treatments for poorly controlled cases. Though it is difficult to generalise in such cases there were some questions raised over the appropriate choice of therapy for some cases.

7. Changing the assay for HbA1.

The changes in mean HbA1 levels observed during 1988-1989 showed an increase in all patient groups - and at face value were considered alarming. This raised the question of whether the laboratory assay had changed. It appears that though no formal change has occurred in the assay the laboratory reports this increase across all its samples covering two other hospitals (the general lesson is reported in Appendix 6 and in a presentation at a BDA meeting Home et al 1991). A decision has been made to change the techniques used.

Without the ability to review longitudinal change provided by this study this problem may not have been detected. It was felt that this may explain the otherwise unprecedented recent increase in referrals to the clinic.

8. Changes in calibration of biothesiometer and albustix

Longitudinal analyses of the albustix test for alubuminuria and the biothesiometer tests

for neuropathy also identified problems in calibration which may otherwise have gone unnoticed. Comparing successive years it became clear that in both cases the number of reported positive results had increased dramatically and far more than could be explained in clinical terms.

Changes such as these could have widespread effects on the diagnosis (and treatment rates) for clinical complications of diabetes. The analysis of results across a population of patients, as in the outcomes study, enabled these problems to be identified.

9. Development of parallel studies

A number of additional related studies of patient outcome have been undertaken or are emerging and the department now has considerably more expertise in the tools required to monitor outcomes. In general there has been a shift towards looking towards better outcome measures rather than traditional measures of process. Perhaps the best example is currently underway and that is a specific examination of the changes when patients are put on to insulin treatment. This simple study is using retrospective and prospective data to address the fundamental question of whether moving patients on to insulin is achieving real results in terms of both blood glucose control and treatment satisfaction. The results could potentially change the criteria that are used in making this decision.

Perhaps the most important consequences of this study lie in the ability of the department to examine its performance in a new way and incorporate patient outcomes as a critical part of that assessment. Thus the value of the project was expressed by one clinician (Appendix 3) as:-

> *"We are now clear that the data can be collected, examined against standards and that the result can be presented in a useful form. All those that have seen the analysis have found it useful.....Given a large boost to our general audit activities and expanded our views of what it might achieve"*

In summary the study of outcomes in diabetes, despite the difficulties posed by a chronic disease, has been successful in agreeing the concepts of outcomes monitoring. It has also

led to the recognition that the data can be useful in identifying and implementing improvements to clinical practice. Clinical staff are keen to continue this work.

Chapter 7. Angioplasty (PTCA)

A. Introduction

The Freeman Hospital includes one of the most prestigious centres for cardiothoracic surgery and medicine in the country. The cardiology department includes 4 full time NHS consultants, 2 part time academic posts and 3 part-time consultants (shared with neighbouring hospitals). Two consultants have led the use of PTCA within the Freeman for the treatment of angina. The procedure has been in use at the Freeman since 1981, and the skills required to perform the procedure are being passed on to other consultants. As a regional specialty patients can come from different parts of the country - though most are from the Northern region.

Angioplasties were selected for the study mainly because it was a relatively new area of medicine which was expanding rapidly. Given the relative novelty of the procedure there existed considerable uncertainty over when this treatment is to be preferred to others.

Indications and natural history

Coronary heart disease (CHD) is one of the most common chronic diseases in the UK. The disease is caused by impaired blood flow in the critical vessels supplying the heart. When sufficiently severe, this results in anginal chest pain of varying degrees of severity, and with a high risk of myocardial infarction and death.

The technique of Percutaneous Transluminal Coronary Angioplasty (PTCA), was first used in 1977 (Gruntzig 1977), and uses a small balloon on the end of a catheter to dilate narrowed vessels. Within the relatively few years since then, PTCA has come to be an established treatment for coronary heart disease. As such the procedure sits between the medical management of angina through a variety of drugs and the more invasive surgical procedure of by-pass grafting (CABG). The use of this procedure has certainly grown

quickly due both to the fact that it is less invasive than surgery and that it is claimed to be less expensive (Treasure 1990). In the US it is reported that PTCA had by 1987 become more common than bypass surgery as a method of revascularisation (Baim & Ignantious 1988).

In its early years the use of PTCA was limited to relatively low risk patients with straightforward single vessel disease (Detre et al 1988) and stable angina. However as techniques 'improved' and equipment became more sophisticated, PTCA came to be used on a wider variety of patients including those with multi-vessel disease (Holmes,Reeder and Vliestra 1988) or in other specialised situations (Sprigings et al 1988).

Despite the rapid growth, "uncontrolled" according to one commentator (Treasure 1990), there is still considerable uncertainty over its long term efficacy. Although PTCA is considered as a cost-effective method of revascularisation for patients with single vessel disease and stable angina its cost-effectiveness over surgical revascularisation is in some doubt. In stable angina and multi-vessel disease the initial complication rate is higher with PTCA than with CABG. The selection of patients for these different revascularisation techniques varies throughout the world indicating there is as yet no consensus (Chassin et al 1987; Brook et al 1988; Naylor et al 1990; Gray et al 1990).

There are a number of long term prospective randomised trails underway including RITA (Randomised Intervention Treatment of Angina) a British funded trial seeking to recruit patients in 15 centres (though not at the Freeman Hospital). Results from these various trials should become available in the next 2 to 5 years.

In the United States, a National Heart, Lung and Blood Institute registry of PTCA's performed has been established (Detre et al 1988). Comparisons between the early years (1977-1981) and later years (1985-86) suggest that there were changes in both the nature of patients treated, that is they were older with more complex disease and treatment histories, yet outcomes observed in hospital were better.

B. Data set and Data Collection

Identifying outcome indicators

The key patient types considered were those with stable and unstable angina. Though the range of outcome indicators for both these groups were similar the observed results were expected to be different for the two groups. All reports therefore showed the results separately for these two groups.

a. Successful procedure

The actual PTCA procedure does not always result in the dilatation of the required vessels, therefore the first consideration was whether dilatation had been achieved - success usually being defined as when the residual narrowing is less than 50% on all vessels attempted and with no adverse events occurring. This is one of the standard measures of angiographic success with rates of the order of 85-88% quoted in the literature (Henderson et al 1989; Glazer et al 1990; Rupprecht et al 1990). There were some early problems in the Freeman study in distinguishing an attempted angioplasty from an investigative angiography (exploration with catheters) - the decision to perform the procedure not always being taken in advance. An attempted angioplasty was therefore defined as one where 'the guide wire was passed into the vessel in order to cross the lesion'. Subsequent failure to dilate the vessel was deemed a 'failed angioplasty'.

b. Death

With coronary disease there will always be a significant mortality rate, in hospital and to follow-up, especially among cases with unstable angina. A number of studies have examined survival after PTCA and observed a mortality rate of up to 2% (higher if PTCA was unsuccessful) after 5 years. (Henderson et al 1989; Detre et al 1988). In general the majority of deaths will be from cardiac causes.

c. CABG - elective or emergency

In patients with stable angina, PTCA should preclude the need for subsequent early

surgery. In some cases when the PTCA is unsuccessful the patient is transferred immediately for an emergency CABG. In other cases elective surgery for stable angina patients may follow at a later date due to either the failure of the PTCA, the recurrence of the dilated lesion or the progression of coronary artery disease. The proportion of cases undergoing CABG is yet another of the standard measures used to gauge PTCA success rates. The anticipated rates range from 5% within the first year to 10% after 5 years. For patients with unstable angina the initial aim is to stabilise the patient and later CABG may be appropriate.

In this study the distinction was made between emergency and elective surgery when collecting data. However in the analysis these groups were pooled.

d. Repeat PTCA

In most cases the procedure should only be necessary once (at least within 5 years), thus the rate at which patients received repeat PTCA's in the same vessel was regarded as a poor outcome. The procedure may be repeated because the first attempt was unsuccessful in dilatation or because of the patient symptoms or a process of re-stenosis of vessels. (Holmes,Reeder & Vliestra 1988; Leimgruber et al 1986; Serruys et al 1988). Repeat PTCAs are said to carry greater risks (Sugrue et al 1987).

e. In-hospital events - post-procedure

As with surgical procedures there are a number of problems that may be manifest before discharge from hospital. These included myocardial infarction, recurrence of angina, and local vascular complications. Detre et al observed an in-hospital non-fatal MI rate of 4.3% and it was expected that such events would happen with a measurable frequency. The goals of treatment are that such problems, though inevitable to some degree in this population of patients, should be as rare as possible.

f. Decreased Anginal Pain

One of the prime indications of success is a reduction in the severity of chest pain felt by the patient. The assessment of anginal pain was based on the standard classification

of the Canadian Cardiovascular Society (CCS)(Cambeau 1976) which distinguishes four classes of pain:

I *"Ordinary physical activity does not cause... angina", such as walking and climbing stairs. Angina with strenuous or rapid or prolonged exertion at work or recreation.*

II *"Slight limitation of ordinary activity". Walking or climbing stairs rapidly, walking up-hill, walking or stair climbing after meals, or in cold, or in wind, or under emotional stress, or only during the few hours after awakening. Walking more than two blocks on the level and climbing more than one flight of ordinary stairs at a normal pace and in normal conditions.*

III *"Marked limitations of ordinary physical activity." Walking one to two blocks on the level and climbing one flight of stairs in normal conditions and at a normal pace.*

IV *"Inability to carry on any physical activity without discomfort - anginal syndrome may be present at rest."*

Though such schemes are part of the basic vocabulary in cardiology, there were problems in the extent to which clinical staff used the same criteria to assess patients (in particular there can be confusion between the CCS scheme and the New York Heart Association definitions). Therefore changes in pain were based on series of questions asked of the patient, the angina class could be estimated from the patient's response to five key questions on the degree of pain they suffered. These questions were adopted so as to be similar to a study, then current, of CABG patients that was being undertaken. In addition to the classification obtained from the patient, the angina score recorded in the notes (albeit sometimes implicitly) was also recorded for validation.

There is relatively less known about the clinical success of PTCA in terms of changes in anginal pain though one study (Henderson et al 1989) has reported that, after 5 years, 70-80% of patients remained free of angina and a further 6-12% had improved at least 2 angina grades (CCS scale). Rupprecht et al found that after 3 years, 59% of stable and 84% of unstable angina patients had experienced an improvement of at least one class in the CCS classification (Rupprecht et al 1990).

g. Increased walking distance

A standard method of assessment of the severity of angina is based on how far the patient can walk before pain starts. After the procedure the patient should be more active and able to walk further before pain commences. This was gauged simply by asking the patient before and after the procedure. More sophisticated methods are available and often used, for example exercise testing, however this was ruled out because though it was common it was not routine on all patients especially at follow-up. It would also require data to be collected by clinical staff which would be more restricting in terms of data collection ie require an out-patient visit for the follow-up.

h. Improved well-being

A number of different measures have been used to monitor quality of life in cardiovascular disease and coronary surgery (Fletcher, Hunt & Bulpitt 1987; Mayou & Bryant 1987). The Nottingham Health Profile (Hunt, McEwen & McKenna 1986) was chosen as a suitable instrument for assessing patient well being. This had the advantage that other studies of heart diseases had used the scale successfully (Wallwork & Caine 1985; O'Brien, Buxton & Ferguson 1987; Buxton et al 1985; O'Brien et al 1988).

Where possible patients were given forms before the PTCA. In the case of emergencies this was not possible, so only follow-up data was available. In some cases the patient may not have had any serious symptoms prior to admission, for example following emergency admission for a first myocardial infarction. In these cases the expected pre-admission (pre-morbid) scores would be approximately zero.

i. Other health events after discharge

In addition to major health events during the hospital stay, it was also necessary to consider some events that the patients might experience after discharge as indicators of poor outcomes. The patient was asked at follow-up whether they had been readmitted to hospital , or visited A&E departments or their GP and if so the reasons why. There were some problems found in interpreting the reasons given by patients and their exact clinical significance in relation to the PTCA.

j. Decreased Potency of Medication

At least one other study has examined medication after PTCA, in that case as a proxy for quality of life (Rupprecht et al 1990). This reported that 45% of stable angina patients were receiving no anti-anginal therapy at three years. There was some debate in this study as to whether decreased medication - essentially a process measure - should be included in our dataset. In the end it was decided that the information may be useful and so was included. The potency of medication was based on a simple six point classification, the aim being to see a drop down the scale at follow-up. Pre-PTCA information on medication came from the notes, at follow-up information came from the patients themselves. Though the latter did cause some concern about its reliability , the validation suggested that the patients were well aware of the nature of the drugs they were taking.

Medication Scale:

1 One drug taken for cardiac problems eg aspirin, diuretic, antiarhythmic
2 One drug for cardiac problems plus one other
3 Two drugs for angina from the following classes - beta-blocker or calcium antagonist or long acting nitrate
4 Two drugs for angina (as 3) plus another drug for cardiac problem (eg aspirin, diuretic, ACE-inhibitor, antiarhythmic)
5 Triple therapy - angina drugs including a beta-blocker and a calcium antagonist and a long-acting nitrate
6 Triple therapy plus another.

Timing of observations

It was agreed that as well as obtaining baseline data before the procedure and brief details on discharge from hospital, follow-up information should be at three months and twelve months in the first instance. As not all the patients returned to the Freeman (some go to local doctors in distant parts of the region) the follow-ups could not be tied to hospital outpatients appointments after discharge.

One of the complicating factors in this study was that a number of patients received a repeat PTCA. The practice for these patients was to consider the first analysis ended at

this point (an endpoint had been reached) and start again on the second PTCA. If the necessary data for the pre-procedure assessment was already available from recent earlier follow-ups, eg if the patient had complete an NHP 2 weeks earlier, this was used rather than repeating data collection.

Other data items

In order to help the analysis of the database additional information was collected :-

a. The severity of the disease and the affected vessels was recorded.
This included whether the patients had single, double or triple vessel disease (taken from CASS 1983a). The degree of stenosis and the particular vessels affected was noted - based on evidence from cardiac catheterisation. An indicator of ventricular function was also used with categories of either poor, moderate or good. If no comment was made and no evidence of dyskinesia was found ventricular function was assumed to be good. In addition a list of basic risk factors and information about previous problems was collected.

b. The details of the procedure undertaken, the vessels attempted and dilated and any problems that emerged during the procedure.

c. The cardiologist's assessment of risk was added some way into the study. This was based on a simple three point classification of the risk of an adverse outcome for that patient.

Data collection

The model for data collection on angioplasty was almost identical to that used for cholecystectomy. Once again a local research nurse was available to undertake the co-ordination of data collection.

1. Patient Identification

Most eligible patients were admitted to one ward and identified by the local research nurse through contacts with ward sisters. In some cases it was not known beforehand if a patient was to receive a PTCA - rather than exploratory catheterisation. If in doubt these were initially included in the study and excluded later. Some patients were emergency cases or too ill to complete forms and data could only be collected retrospectively and at follow-up.

2. Pre-procedure questionnaire and background information

Patients joining the study were given a letter of introduction and asked to complete a NHP and questionnaire checking details of the severity of symptoms (see Appendix 7). On completion these were handed back to the research nurse. The basic clinical details were collected from the medical notes, typically after discharge when the notes were more complete. There were a number of areas where a good working knowledge of cardiology was required to elicit information that could be hidden in test results in the notes or in laboratory reports.

3. Discharge information - complications

Brief details of events during the hospital stay were completed by the nurse after the patient had been discharged. A simple pro-forma listing the main possibilities was used with a category to deal with other conditions.

4 Follow-up information

The follow-ups at three and twelve months were carried out via postal questionnaires to patients. These reproduced the pre-PTCA questions on symptoms and the NHP. In addition there were questions concerning medication, re-admissions and visits to A&E departments or to their GP. If patients failed to respond to their first follow-up within one month they were automatically sent another questionnaire. If no response was forthcoming after the second questionnaire, the local GP was checked to see that the patient was still alive.

5 Data inputting

All data was input through locally developed software into a series of databases. Analysis and reporting was also through local software.

Validation

A number of approaches to validation were adopted for different items in the data set.

1. Patient Interview

A semi-structured interview was used to validate some of the general details of the patients experience. The interviews were carried out shortly after follow-up questionnaires had been returned, though these were not seen by the interviewer beforehand.

Time since anginal pain began:- There was general agreement on the time since anginal pain started between information given at interview and the result of the questionnaire. Discrepancies were noted in only one case which was not clinically significant (ie 2 years versus three years).

Previous medical history:- In most cases the interview and questionnaire revealed the same details though in two cases it appeared the questionnaires missed important prior events, one being a myocardial infarction, the other a previous PTCA.

Walking distance:- There were some problems observed in estimating the 'walking distance' before the PTCA. In some cases at interview it was confused with the state before the angina. This made comparisons difficult. The two methods did agree on the two extremes of the spectrum ie cases who walked over 2 miles with no problems or cases where walking was very limited. Assessments of walking distance at follow-up closer to the actual times of interview showed greater agreement with only one case showing a difference (less than half a mile

at interview but 2 miles on the questionnaire).

2. Comparison of angina scores

The duplication of angina scores, one obtained from the notes and the other from the patient allowed a comparison between the two methods. There was a significant association between the two scores ($p < .05$) though there was an exact agreement in only 36% of cases, with 31% of cases higher on the patient's score, and the remaining 33% higher in the notes. The patient completed score showed a significantly lower proportion of patients recorded as class III angina. The pattern of responses required to fall into this class demanded correctly interpreting the question about whether pain was observed only during certain conditions - a more complex question than the others. It is likely that in some cases patients have ticked the boxes incorrectly and this has resulted in an assignment of class II rather than class III angina. In total more patients reported having the most severe class (38%) than did the notes (26%). It was noted during the course of the study that there were some differences in the definition of class IV angina (the New York Heart Association scale differs from the Canadian one used here). This may explain the discrepancy (Cambeau 1976).

It should be noted therefore that in order to be consistent between pre-PTCA and follow-up, the angina classes quoted in the data analysis section may not correspond exactly to those defined by the clinicians but are drawn from the responses given by the patient. The outcome indicator therefore relates to changes in pain based on the pattern of responses given by the patient.

3. Internal consistency

A number of checks were made for internal consistency between data items and these are reported in the data analysis section. In particular the NHP 'Pain' score was found to be significantly associated with the class of angina (see Data Analysis section).

4. Check against another database

The cardiologist primarily responsible for developing PTCA at the Freeman kept a

separate list of all the patients on whom he had performed a PTCA. The list was updated with major events that may have happened. The CASPE database was periodically checked against this list. In general the two matched up. The biggest problem occurred when cases were included in the CASPE study following exploratory catheterisation and PTCA was not attempted. This led to tightening the definitions of attempted PTCA. The classification of patients as either stable or unstable angina was taken from this database.

C. Data Analysis

The analysis of the data concerned three main questions.

1. Do the indicators show significant changes following PTCA?
2. Are the indicators related to each other?
3. Are the results influenced by the presenting characteristics of patients?

Response rates

As with cholecystectomy, the response rates were extremely good. From 202 patients (to November 1990) only 10 cases had no 3 month follow-up (95.0% response rate).

One problem was incomplete baseline data for emergency admissions due either to the patient being admitted too quickly or too ill to complete the forms. This is unavoidable but means that for unstable angina patients baseline data was only available in 53% of cases which form the sample used in the longitudinal analyses.

Changes to 3 months (Table 7.1)

Table 7.1 summarises the main events of patients to three months. The results have been split into two groups, stable and unstable angina. The population is further divided into

patients who had a successful angioplasty against those where the procedure failed. Some additional data on events within hospital after PTCA were collected (for example, return of angina on the ward, post-procedure complications) however the data has not been used in the analysis as it tends to overlap with other information.

Failed PTCA

In total 22 of the 202 patients (10.9%) had PTCAs that were unsuccessful - with a higher incidence amongst the unstable than stable angina cases. This figure roughly correspond to the 'failure' rate that is expected and compares with values of 5-15% quoted in the literature.

Events to three months

In total four patients died, three in hospital and one after discharge. Of these 2 had severe disease and had been considered unsuitable for surgery. One patient died as a result of renal failure and another after a heart transplant. None of these was a straightforward case and these results were not surprising at the time.

The most common event up to 3 months after the procedure was patients going for either elective or emergency surgery - in total 17 cases. This group included as expected a large proportion (though not all) of the failed PTCA cases and a high proportion of cases with unstable angina. Only one case had a subsequent myocardial infarction with no other procedure. Table 7.1 clearly shows that unstable angina patients were much more likely to have another event after the PTCA. Perhaps more surprising is that the stable angina category with successful PTCA included a number of repeat PTCAs to the same vessel. Some of these would have been planned beforehand as part of the treatment strategy for that patient.

The readmissions were a large category and included a variety of reasons other than repeat PTCA's, CABG or investigatory angiograms. The reasons quoted (from all cases) included:-

Table 7.1 Summary of events to three months by patient type
to 1/11/90 n=202 (10 Cases have no 3 month follow-up)

	Angina Type			
	Stable Angina		Unstable Angina	
	Successful PTCA?		Successful PTCA?	
	Yes	No	Yes	No
OUTCOME				
Died in hospital		1	2	
CABG in hospital	1	3	1	5
MI in hospital			1	
Died after discharge			1	
CABG after discharge		3	3	1
Repeat PTCA	5		3	
Readmitted	7		3	1
None of the above	98	6	45	2

Table 7.2 Changes in key indictors to three months by patient type.

		Stable Ok	Unstable Ok	CABG	Fail nec	Readmitted
No. cases (with pre+post data)		98 (80)	45 (24)	17 (12)	8 (5)	11 (11)
Anginal Pain	% Better	52.6***	46.7	62.5*	25.0	50.0
	% Same	42.1	33.3	25.0	50.0	33.3
	% Worse	5.3	20.0	12.5	25.0	16.7
% Class IV Start->3m		31->14%	44 ->20%	40 ->13%	25 ->25%	44 ->17%
Walking Dist.	% Better	51.4*	50.0	58.3	50.0	45.5
	% Same	40.0	35.0	25.0	25.0	36.4
	% Worse	8.6	15.0	16.7	25.0	18.2
Medication	% Better	36.3**	16.7	58.3	20.0	9.1
	% Same	42.5	45.8	25.0	80.0	81.8
	% Worse	21.3	37.5	16.7	-	9.1
% Triple Therapy Start -> 3m		56 ->48%	20 ->53%	41 ->33%	63->100%	55->64%
NHP mean admission	Energy	46.4	45.9	49.9	34.4	40.6
	Pain	25.8	26.1	25.8	26.0	20.2
	Emot Reac	24.0	24.2	27.4	14.5	33.1
	Sleep	34.8	25.3	30.5	17.1	46.7
	Soc Isol	10.9	12.7	18.8	12.1	11.3
	Mobility	16.2	19.9	24.8	21.2	20.5
	Part II	51.9	48.8	53.6	54.3	57.2
	Combined	26.8	25.7	29.5	20.7	28.7
NHP Follow-up	Energy	29.0	25.2	22.7	34.4	41.7
	Pain	10.7	13.1	12.1	12.6	15.8
	Emot. Reac	14.8	19.6	15.3	15.1	22.0
	Sleep	28.5	20.6	33.7	17.1	44.0
	Soc Isol	5.5	5.5	1.7	8.3	13.4
	Mobility	10.6	13.1	17.1	22.0	16.0
	Part II	27.2	35.7	33.3	57.2	40.3
	Combined	16.5	16.2	17.1	18.2	25.5
Change in NHP	Energy	17.5***	20.6**	27.2*	0	-1.2
	Pain	15.1***	13.0*	13.5*	13.5	4.4
	Emot. Reac	11.6***	4.6	12.1	-.5	11.1
	Sleep	6.3	4.7	-3.2	0	2.7
	Soc Isol	5.4*	7.2	17.1*	3.9	-2.0
	Mobility	5.6*	6.8	7.8	-1.9	4.5
	Part II	23.9***	13.1	20.2**	-2.9	16.9*
	Combined	10.2***	9.5*	12.4**	2.5	3.3

*** p<.001, ** p<.01 , * p<.05 Changes to 3 months Wilcoxon Rank sum test.

F 'Chest pains, pains in neck & jaw'
F 'Blood clots above and below angioplasty'
F 'Severe chest pain'
F 'Chest pains'
O 'Virus chest infection'
O 'Many times chest pain '
O 'Gastroscopy'
F 'Chest pain'
O 'Pains in both arms and up neck'
F 'Gastric reflux'
F 'Kidney transplant'
F 'Clot in femoral artery. Due to angioplasty'

(Cases marked with F were admitted to Freeman, O to Other hospitals). It became clear that the reasons for re-admission may not necessarily be related to the PTCA (the same also applies to visits to A&E departments) and that cases in this category needed careful interpretation.

Table 7.1 makes clear that it is difficult to isolate the one event of PTCA in these patients when there are a considerable number of other events and interventions happening in this chronic disease. Thus examining the outcomes of PTCA becomes more complex as the patient population fragments into different groups as different treatment options are used.

Subsequent analyses grouped patients into five categories as defined below:

a. Stable angina no events. (Stable Ok)
b. Unstable angina no events (Unstable Ok).
c. Patients who have had CABG (CABG).
d. Patients whose PTCA was unsuccessful (not elsewhere classified) (Failed nec)
e. Patients who have been readmitted other than for CABG or PTCA (Readmit).

Some patients do not fall into any of these groups, including repeat PTCA patients and patients who have had only a myocardial infarction. For patients with a repeat PTCA at the Freeman Hospital, it was the practice to stop data collection from the first procedure and start on the second, effectively recognising that for assessing the outcome

of the first PTCA an appropriate end-point had been reached.

Changes in anginal pain (Table 7.2)

The analysis of anginal pain was based on the patients response to the questionnaire and not on clinical assessments of their angina category. The caveats about whether the resulting score represents the same value as that made by a clinician must be borne in mind - however the main interest is in longitudinal change. Results show the proportion of patients who had improved by at least one class, stayed the same or got worse. In addition the proportion of patients with the most severe class of angina (IV) before and after the procedure are shown.

The results indicated a general improvement for most groups with around 50% of patients recording an improvement of at least one class.

Considering first the stable angina group, about half showed some improvement whilst only 5% appear worse. After the procedure only 14% of these patients had Class IV angina compared with 32% at the start. Thus this group showed a net improvement and one that was statistically significant. It is possible that there were beneficial changes in the patients who stayed in the same angina category but that the instrument was not sufficiently sensitive to the scale of change.

The unstable angina group also showed a net improvement though in this case a higher proportion of patients appeared worse. This was presumably a reflection of the 'instability' of the angina. The results cannot show what proportion of patients would have got worse without treatment (or with other treatments). The changes in CABG patients and those readmitted are similar to these two groups.

The worst results are shown for the 'Failed PTCA nec' group where as many patients got worse as got better and the proportion of cases with Class IV angina stayed the same. This is as one would expect for this group of patients. The lack of any change is an encouraging sign of the validity of the measure and suggests that the improvements seen

in the other patient groups can be attributed to the successful PTCA rather than other factors.

Change in walking distance

The walking distance for patients was classified into 4 groups.

1. 'Never leaves the house' or 'Drives everywhere'
2. 'Walks less than 1/4 of a mile'
3. 'Walks 1/4 to 1/2 mile'
4. 'Walks more than 1/2 a mile'

As with the anginal pain, cases were classified as better, the same or worse on this ordinal scale corresponding to the patient's progress to three months. All patient groups showed increased walking distance in about half of the cases. All groups also showed a noticeable proportion of cases where the walking distance is less at follow-up than at admission, the worst being the 'Failed nec' and 'Readmitted ' categories. The groups of 'Stable Ok' patients seem to fare best. As with the pain scores the change in walking distance reveals a considerable proportion of patients who are the same or worse - a fact which raises questions over the sensitivity of the score.

Changes in medication

The six point medication scale described earlier was used, the less potent the medication the 'better'. The last two categories of this scale cover triple therapy. The pattern of changes in medication were different across patient groups. 'CABG' patients showed the best improvements. The 'Stable Ok' group showed a slight reduction in medication whilst the 'Unstable Ok' cases appeared worse. Similarly the proportion of cases on triple therapy showed little reduction from before PTCA to follow-up. This overall lack of change in medication has been raised by clinical staff as a surprising result. It was suspected that any review of medication following PTCA was occurring after three months. In all cases there still remained a high proportion of cases on triple therapy after the procedure.

Changes in health status

The mean NHP scores at admission and changes in score to follow-up are shown in Fig 7.1 and detailed figures given in Table 7.2. The NHP scores on admission were largely similar for all patient groups. Higher scores were observed for 'Energy', 'Sleep' and Part II of the profile. The latter perhaps showing how the disease itself appeared to have a fairly large impact on everyday living amongst these patients.

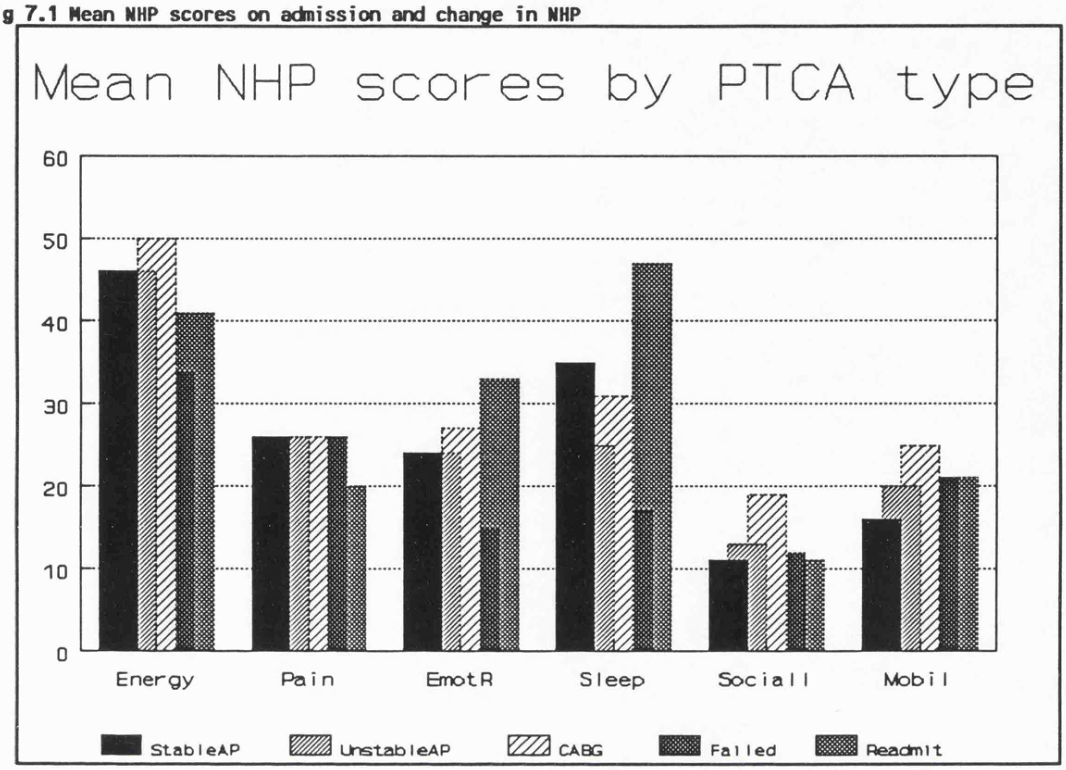
At follow up the scores for most groups were significantly lower yet the relative scores between dimensions stay largely the same. Thus for example 'Energy' and 'Sleep' which were high scoring dimensions before PTCA were also the highest at follow-up. The scale of improvements amongst the stable, unstable and CABG groups are roughly comparable and the changes in the higher scoring dimensions tend to be significant for all these groups.

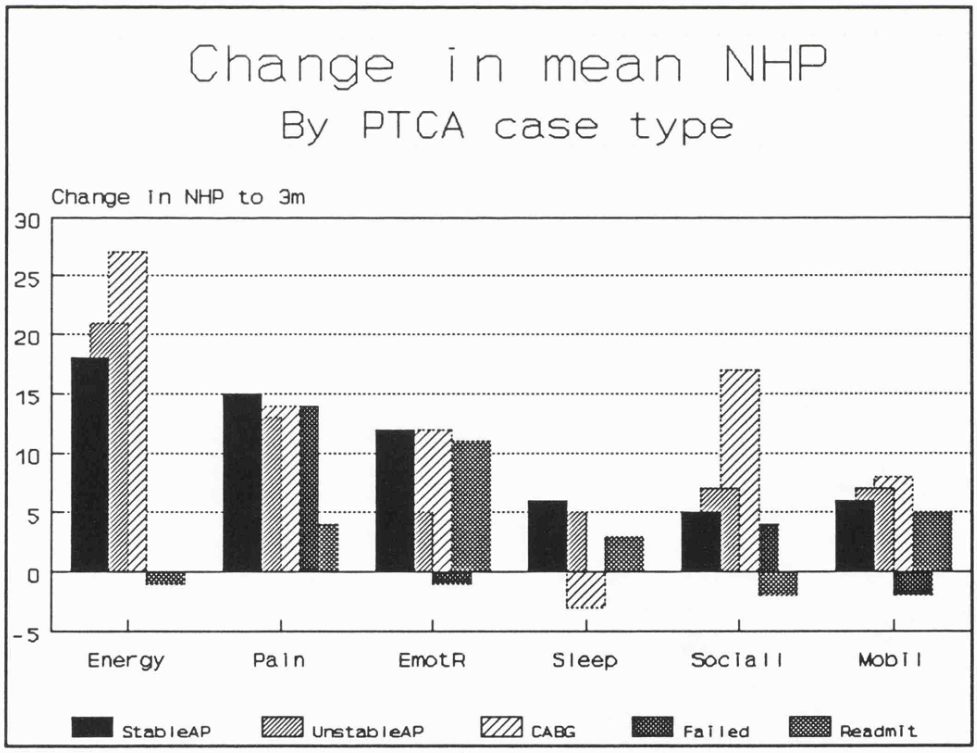
The exceptions to this general pattern of improvement in NHP are the group of failed PTCA and readmitted cases where the observed improvements are noticeably less. On the latter group the high scores at follow-up reinforce the idea that there is something different about these patients that has led to their readmission. It is interesting to observe that the 'Failed PTCA' patients show little change on most NHP dimensions - which is what one would hope to see if the instrument is working correctly.

Changes to twelve months (Table 7.3)

Table 7.3 summarises the main events in those patients who have reached the stage of a twelve month follow-up. It is clear comparing this table to the events at three months (Table 7.1) that there were a substantial number of changes between 3 and 12 months. In particular, patients were still liable to go for CABG or be readmitted for a repeat PTCA. Once again activity amongst the unstable group is higher than for stable patients. No additional patients died during this period.

Fig 7.1 Mean NHP scores on admission and change in NHP





Twelve months after the PTCA only 49% of stable patients and 29% of unstable patients had had a successful PTCA and had not had another intervention, an MI or been readmitted to hospital. Very few of the patients have died (3%) or had an MI (2%), but from our original sample roughly 17% have had a CABG, 14% a repeat PTCA and 21% had been readmitted for some other reason.

Table 7.3 Events after PTCA to 12 months

n=95 2 cases not respond at 12m

	Angina Type			
	Stable Angina		Unstable Angina	
	Successful PTCA?		Successful PTCA?	
	Yes	No	Yes	No
OUTCOME				
Died		1	2	
CABG	4	6	4	2
MI	1			1
Repeat PTCA	8		5	
Readmitted	10	1	9	
None of the above	30		10	1

Table 7.4(a) summarises the changes in indicators between three months and twelve months. The general picture is of relative little change beyond the improvements seen to three months for those patients who are not readmitted. It appears that the beneficial effects of the PTCA are manifested within the first three months. Beyond three months other events may occur in a large number of cases and tend to reduce the overall success rate of the procedure.

The number of cases in each category is relatively small so it is difficult to be confident about the results at this stage. However results for patients who have been readmitted appear generally worse than the other groups, once again reinforcing the idea that the fact of readmission and a worse health status are linked. Once again there are a variety of reasons for these readmissions apart from CABG, PTCA and angiograms. As with cholecystectomy there are more unrelated admissions between 3 and 12 months than before the 3 month follow-up.

Table 7.4(a) Changes in key indicators to 12 months by patient type

		Stable Ok	Unstable Ok	CABG	Fail nec	Readmit
No. cases with pre+post data		25	10	14	-	18
Anginal Pain	% Better	67	50	64	-	39
	% Same	29	50	27		31
	% Worse	5	0	9		31
% Class IV Before ->12m		29->5	50->25	55->18		46->46
Walking Dist.	% Better	53	80	44	-	60
	% Same	33	20	28		15
	% Worse	13	-	28		25
Medication	% Better	64	25	43		36
	% Same	18	50	29		29
	% Worse	18	25	29		36
% Triple Therapy Before->12m		50->23	75->75	43->57		57-49
NHP mean admission	Energy	47.5	31.0	42.7	-	56.2
	Pain	18.9	11.0	27.8		42.7
	Emot. Reac	19.8	2.6	22.2		30.1
	Sleep	22.4	13.5	35.4		50.2
	Soc Isol	6.5	0	10.3		14.0
	Mobility	15.0	8.5	14.2		17.5
	Part II	47.4	46.4	50.0		63.3
	Combined	21.8	11.1	25.4		36.8
NHP combined score at 3 mnths		13.1	7.2	19.4	-	24.1
NHP 12 mnth	Energy	23.7	0	40.9	-	44.7
Follow up	Pain	9.6	0	13.8		21.8
	Emot. Reac	14.5	0	20.7		20.7
	Sleep	22.9	16.6	36.2		31.5
	Soc Isol	2.4	0	5.5		10.3
	Mobility	12.7	0	11.5		24.6
	Part II	23.4	21.4	43.9		33.7
	Combined	15.1	2.8	21.4		25.6

Table 7.4(b) Changes in key indicators from three months to 12 months by patient type

		Stable Ok	Unstable Ok	CABG	Fail nec	Readmit
No. cases with pre+post data		25	10	14	1	18
Anginal Pain	% Better	17.4	12.5	36.4	-	16.7
3m-12m	% Same	56.5	75.0	36.4	-	38.9
	% Worse	26.1	12.5	27.3		44.0
% Class IV 3m-12m		4 ->4	13 ->13	9 ->18		39->11
Walking Dist.	% Better	10.0	10.0	38.9	-	20.0
	% Same	83.3	90.0	44.4		55.0
	% Worse	6.7		16.7		25.0
Medication	% Better	44.0	30.0	35.7	-	33.3
	% Same	24.0	50.0	28.6		44.4
	% Worse	32.0	20.0	35.7		22.2
% Triple Therapy 3m->12m		52->24	60->40	50->29		50->50
NHP	Admission	21.7	11.1	25.4	-	36.8
Combined score	3 months	13.1	7.2	19.4		24.1
mean	12 months	15.1	2.8	21.4		25.6

The reasons for readmission after three months included:-

O 'Chest and viral infection'
F 'Vascular op and blood clots'
F 'Infect in leg'
F '?Pericarditis and test as CABG did not work'
O 'Hysterectomy'
O 'D&C - for slight bleeding'
O 'Colitis'
O 'Fitting Tenckoff catheter, NGH-chemo'
O 'Severe anaemia'
F 'Repeat camera tracings around heart'
F 'Chest pain + repeat angiogram'
O 'Angina attack'

Table 7.4(b) shows the changes from before PTCA to the twelve month follow-up. This table therefore covers the improvements seen to three months and possible deteriorations beyond that. The stable angina patients show clear net improvements in walking distance, anginal pain, and NHP scores. For this group it also appears that there has been some change in medication presumably in the months after first follow-up.

However there remains a substantial proportion of patients on the most potent medication - triple therapy - one year after PTCA. The unstable patients (of which there are relatively few who have not been re-admitted within a year) showed changes in anginal pain and walking distance and extremely good NHP scores after one year. For this small group the outcomes looked very good - not only has the PTCA worked to prevent an MI, but there appears to have been a net improvement in the general health status which was not really expected. More numbers are needed to confirm this pattern.

The outcomes of the group of CABG patients appeared to show a net improvement up to one year after the PTCA - though the degree of change was not as marked. Analysis of this group could be confused by the fact that the time since the CABG and the follow-up could vary between 1 day to one year. Nevertheless improvements in anginal pain and walking distance were observed. Similarly improvements were seen in some dimensions of the NHP but the scale of the change was less than for the previously mentioned categories.

The group of patients who have been readmitted did not show clear improvements. Walking distances were slightly better but the anginal pain and medication appeared largely the same. The NHP scores tended to be higher at admission and at follow-up for this group.

Links between outcome measures

The various indicators were compared on admission and at follow-up. In addition the observed changes in indicators between admission and follow-up were compared. In summary the results of these comparisons suggest that there appears to be some overlap between the pain scores, the NHP scores and the walking distance when viewed at admission or at follow-up. Though the indicators do not reveal identical results the general messages were comparable. The following section gives some examples of these links.

Table 7.5(a) shows the relationship between the pain score at admission and the mean NHP scores on selected dimensions. As discussed earlier the NHP 'Pain' score was significantly related to anginal pain class, as were the 'Mobility' and combined scores in the direction predicted. The higher the pain scores the higher the mean NHP scores, a relationship that appeared particularly strong for class IV angina cases.

Table 7.5(a) Relationship between selected dimensions of NHP and admitting pain score (n=179)

	Pain	Mobility	Combined	% Triple Therapy
Pain score=0	12.7	8.7	26.2	50%
=1	20.6	16.7	24.6	67%
=2	20.4	14.2	22.1	58%
=3	21.2	9.6	23.5	45%
=4	38.7	26.1	34.8	60%
Significance	p=.0002	p=.0007	p=.005	ns

Table 7.5(b) examines NHP scores and other indicators against walking distance. Excluding the group "Drives every where, never leave house", the relationship was as

expected. The greater the claimed walking distance the lower the NHP scores, the gradient applying across all dimensions. Similarly the proportion of patients with class IV angina reduced across these categories showing the link to pain scores. Thus there

Table 7.5(b) Mean NHP scores, %Class IV angina and %triple therapy by walking distance categories (n=179).

	Dist=1 Home/drives	Dist=2 <1/4 m	Dist=3 1/4-1/2m	Dist=4 >1/2m
Energy	42.2	64.7	48.1	25.6
Pain	25.4	39.4	25.6	16.9
Emotional R	22.2	33.1	26.2	19.8
Sleep	31.4	43.7	33.6	25.4
Soc Isol	10.9	19.7	11.2	5.8
Mobil	17.4	24.2	20.0	8.5
Part II	50.7	60.9	55.5	36.0
Combined (p=.005)	24.9	37.5	27.4	17.0
% Class IV	49%	61%	29%	27%
% Triple Therapy	62%	65%	62%	50%

appeared a consistent picture that patients who were more active tended to have lower pain scores and lower NHP scores. This relationship was also observed when follow-up scores are compared.

The exception was for category 1 ("Drives everywhere" or "Never leaves home") which did not fit this pattern. It appeared that patients falling into this category may not have been giving a genuine reflection of their overall health though they may well have been answering the question truthfully. It is suggested that because of this these cases should therefore be excluded from the ordinal scale.

The links between these indicators and the medication category were more complex and difficult to interpret. There was no consistent pattern whereby the patients on more potent medication appeared less active or healthy as is shown in Table 7.5(c).

Table 7.5(c) Mean combined NHP scores by medication category.

Medication	Category	No.	NHP Combined Mean	Std. Dev
Single	2	13	27.4	12.03
Double	3	9	21.6	15.82
	4	35	27.6	18.45
Triple	5	17	30.0	16.99
	6	77	25.3	21.08
Total	-	154	26.2	19.00

Comparing changes indicators.

Table 7.6 shows the correlations between the dimensions of the NHP at admission , follow-up and the changes between admission and follow-up. There was a high degree of correlation between the scores of individual dimensions of the NHP - most correlations were significant to the 0.1% level. The most important aspect of the table is that the fact that changes in score were related to the initial score. That is the higher the admission the higher was the change in score. All the NHP scores showed skewed distributions with a high proportion of cases scoring zero - particularly at follow-up. It is possible that there are some 'end-effects' caused by using a scale where improvement in score will tend towards zero.

Table 7.7 summarises the relationship between changes in certain indicators from admission to three months and shows average change in the combined NHP score for patients who appear better or worse on three indicators. The relationships were not the simple pattern one might expect given the correlation observed earlier when the indicators are compared in cross-section. Whilst it appeared that patients who were 'better' on the pain score and in terms of walking distance showed larger average improvements in NHP, the differences between patients who were the 'same' or 'worse' was not as one might expect. Patients who had the 'same' anginal pain had lower improvement in NHP than those who got 'worse'- which was not as expected.

Table 7.6 Correlations between admitting NHP, follow-up NHP and changes in NHP (n=132).

Correlations:	Energy	Pain	Emot. Reac	Sleep	Soc Isol	Mobil	Part II	Comb.
Admission								
Energy	1.0000**	.4799**	.4727**	.3875**	.4136**	.4974**	.4630**	.7997**
Pain	.4799**	1.0000**	.4601**	.3905**	.4251**	.6289**	.4461**	.7534**
Emot Reac	.4727**	.4601**	1.0000**	.4942**	.6572**	.3933**	.4699**	.7832**
Sleep	.3875**	.3905**	.4942**	1.0000**	.1276	.3204**	.3039**	.6601**
Soc Isol	.4136**	.4251**	.6572**	.1276	1.0000**	.3261**	.3820**	.6343**
Mobility	.4974**	.6289**	.3933**	.3204**	.3261**	1.0000**	.4833**	.6962**
Part II	.4630**	.4461**	.4699**	.3039**	.3820**	.4833**	1.0000**	.5801**
Combined	.7997**	.7534**	.7832**	.6601**	.6343**	.6962**	.5801**	1.0000**
Follow -up (3mnth)								
Energy	.4849**	.3069**	.2200*	.2661**	.2187*	.3500**	.1438	.4437**
Pain	.3163**	.3136**	.1712	.1573	.1687	.2754**	.1297	.3271**
Emot Reac	.3158**	.2576*	.4189**	.1044	.3583**	.3014**	.1368	.3922**
Sleep	.3108**	.3297**	.2967**	.5238**	.2062*	.2437*	.1414	.4537**
Soc Isol	.1739	.1312	.3818**	.1093	.3411**	.1859	.0415	.2886**
Mobility	.4001**	.4086**	.2007*	.1931	.2399*	.4863**	.1209	.4396**
Part II	.3577**	.3691**	.2458*	.0647	.2049*	.3997**	.3488**	.3716**
Combined	.4666**	.3942**	.3678**	.3366**	.3287**	.4076**	.1667	.5351**
Change in NHP								
Energy	.5624**	.1998*	.2746**	.1439	.2155*	.1768	.3367**	.3969**
Pain	.2061*	.6920**	.2980**	.2437*	.2673**	.3758**	.3165**	.4524**
Emot Reac	.1774	.2167*	.5956**	.3851**	.3180**	.1134	.3330**	.4098**
Sleep	.0801	.0639	.2041*	.4913**	-.0798	.0799	.1675	.2139*
Soc Isol	.2623**	.3060**	.3361**	.0384	.6949**	.1695	.3333**	.3855**
Mobility	.1489	.2779**	.2243*	.1555	.1183	.5925**	.3916**	.3194**
Part II	.0908	.0659	.1950	.2089*	.1541	.0716	.5681**	.1809
Combined	.3913**	.4139**	.4722**	.3714**	.3516**	.3392**	.4552**	.5375**

** p<.001 *p<.01

For the medication, patients moving on to less potent drug regimes (better) show a worsening NHP score, whilst those where the drugs are more potent show an improvement. This may be because a reduction in medication makes people feel worse though further investigation of this point is needed.

Table 7.7 Comparing changes in indicators. Mean combined NHP scores according to changes in pain, walking distance and medication.

	Better	Same	Worse
Pain	30.6	7.4	14.2
Medication	-5.4	10.6	13.8
Walking Distance	14.1	12.1	3.3

Effects of presenting variables

Presenting variables within patient groups

Table 7.8 summarises the prevalence of certain features of the patients on admission for each of the five main patient groups discussed earlier. In brief there were few significant differences between the presenting characteristics of the various patient groups.

There was little difference in the mean age of the groups ranging from 49 to 54 years. The proportion of women was slightly higher in the 'Unstable' angina group (36%) and lower in the 'Failed PTCA' group (13%) with on average 25% of patients being women.

The percentage of patients who had had a previous procedure (either PTCA or CABG) varied around the average of 15%, being lower in the 'Unstable' angina group and the CABG patients. Of the 'Failed PTCA' cases, 25% had had some previous procedure. The proportion of patients with a previous MI was more constant at around 32% and was highest in the CABG group (41%).

The complications (including hypertension, diabetes, smoking etc) showed few variations, with around 40% of cases having none of these additional problems or risk factors. The percentage of such 'non-complicated' cases appears slightly lower in the 'Unstable' angina group and the 'Failed PTCA' group.

The proportion of cases with disease of a single vessel (as opposed to two or three vessels) was lower in the 'Failed PTCA' cases. Together with the slightly higher 'stenosis score' it appears that the pathology of the disease in the 'Failed' cases was greater than for the successful cases. This observation had been made during the course of the project and was broadly in line with the view accepted by the clinicians.

Ventricular function was re-classified into either 'Moderate or poor' and the rest (very few cases were recorded as having poor ventricular function and most were 'good'). There appeared to be no large differences between the patient groups.

Table 7.8 Characteristics of patients on admission for main patient groups

	Total	Stable Ok	Unstable Ok	CABG	Failed nec	Readmit	All failed
No.	179	98	45	17	8	11	22
Mean Age (yrs)	53.1	53.4	53.7	50.9	53.9	49.7	52.3
% Male Female	74.5% 25.5%	77.6% 22.4%	64.4% 35.6%	76.5% 23.5%	87.5% 12.5%	81.8% 18.2%	81.8% 18.2%
% Prev op	14.6%	19.4%	8.9%	5.9%	25.0%	0	13.6%
% Prev MI	31.8%	30.6%	31.1%	41.2%	25.0%	36.4%	36.4%
Complic- =None ations =1 >1	42.7 47.9 9.4	48.0 44.9 7.1	33.3 51.1 15.6	47.1 47.1 5.9	25.0 75.0 -	45.5 54.5 -	40.9 59.1 -
% Single Vessel	59.9	63.3	60.0	58.8	37.5	63.6	40.9
% Vent.function Moderate/poor	14.6	19.4	8.9	11.8	-	18.2	9.1
Mean Stenosis Score	2.85	2.58	3.16	3.06	3.63	2.90	3.54

In general this limited analysis shows that there were not enormous differences between the presenting characteristics of these groups. The exception to this was the relationship between the more severe pathology of cardiac disease in the failed PTCA cases. Though there were relatively few cases where the PTCA was unsuccessful the group did appear to have had a greater degree of vessel involvement, more serious stenosis of vessels and a slightly higher rate of complications or risk factors. These factors add up to more clinically severe disease in these cases. Thus it appears that the simple rate of failed PTCA's will be especially sensitive to the severity of presenting disease and this factor must be examined in comparative studies.

For the other groups the approach used here suggests that these factors are not strongly predictive of these immediate outcomes and that outcomes expressed in terms of these groups will not be unduly sensitive to changes in the mix of patients treated.

Presenting variables by NHP scores

Table 7.9 summarises the relationships between the presenting variables and the NHP scores (using the combined score) for the group of stable angina patients with no other intervention to 3 months.

Table 7.9 NHP scores (mean combined scores) by presenting characteristics (n=132)
Significance tests by Kruskal-Wallis analysis of variance with ranks.

	Mean NHP scores (Combined)		
	Pre	Post	Change
Age Band			
<40	26.4	18.8	7.6
41-50	24.4	13.8	10.7
51-60	30.5	18.5	11.9
>60	23.3	16.4	6.9
	ns	ns	ns
Sex			
Male	24.3	15.2	9.1
Female	34.6	20.7	13.9
	ns	ns	ns
Previous Surgery?			
No	26.5	15.4	11.1
Yes	27.9	21.3	6.6
	ns	ns	ns
Previous MI?			
No	23.6	13.1	10.5
Yes	33.5	23.7	9.7
	p=.046	p=.037	ns
Complications			
0	23.7	11.2	12.5
1	29.1	20.3	8.8
2	31.3	26.2	5.1
	ns	ns	ns
Number vessels			
Single	26.4	17.4	8.9
Double	26.0	13.8	12.2
Triple	34.9	20.4	14.5
	ns	ns	ns
Stenosis score			
1	22.8	2.0	20.8
2	28.6	22.8	5.8
3	27.0	16.7	10.2
4	17.1	6.1	11.0
5	25.6	12.7	13.0
7	48.0	43.1	4.9
	ns	ns	ns
Ventricular Function			
Good	25.7	16.1	9.6
Poor/Mod	30.2	17.9	12.3
	ns	ns	ns

There appeared to be differences between the average NHP scores on admission according to sex though these are not all significant using non-parametric tests. In particular women tended to score higher on most dimensions and thus on the combined score. Interestingly the higher scores for women were typically observed on both admission, follow-up and change in NHP. Perhaps surprisingly there were no clear patterns in NHP scores with respect to age and certainly no statistically significant ones. It might be expected that older patients would tend to show higher NHP scores but this is not evident. However it should be noted that the age range in this group of patients

was generally lower than in other studies at the Freeman and there were few very elderly patients receiving PTCA.

When regarding the various presenting features of patients which will suggest clinically more difficult patients there was a generally consistent relationship. Patients who have had a previous operation, have had a myocardial infarction, poor ventricular function, greater vessel involvement and more co-morbid conditions or complications tend to have higher initial NHP scores (though using non-parametric tests none of these relationships in isolation is statistically significant at the 5% level). For all these variables it appears that the NHP scores at follow-up tend to be higher as well though all cases show a general improvement in NHP scores.

Differences in the scale of the change in the NHP scores were not consistent for the various presenting characteristics. Clinical complexity was associated with less change in NHP when considering previous surgery, previous myocardial infarction and the presence of comorbid complications. With all these variables the more complex cases showed a smaller net change in NHP and so one can infer derived slightly less benefit from the PTCA.

The opposite picture was observed for patients with poor ventricular function and multiple vessel disease where the clinically more difficult patients showed a larger change in NHP and so one infers a greater benefit. The score of the degree of stenosis in cardiac vessels showed no consistent relationships to any of the NHP scores.

These results do show some of the relationships one might expect between the various clinical problems and the health status as measured by the various dimensions of the NHP. However there is no consistent and statistically significant relationship to the change in NHP observed between admission and follow-up at three months. If this latter value is used as the indicator of successful outcome then it would appear that, for the purposes of comparative studies, the results suggest that standardisation for these clinical factors may not be necessary. A larger number of cases would provide more confidence in such a conclusion.

Presenting Variables by Other Outcome Indicators

Table 7.10 shows the relationships between changes in walking distance, medication and anginal pain, according to the presenting characteristics of patients. The values in the table represent the proportion of patients with a given characteristic who are observed to be better, the same or worse on each of the indicators. For example of the patients aged over 60 and using walking Distance as the outcome indicators, 58% of the over 60's are better, 36% the same and 12% have lower reported walking distances.

There appeared to be few consistent relationships which hold across all indicators. Thus considering the age categories, the older age band appeared to fare slightly better in terms of walking distance or anginal pain in that fewer cases appeared to be 'worse' at follow-up than admission when compared to younger age bands. However the older age group showed less change in the medication.

Sex did not appear to be strongly related to changes in any of the indicators, the pattern for men and women being largely the same. The biggest difference being the slightly higher proportion of women who are 'worse' with respect to walking distance at follow-up (21% for women versus 12% for men). Patients who had had previous surgery seemed to fare better with respect to walking distance and medication yet worse with respect to anginal pain (36% 'better' against 57% of those with no prior surgery). Similar results were seen comparing improvements in walking distance and anginal pain against previous myocardial infarction. The patients who had had an MI appeared slightly better with respect to walking distance (65% 'better' versus 48% of those without prior MI) however they appear to do worse in terms of anginal pain (30% better versus 63% without MI).

Patients with more complications/co-morbidities showed less change with respect to medication though a greater percentage showed improvements in anginal pain. Comparing improvements in patients with single vessel disease against those with two affected vessels, it appeared walking distance and pain indicators agree that greater

Table 7.10 Relationship between presenting characteristics and subsequent changes in walking distance, medication and anginal pain. Percentage of cases 'Better', 'Same' or 'Worse' at three months after PTCA (n=132).

	Walking Distance			Medication			Anginal Pain		
	Better	Same	Worse	Better	Same	Worse	Better	Same	Worse
Age Band									
<40	0	100	0	50	-	50	50	50	0
40-50	52	36	12	44	28	28	53	42	5
50-60	56	24	21	38	50	12	50	42	8
>60	58	37	5	21	53	26	58	42	-
Sex									
Male	53	36	12	36	43	21	55	39	7
Female	58	21	21	37	42	21	46	54	-
Previous Surgery?									
No	54	34	12	34	46	20	57	39	4
Yes	54	27	20	47	27	27	36	55	9
Previous MI?									
No	48	33	19	43	41	17	63	30	8
Yes	65	31	4	23	46	31	30	71	-
Complications									
=0	51	27	22	41	32	27	61	30	9
=1	57	35	8	35	51	14	40	60	-
=2	50	50	-	17	50	33	50	50	-
Number Vessels									
Single	49	35	16	37	41	22	44	47	8
Double	67	21	13	29	46	25	59	41	-
Triple	40	60	-	60	40	-	100	-	-
Ventricular Function									
Good	56	30	15	39	36	25	54	39	7
Poor/mod	47	42	11	26	63	11	50	50	-
Stenosis									
=1	-	100	-	33	67	-	100	-	-
=2	52	35	13	39	35	26	40	53	7
=3	50	31	20	39	42	19	50	43	7
=4	63	25	13	25	50	25	80	20	-
>4	60	40	-	40	60	-	40	60	-

improvements in both indicators are seen in patients with two vessel rather than single vessel disease. There are no clear patterns with regard to the effects of either ventricular function or the stenosis score on the indicators.

As with the NHP scores the general pattern is largely as one might expect yet using this crude analysis there are few relationships which appear to be statistically significant or send consistently strong messages across all the outcome indicators.

Conclusions on data analysis.

1. The analysis of results was complicated by the complexity of different treatment patterns that emerge in patients after a PTCA. The picture is different from that seen in cholecystectomy where there tends to be an isolated event in hospital the consequences of which are fairly easy to follow. CHD is of course a chronic disease and it is clear that PTCA does not make the problems disappear completely. The fact that patients are likely to receive additional interventions after the PTCA makes the analysis of the data considerably more complex by requiring the population of patients to be split into a variety of groups. The numbers of cases in each group tend to be small and conclusions about the behaviour of the group correspondingly less reliable.

2. The pattern of change in stable angina patients is different from that in unstable angina patients. These groups have been examined separately throughout this study and it is clear that though similar measures can be used for these two groups the expected outcomes will differ. Thus for example only a minority of patients (30%) with unstable angina will not have been readmitted for some other intervention a year after the initial PTCA.

3. The analysis of outcomes in terms of the major events following PTCA eg success, death, CABG etc. shows results more or less as expected and roughly in line with those quoted in the literature. The frequency with which these events occur in the study population is sufficient to make reporting of such events reliable and these simple descriptions of process can become a useful proxy outcome indicator.

4. The more direct measures of patient health ie NHP, angina scores, walking distance, do show significant improvements to three months following successful dilatation. For angina scores and walking distance the categorisations based on the patient's response to a simple questionnaire were crude but still sufficiently sensitive to show a net improvement. However it is clear that there are substantial proportions of patients who do not show any improvement with these indicators and a few who appear worse at

follow-up. The results suggest that one can expect about half the patients to show an improvement to three months on the two indicators. Though the NHP scores are highly skewed they do show significant improvements in the most important dimensions using non-parametric tests of significance. For the few cases whose PTCA failed, there are no major changes in these three indicators - a fact which supports the conclusion that changes are related to the PTCA rather than either random variation or some other effect.

5. The benefits of the PTCA - as measured in relation to patient health - are manifested at three months with no great changes for the better to 12 months. However the negative aspects of further events do appear up to 12 months and probably beyond and therefore follow-up to a year at least is recommended. The proportion of patients who receive no other intervention falls to about 50% of stable angina patients by 12 months.

6. Though in general different measures of patient health are related - when compared in cross section - there also appears to be considerable degrees of variability in any one measure. Thus higher scores of anginal pain will tend to be associated with higher NHP scores in the population of patients, but the relationship does not hold true for all patients. Nevertheless statistically satisfactory relationships can be seen between NHP scores, reported walking distances and pain scores and suggest these indicators are demonstrating some degree of convergent validity. The relationships between longitudinal changes in these indicators tend to be much weaker.

7. There is no consistent link between the potency of medication and the other health indicators. This is probably because health status and medication are in some form of equilibrium, as pain increases so the medication should increase to counteract this and so lead to lower pain. The problem in this case is that the potency of medication is really a process measure and not particularly useful as a proxy for health. The results have suggested that medication has not changed to the degree expected at three months though changes at 12 months are observed for some patient types. This is as expected given the current policy at the Freeman Hospital.

8. The relationships between presenting characteristics and eventual outcome appear rather weak. The most significant observation is the relationship between more disease pathology, complicating conditions (risk factors) and failure of the PTCA procedure. Some differences between NHP score and sex have been observed.

D. Review Process

Though a number of meetings were held with the main cardiologist concerned only two wider meetings were held to discuss the results. The review process has been handicapped by the lower then expected rate of PTCAs which increased the time taken for sufficient numbers of cases to emerge. This has been coupled with the necessity to analyse distinct groups of patients separately, so reducing the numbers in any one group.

Nevertheless the clinical staff concerned have expressed an interest in the results and are keen to continue with data collection if possible. Some of the main areas of interest have been:

1. Discussions have been made of the relative value of the usual outcomes or endpoints for clinical trials - mainly process measures and deaths - versus the more patient centred measures - pain relief, health status etc. Though the NHP has performed well it has been less well accepted than in cholecystectomy. In terms of the relative impact of the information the traditional process measures were considered more useful and easier to understand.

2. The reactions to the results were mixed and though two of the indicators, deaths and failed PTCAs, were known in advance - there was a genuine uncertainty over what outcomes to actually expect. In particular, limited comparisons to the literature have been used to place the results in some context - though this is not available for the more sophisticated measures. There has been interest in comparing the results to a similar study of CABG patients that has been carried out recently at the Freeman Hospital. In fact some of the data have been designed to be compatible.

3. The lack of change in the medication was greeted with some surprise and it was suggested that perhaps clinicians were being a little conservative in their use of drugs following PTCA. The cardiologists are considering reducing the potency of drug regimes at an earlier stage than at present. This is one area where a possible clinical change may result.

4. There were some doubts expressed over the value of using the patients' perception as a way to measure outcome in this group. More specifically there is felt to be a wide variation in how patients cope with their condition which may not be in proportion to the severity of their disease. This question concerns not just the process of assessing health status by asking the patient but also the variability between patients in the way that they react to what in one sense can be considered essentially similar health states. Opinions were divided on the question of whether the selection criteria for patients (PTCA or other treatments) include some assessment of those who were felt to be 'psychologically' more likely to benefit from the procedure?

5. One of the key areas of debate was the process whereby patients were selected for PTCA and this was felt to be the most important determinant of the outcomes. There was correspondingly relatively little debate over the specific process of care adopted for patients. Thus, as with the cholecystectomy study, the data has served to focus attention on one key issue which is felt to be the most important in improving outcomes of this particular procedure.

Chapter 8 Other Specialties
Orthopaedics, Rheumatology, Urology

In addition to the main specialties the study has undertaken some work in other specialties. In orthopaedics a study of knee replacements has been underway for some time and the work has developed along the lines of that seen in cholecystectomy and angioplasty. In rheumatology there have been discussions about the dataset and a number of ad-hoc studies designed to help the rheumatologists with their audit process. In urology the work has been limited to a discussion of the outcome data set for an examination of prostatectomy.

This chapter gives a brief overview of the work in these specialties and a limited analysis of the results to date. They have been included because they have provided useful insights into the main project and how its findings are applicable to other specialties and contribute to the results of the study as a whole.

A. Orthopaedics - Knee replacements.

Work in orthopaedics originally began because the area was of keen local interest and because of the proposed development of a local clinical information system. It was recognised from the start that the commitment of the study team would have to be limited. The case type chosen was knee replacements - with the likelihood that hip replacements would be examined later. Though there are a number of consultant orthopaedic surgeons at the Freeman Hospital, only two carry out knee replacements and these have been involved in the development of the outcome indicators and the review of the results.

Outcome indicators

1. Knee function

There is a standard method of assessing knee function before and after knee replacement

that has been used for some time and has formed the basis for studies of the relative effectiveness of different prostheses (Insall et al 1976). The assessment made by the doctor is based on the doctor-patient interview and includes basic questions on the degree of pain and patient function, as well as measures of the degree of damage to the affected joint. The scoring system is rather arbitrary with a maximum score of 100 , the higher the score the better the knee function. Both surgeons used this structure in assessments of the patients therefore a standard pro-forma was adopted which included the items of the knee score for completion during the patient interview.

2. Improvement in health status.

It was agreed that in addition to improving knee function, there was also the objective to improve health from a wider perspective. Early on in the study the intention was to identify datasets that would be compatible between rheumatology and orthopaedics. The two specialties work very closely together at the Freeman sharing resources and patients. Given this commonality the first choice of a general health status measure was the health assessment questionnaire (HAQ) which has come to be something of a standard in rheumatology (Fries 1983; Kirwan & Reeback 1986; Leighton-Read et al 1987). However it was realised fairly early on that this questionnaire did not cover the main problems experienced by the knee patients. In particular it included a variety of questions concerned with the upper body and did not include more general psychological or social functioning. It was therefore decided to try the Nottingham Health Profile which had been proving satisfactory in the other specialties. This has been the instrument of choice for the rest of the study.

3. Peri-operative complications

As with other surgical studies, a simple form was used to check for problems during the procedure or anaesthetic. Free text was used in this case rather than pre-specified responses.

4. Post-operative complications

With any surgical procedure there are a number of problems that may arise post-operatively. A list of potential problems was agreed and included:

- Wound infection
- Post-operative bleeding
- DVT/ Pulmonary embolus
- Respiratory infection /complications
- Cardiovascular complications
- CNS complications
- Urinary complication (eg retention, infection)
- Septicemia
- Renal failure
- Other (specify)

5. Joint loosening - X ray evidence

One of the problems that can occur following knee replacement is that the prosthesis becomes loosened. The assessment of this is based on a very simple scoring of the evidence of translucent zones revealed by post-operative X-ray. The scale and position of any translucent zones was noted before the operation and at subsequent follow-ups. The recording is based on simple diagrams of the joint used by Goldberg and colleagues (Goldberg et al 1988). A serious increase in these zones is an indication that the joint is loosening.

Other data items

The intention that the orthopaedic study would eventually contribute to the development of a local clinical database meant that the scale of additional information collected was larger than needed purely for outcomes measurement. In particular a number of details concerning the patient before the operation are collected. It is intended that the relationships between these factors and eventual observed outcomes are explored in some depth.

Key items recorded include:

Administrative - Hospital number, name, address etc.

Indications for surgery

Extent of arthritis

Previous surgery

Concurrent conditions

Type of prosthesis

Knee detail - Lachman's Test, pain.

Pre-operative X-ray - Pathology of knee disease etc.

Post-operative X-ray

Timing of observations

The timing of observations for knee replacements is rather different from that in other specialties with longer time scales being considered more appropriate. This stems from the evidence on various joint replacement procedures that the prosthesis itself will eventually have to be replaced, typically the time scale for hips being 10-15 years. Goldberg and colleagues (Goldberg et al 1988) followed 82 patients who had had a condylar knee prosthesis for an average of nine years and found that ten had had a revision for various reasons. Thus the Freeman surgeons were interested in knowing how long their joints would stay in place as well as the immediate benefits of the procedure.

For the purposes of continuous monitoring of outcomes prospective studies over 10-20 years are unlikely to have much managerial impact. Therefore benefit over shorter time scales were measures using three months (or nearest out-patient visit) and twelve months after the procedure.

Data Analysis

This data analysis section records only brief descriptions of some of the results rather than a comprehensive examination of all the relevant variables. The selection of issues in this report is based on those that have developed during the course of the study and discussions with the orthopaedic surgeons.

1. Changes in Knee and NHP scores.

Table 8.1 shows the changes in the knee score and NHP scores to 3 months, and to

twelve months for the smaller subset of patients who have reached that stage.

These results show the changes for patients who have received a single operation only and give a clear idea about the sensitivity of the measure to changes in patients health.

Table 8.1 Changes in Knee Score and NHP scores to 3 months and 12 months

	Mean scores to 3 months n=137			Mean scores to 12 months n=43			
	Before	+3mnths	Sig	Before	+3 months	+12 months	Sig
Knee Score (sd)	44.7 (13.6)	72.9 (17.4)	.000	46.4 (13.1)	75.7 (12.4)	79.7 (14.6)	ns
Energy	41.5	29.1	.000	45.7	24.6	19.8	ns
Pain	64.1	33.4	.000	71.1	35.2	30.5	ns
Emot.Reacs	22.5	11.7	.000	20.7	10.5	11.5	ns
Sleep	48.1	35.4	.000	63.2	47.4	42.1	ns
Soc Isol	14.6	7.1	.000	14.8	3.3	5.1	ns
Mobility	46.1	31.8	.000	47.7	29.6	31.4	ns
Part II	42.7	29.0	.000	37.4	23.6	23.6	ns
Combined	39.5	24.7	.000	43.9	25.1	23.4	ns

The knee scores show significant improvements after the procedure typically moving from an average of about 45 before the operation to 70-80 afterwards. These changes are roughly in line with those observed by Insall et al (Insall et al 1976). For the subset of cases who have reached the twelve month follow-up (without revision) there is only a small change in knee score beyond three months.

Similarly the NHP scores on all dimensions show significant improvements. Fig 8.1 shows the distributions of the six dimensions of the NHP before the procedure. The highest scoring dimension is 'Pain', followed by 'Mobility', 'Energy' and 'Sleep'. 'Social Isolation' and 'Emotional reaction' score lower, with highly skewed distributions, though the mean values are still above zero. Three months after the procedure the scores are all significantly lower. The residual scores at three months however are still above zero particularly for 'Pain', 'Energy', 'Mobility' and 'Sleep'. These results suggest that there is a marked improvements related to the procedure but that the patients still have quite significant health problems afterwards. The scores on the NHP are fairly high both

before and after the procedure - when compared to other specialties.

2. Correlations between knee score and NHP

Table 8.3 shows the correlations between the knee scores and the dimensions of the NHP. Interestingly only pain and mobility (and the combined score) are significantly related to the knee score for the pre-operative assessment. At follow-up only energy and mobility (and combined) score are correlated. There are no significant correlations between the change in NHP and the changes in knee score.

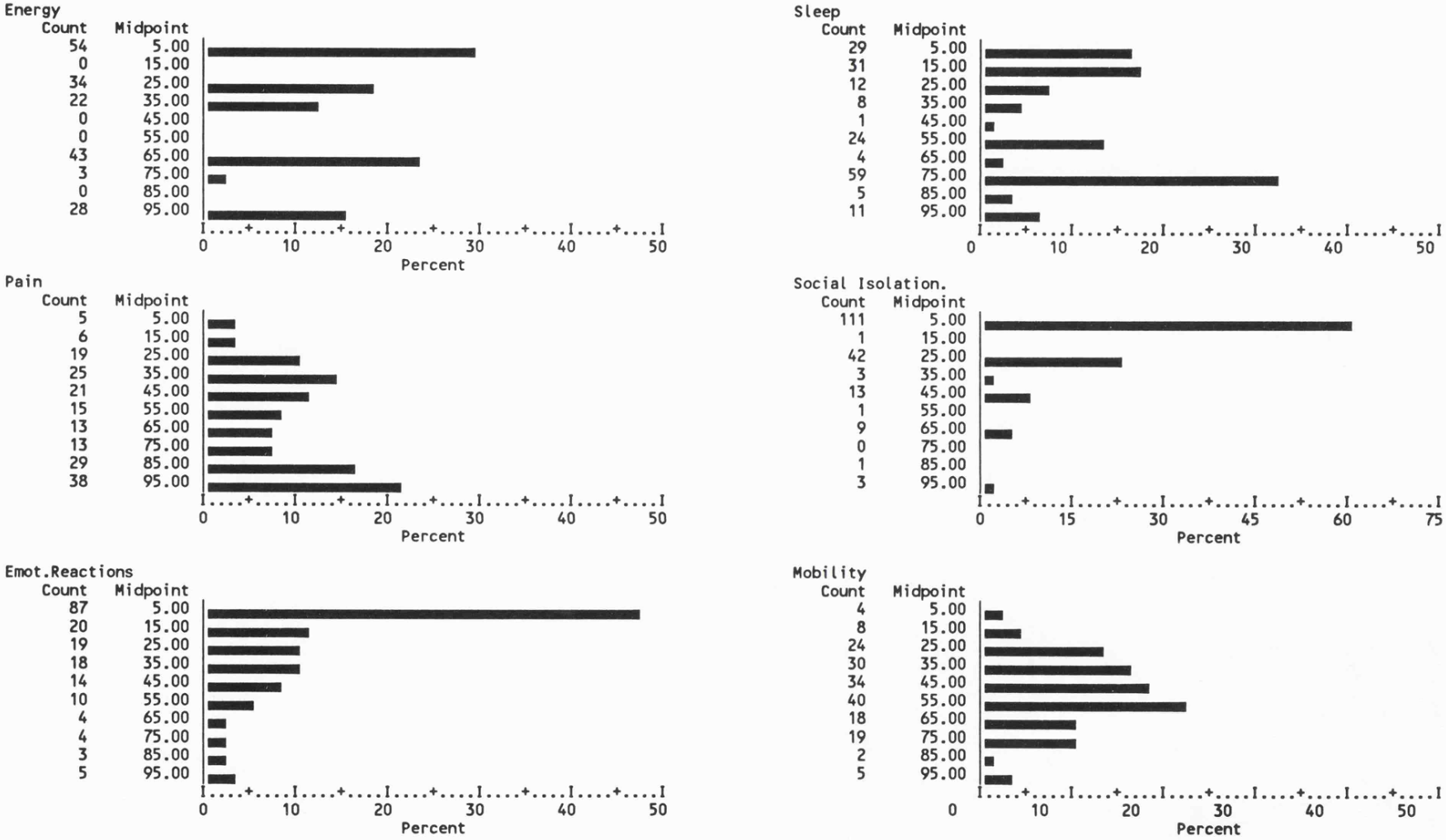
Given the earlier similarities between the way in which the two sets of scores behave after the procedure this lack of strong links between the two is a little surprising. However it is clear from the content of the two scales that they are often measuring different things. In terms of content the pain element in the knee score is quite large and some correlation with the NHP is therefore observed for this element (though the weightings used in calculating the knee scores are rather arbitrary). However apart from this the score does focus on the pathology of the knee itself - a rather narrow view of the patient. The NHP on the other hand incorporates the far broader picture of general health.

Table 8.2 Correlation (rank) between NHP scores and Knee scores (n=123)

	Knee Scores		
Correlations:	Initial Score	Follow-up Score	Change Score
Initial NHP			
Energy	-.2047	-.2586*	-.0816
Pain	-.3399**	-.1945	.0509
Emot.Reac	-.1856	-.0800	.0486
Sleep	-.1534	-.1628	-.0365
SocIsol	-.1305	-.1490	-.0395
Mobil	-.2708*	-.2486*	-.0337
PartII	-.0966	-.1738	-.0796
Combined	-.3200**	-.2792**	-.0283
Change to 3m			
Energy	.0112	.0497	.0327
Pain	-.0565	.1190	.1286
Emot.Reac	.0407	-.0226	-.0425
Sleep	-.0282	.0332	.0434
SocIsol	.0479	-.1406	-.1405
Mobil	-.0444	.1274	.1278
Part II	.0920	-.0493	-.0946
Combined	-.0123	.0536	.0500

1-tailed Signif: * - .01 ** - .001

Fig 8.1 Distributions of NHP scores before admission for knee replacement (n=184)



3. Effects of age and type of arthritis.

Table 8.3 shows the observed changes in both NHP by age band and for osteoarthritis and rheumatoid arthritis separately. With regard to the age, the results suggest that the younger age band, under 60 years of age do not do as well as the older age groups when the NHP is examined. The observed scores for the younger age band before the operation are not greatly different from the older groups. There are no differences across age groups in terms of the change in knee score. A number of explanations for this have been offered during the study and none have been found to be wholly satisfactory. It is clearly due to some form of interaction between overall health perceptions, and possibly expectations, being different with age. Burton et al have suggested that patient expectations from hip replacements are an important part of the assessment of outcomes and may not necessarily correlate with the technical success of the procedure (Burton, Wright & Richards 1979). The results are being investigated further while the issue of whether differences in expectations can (and should) be met from within existing technologies and resources needs to be addressed.

Finally, Table 8.3 also shows that the benefits of surgery, both in terms of the knee score and overall health status, are not significantly different for patients with osteoarthritis and those with rheumatoid arthritis. This differed from the expected results that rheumatoid

Table 8.3 Initial and change in Knee score and average change in NHP for patients according to age group and type of arthritis. ns=not significant p>.05

	Differences by age			Difference by arthritis	
	Age<60 n=45	60-70 n=51	>70 n=68	OA n=104	RA n=55
Initial Knee	44.9	41.2	45.6 ns	44.4	44.2
Change Knee	27.3	31.3	24.7 ns	28.8	27.1
Energy	-1.4	13.6	21.2 .049	10.7	16.8
Pain	23.0	40.1	41.1 .017	37.9	33.2
Emot. Reacs	8.5	13.4	19.5 ns	14.3	14.9
Sleep	15.8	16.1	26.4 ns	21.7	17.3
Soc Isol	7.1	7.3	14.4 p=.050	9.9	7.1
Mobil	11.0	11.1	26.9 p=.027	19.4	14.3
Part II	17.3	25.2	18.4 ns	21.7	12.5
Combined	6.9	16.3	18.6 p=.039	14.4	15.1

patients will not benefit as much since they would have more serious longer term

problems which may overshadow the benefits gained in a single joint.

Using the information

Data on outcomes has been presented to the relevant clinicians in the form of simple reports as in the other specialties. These reports have been discussed in meetings as in other specialties. Discussion have included :

- issues of data collection
- results to data
- validity of the various measures used
- effects of presenting characteristics
- individual case review of 'non-responding patients'

The two consultants concerned, and their senior registrar when relevant have shown a commitment to the project and in fact have been responsible for most of the basic data collection. The research team has been responsible for co-ordinating the data collection and chasing up when necessary but the input in terms of resource has been considerably less than in other specialties. Nevertheless the clinicians have shown interest in the results and analysis and found them useful and their comments n teh evaluation form included (Appendix 3):

> *"As yet in orthopaedics still not at our one year objective which will be of value. We have certainly learnt the value of measurement but require further analysis to assess our conclusions"*

> *"Enables detailed review after a busy clinic with highlighting of problems"*

Though the NHP was at first regarded with some suspicion, it has proved itself to be an important part of the assessment of patient outcomes and has come to be accepted by the clinicians themselves. Once again the validity of the instrument in monitoring changes in this type of surgery is an important finding of the study. It is now accepted that the

wider view of health offered by such measures is critical in the assessment of outcome.

There has been some conflict between the demand for basic information over long time scales and in particular the expected life of prostheses, which was of particular interest to the clinicians, and the shorter term reviews necessary to complete an audit cycle as advocated by the research team. Thus the results to date are only seen as addressing outcomes in the short term.

The clinicians were already familiar with the knee score and its interpretation raised fewer problems in the early stages. Thus both surgeons were aware of the expected range of scores to be seen before the operation. However the scale of observed change to follow-up was not known locally. The improvement has been reassuring and conforms to figures quoted in the literature. Moreover simple diagrams showing the change in score made it easy to identify those patients where improvements had not been for the better. These patients were reviewed individually using all the available data.

Similarly, patients who had poor outcomes either in terms of the NHP or post-operative problems were reviewed individually. Though discussion of these individual patients was interesting there were few general lessons about how practice could be improved in future. A variety of explanations were offered but the two most common reasons were:-

- that patients had severe disease in a number of joints (or other health problems) which were limiting the scale of improvement.
- that, though the operation was 'technically' successful, patients were not willing to either exercise or test the knee, or change their life-styles. It was suggested that some form of screening to identify these patients in advance be undertaken if one wished to secure the maximum benefit from the procedure.

The review of data in orthopaedics cannot point to any specific changes in practice that have yet resulted. However clinicians have expressed their greater awareness of a number of issues. Thus the observed, and as yet unexplained, relatively poor

improvement in the younger age band is causing some concern, as is the question of selecting those patients who are most likely to benefit. The observation of the relative improvement in the rheumatoids has also brought comment and in effect confirmed their current practice.

Conclusions

The study of outcomes in orthopaedics has shown that, with commitment from the clinical staff, it is possible to link in data collection for outcome measurement with existing processes of data collection. The scale of involvement for the research team in this study is limited to the co-ordination and analysis of data. The results have been encouraging both in terms of the changes in outcome indicators that are observed after surgery and the behaviour of the instruments chosen. It appeared that though both the traditional Knee scoring system and the NHP show improvements after the operation, they are describing different types of health benefit that the patients received.

B. Rheumatology - Rheumatoid arthritis

The work undertaken in collaboration with the rheumatologists, is another example of where the clinicians shared an interest and enthusiasm for research into outcome measurement. The study has become involved in a number of pieces of work with the clinicians which have explored possible outcome measures in rheumatology. The work has been rather opportunistic and has been limited by the resources that the research team could devote. Nevertheless some interesting results have been revealed and useful issues addressed.

There are five consultant rheumatologists shared between the Freeman Hospital and the Royal Victoria Infirmary. In addition to out-patient sessions, which is the focus for most work in rheumatology, they have some beds at the Freeman and also work closely with orthopaedic surgeons.

Outcome measurement in rheumatoid arthritis

In the early stages discussions were held with the clinicians to agree a basic data set for outcome measurement in rheumatology. The patient group selected was rheumatoid arthritis which comprises the bulk of the specialty workload.

Rheumatoid arthritis is a chronic disease that presents some specific problems for outcome measurement. The longer term outcomes of treatment of rheumatoid arthritis are not especially promising and results after 20 years have shown a general deterioration in patient function (Scott et al 1987). Though slightly higher mortality rates are observed in rheumatoids (Symmons 1988) the main effects of the disease are pain and a progressive loss of function. The measurement of outcomes in rheumatoid arthritis is complicated by a number of factors:-

i. The disease progresses over fairly long time scales and loss of function can be slow and so the measurement of improvement or maintenance of function is correspondingly more difficult.
ii The disease typically exhibits periods of short-term improvement or deterioration.
iii The causes of the disease are largely unknown, and the links between short-term clinical measures and longer term health outcomes difficult to trace (McKenna 1988)
iv Rheumatoid arthritis itself can take a variety of forms , and is related to a number of other similar inflammatory disease, it may be that the term covers a whole family of more specific case types (Woolf 1988).

Despite, or possibly because of, these problems, a considerable amount of work has been done in rheumatology to explore outcome measurement (Fries 1983; Thompson 1988). In particular, the study of the effects of the disease on everyday function and quality of life is well established (Deyo 1988). A simple four point classification to describe the degree of functional impairment has been in use since 1949 (Steinbrocker, Trager & Batterman 1949).

The Health Assessment Questionnaire (HAQ), developed by Fries and colleagues (Fries et al 1980) and later tested in the UK (Kirwan & Reeback 1986) is a simple scale for measuring disability in arthritis and has come to be something of a standard in the field. Other measures have also been widely used including the arthritis specific 'Arthritis Impact Measurement' (AIMs) (Meenan, Gertman & Mason 1980) and the general Sickness Impact Profile (Bergner at al 1981). There are also examples of the use of such scales in clinical trials (Meenan et al 1984; Bombardier et al 1986; Speigel et al 1986). Comparative studies have previously suggested that in this group of patients different measures are highly correlated and though favourites can be chosen they behave in similar ways (Liaing et al 1985; Fitzpatrick et al 1989).

There is therefore a well-established field of work in measuring these issues with respect to rheumatoid patients. There is in addition a number of clinical and laboratory measures which are accepted as standard assessment tools. These include:-

i Laboratory measures include haemoglobin levels, rheumatoid factor and erythrocyte sedimentation rates (ESR) - the latter being voted in one study *'the most reliable single indicator of disease severity'* (Bull et al 1989). However there was little evidence of the link between this score and patient health outcomes and such process measures have been described as having no inherent value to the patients or society (Fries 1983).

ii Measures of specific function eg grip strength, walking time, morning stiffness are widely used and tend to be specific to certain joints and subject to some unreliability in measurement.

iii Measures of the number and severity of painful,tender and swollen joints often exist in the form of an index, for example the Ritchie index (Ritchie et al 1968)

iv Radiographic evidence is sometimes used, for example with scoring systems. However these can be very labour intensive.

Combinations of these measures have been used to classify patients according to the severity of disease (Dawes et al 1989) or to measure specific disease activity (Mallya & Mace 1981).

Discussion of these various measures were held with the clinicians and a basic data set drawn up. One of the basic problems was the gap between the short term measures of disease activity which tend to be couched in terms of laboratory and clinical measures and longer term measures of outcomes which reflect patient's function and health. For a full evaluation of outcome both types of measure were felt to be important. However the collection of data on disease activity is time-consuming to collect. It was recognised that additional resources would be required if this data set were to be used as in the other specialties.

Cross-sectional comparisons of health status measures

Though the HAQ has come to be something of a standard measure in rheumatology, it still has some problems. One is the fact that the results may be affected by disease of the upper limb more than the lower limb. A second problem was that if aids were given to patients, the scores automatically went up (using one scoring system) even if patient function was the same. In addition a concern was expressed over the relative lack of sensitivity of the score to changes in very severely ill patients. It was possible that some of the alternative general health status measures would prove more useful. Therefore a study was undertaken, as a pilot project to examine the use of three measures (the HAQ, the SIP and the NHP) in an outpatient population of rheumatoid patients.

More specifically the aims of the study were to :

i. Test whether the instruments could be administered to an outpatient population.
ii. Compare the distributions of scores obtained and in particular examine the ranges of scores obtained on the SIP and NHP scales for patients with high values on the HAQ.

Questionnaires were distributed via the outpatient clinic to 100 consecutive patients and

returned by post in pre-paid envelopes. In total 80 sets of forms were returned (there was no chasing of non-responders).

Briefly, the results confirmed that such questionnaires could be given out by staff and would be returned by patients with satisfactory response rates. For one-off cross sectional studies the methods were satisfactory. It was clear from the scores returned that all three measures were sufficiently sensitive to detect health problems in these patient groups and that the mean scores on most dimensions were well above those expected in a 'healthy' population (results of the NHP are discussed with those obtained from other specialties in Chapter 9).

All three scales show generally high correlations with each other. These relationships were strongest between the HAQ and those dimensions of the NHP and SIP which measured basic function (Table 8.4(a)). In particular 'Pain' and 'Mobility' in the NHP and 'Ambulation' and 'Self-care' in the SIP. Thus with respect to the basic physical/functional dimensions of health the measures were in broad agreement.

It was also clear that patients were scoring high on dimensions of the SIP and NHP that were concerned with social and psychological problems, dimensions of health that are not covered by the HAQ. Table 8.4(b) shows the high internal correlations between the SIP categories of 'Social interaction', 'Alertness' and 'Emotional Behaviour' and the NHP categories of 'Social Isolation' and 'Emotional Reactions'. The correlations between these dimension and the HAQ are noticeably poorer. Thus it appears the HAQ is failing to pick up significant problems that these patients had in these respects.
There was little evidence to suggest that the NHP or SIP dimensions were more sensitive, and more capable of showing improvement, for patients with high HAQ scores. In particular it was felt that at the high end the HAQ scale the score was insensitive to improvements in patients health therefore the variability in NHP and SIP scores were compared for patients with low HAQ scores against those with high scores.

The analysis of the distribution of the HAQ and SIP scores for patients with high HAQ showed no greater range of values than for patients with a low HAQ score. Thus there

was no evidence to suggest that, in this population, the NHP or SIP would be capable of detecting any more improvements in severely disabled patients than the HAQ.

Table 8.4(a) Correlation coefficients between HAQ and selected dimensions of NHP and SIP dealing with basic physical function (n=80).

NHP dimensions: Pain, Mobility (Mobi)
SIP categories: Self care & movement (S.Care), Ambulation (Ambl)

	----NHP-----		----------SIP----------			
	Pain	Mobil	Total	S.Care	Ambl	Mobil
HAQ	.47	.66	.59	.69	.57	.59
NHP Pain	-	.60	.63	.54	.49	.44
NHP Mobil	-	-	.67	.66	.60	.58
SIP Total	-	-	-	.84	.71	.83
SIP S.Care	-	-	-	-	.69	.69
Sip Ambl	-	-	-	-	-	.64

Table 8.4(b) Correlation coefficients between HAQ and selected dimensions of NHP and SIP dealing with social and psychological function (n=80).

NHP dimensions: Emotional Reactions (EmR) and Social Isolation (SocIs)
SIP categories: Social Interaction (SocInt), Alertness (Alert), Emotional Behaviour (Emob), Communication (Comm)

	----NHP-----		------------SIP-------------------				
	EmR	SocIs	Total	SocInt	Alert	Emob	Comm
HAQ	.22	.21	.59	.27	.05	.36	.16
NHP EmR	-	.73	.56	.53	.46	.52	.56
NHP Soc Is	-	-	.49	.38	.36	.35	.58
SIP Tot	-	-	-	.76	.56	.69	.55
SIP SocInt	-	-	-	-	.60	.62	.33
SIP Alert	-	-	-	-	-	.51	.34
SIP Emob	-	-	-	-	-	-	.50

One explanation for this may be that the outpatient population is, relatively speaking, at the less disabled end of the spectrum of rheumatoid patients. Later studies were therefore undertaken to address, among other things, whether the same results were obtained with an inpatient population.

In-patient follow-up study.

At about the same time this study started, the rheumatologists themselves were examining possible audit procedures for the specialty. They developed a fairly sophisticated system of identifying goals for individual patients and recording whether these goals were met at discharge. The data, along with a comprehensive description of the patients diagnosis, was collected on a locally developed computer database.

It was agreed that the research team would undertake a follow-up of these patients approximately one year after discharge. This would be based on a simple questionnaire sent to patients that would include the HAQ. The clinical team used the opportunity to ask patients about the stay to explore some issue of patients' satisfaction with care and their view of the value of the in-patient stay. The aims of this study were to:-

i Test whether improvements in hospital were matched by improved function at follow-up.
ii Validate changes in the HAQ by comparison to other questions on changes in health and any problems post-discharge.
iii Survey patients' satisfaction with the inpatient admission
iv Examine the changes in HAQ after discharge against the goals achieved in hospital.

The questionnaire (Appendix 9) therefore included an additional page with the following series of questions:

- Do you feel your stay in hospital improved the state of your health in general?
- Do you feel your stay in hospital improved your arthritis?
- How well controlled do you feel your arthritis is now?
- How often have you felt the level of pain you experience has been unacceptably high?
- How well do you feel you understand your arthritis now?
- What do you see as your main problems due to your arthritis?
- Do you have any health problems, other than those due to your arthritis?
- Which social services do you receive?

Patients were contacted using the addresses supplied by the clinical database.

Results

The results of this exercise have yet to be studied in full and the following paragraphs only describes those issues discussed to date.

Survival

There were a greater number of patients who had died since admission than expected.

The exercise revealed a number of administrative problems in identifying when patients had died and local information systems were unaware of the fact.

The observed death rates were far higher than regional age and sex related mortality rates for the general population would suggest and were a great surprise to the clinical staff. Though life expectancy with RA is slightly reduced (Symmons 1988) this would not explain the observed incidence. This simple observation of how many patients had died within a year of discharge prompted a more specific process of case review on the individuals concerned. This was discussed at a rheumatology audit meeting. As a result of that review it was recognised that there was a need to examine the care of cervical myelopathy in the district and a review is underway. This is one example of where a simple piece of information on what happens to patients after they leave hospital has been used to examine the quality of care and promote practical changes as a result.

Changes in HAQ Score

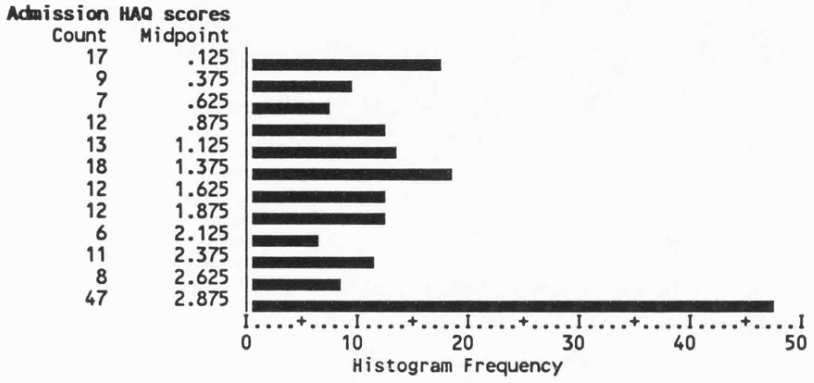
The results showed that during the inpatient stay the average HAQ score does not show any significant change between the initial value of (1.75) and the final value of (1.72). Though the average of the scores remained the same approximately 36% of cases showed better scores and 27% worse scores from admission to discharge.

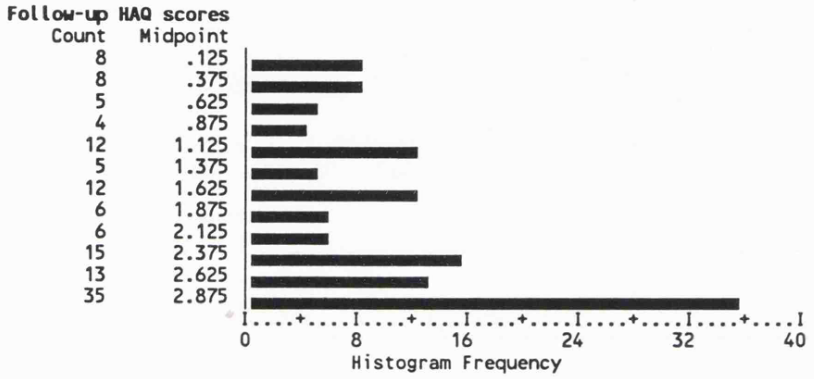
When the follow-up scores were examined there was a significant worsening of HAQ scores from 1.65 to 1.85 ($p=.003$). In this group 31% of patients are better and 51% worse at follow-up. Fig 8.2 shows the distribution of initial and follow-up HAQ scores, and the changes in-between.

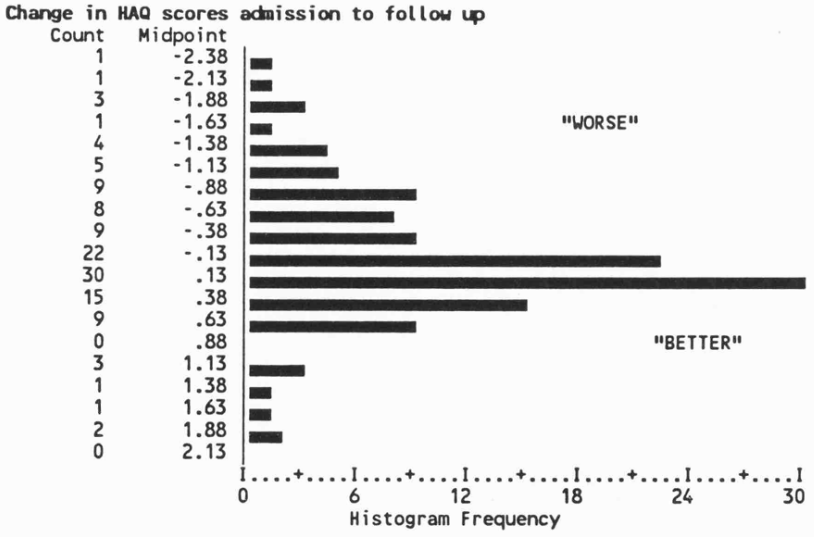
General Questions

Patients were asked a series of general questions about the hospital stay and the problems they faced. A majority (65%) felt that the hospital stay had improved their health, whilst slightly fewer (52%) felt the stay had improved their arthritis. Though the answers to these questions are related there were a proportion who felt that the stay improved their general health but not their arthritis. When the changes in HAQ

Fig 8.2 Distributions of HAQ scores at admission, at follow up after one year and the change from admission to follow-up.







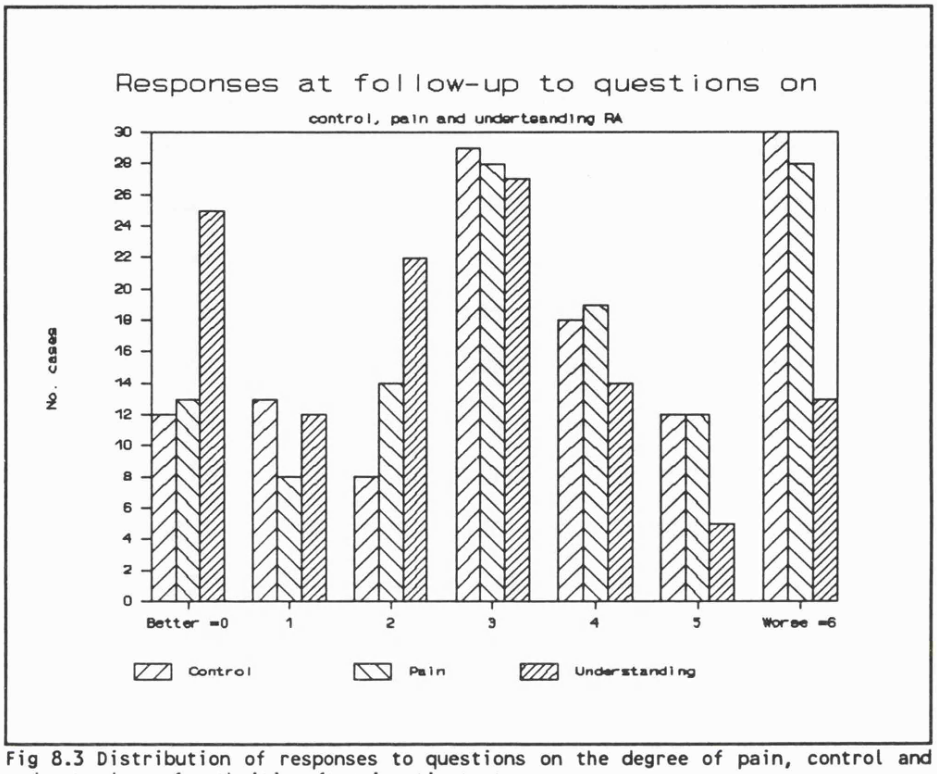


Fig 8.3 Distribution of responses to questions on the degree of pain, control and understandung of arthritis after inpatient stay.

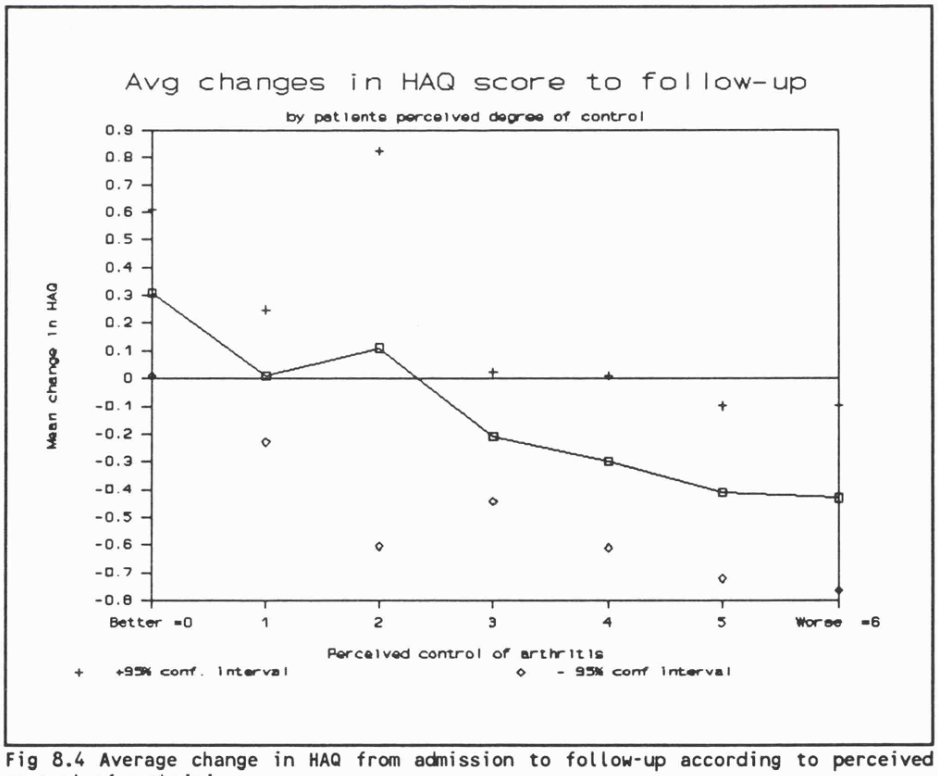


Fig 8.4 Average change in HAQ from admission to follow-up according to perceived control of arthritis.

are examined against these responses the findings broadly agree in that patients who felt the stay had not improved their health or arthritis showed greater average increases in HAQ score. Patients were also asked to grade (on a scale of 0 to 6) whether they felt their arthritis was better controlled, whether they understood their condition better, and whether their pain was better. The replies to all three questions showed a broad distribution of scores involving both extremes of the scale. The patients views on whether their health was any better largely mirror the changes observed in the HAQ, a large proportion replied that they were much worse with respect to control of their arthritis and pain (Fig 8.3). However the responses did show that patients tended to have a better understanding of their condition at follow-up. The responses to the questions on pain and understanding of arthritis were related, patients who felt they understood the disease scored lower for pain.

The perception of control of arthritis was found to be related to the change in HAQ score. Patients who felt the disease was better controlled showing, on average, an improvement in HAQ whilst those who felt worse controlled showed larger reductions in HAQ (Fig 8.4)

Comparison of health status measures on in-patient population.

The study of health status measurement on in-patients was prompted by the clinical staff for three reasons:-

1. The earlier study of outpatients had failed to show the expected lack of sensitivity at the extremes of the HAQ. The comparison of inpatients would address a similar question for a more severely disabled patient group.

2. Comparison of NHP scores derived from the rheumatology outpatient study with NHP scores for other conditions had shown relatively high scores for RA patients. This prompted questions over how inpatients would score on the NHP and whether changes in health status could be observed.

3. Nurses on the orthopaedic ward had expressed an interest in the evaluations provided by the NHP and suggested they were useful in gaining an overall picture of the impact of the disease on the patient's everyday life. It was thought possible that nurses in Rheumatology may also find the instrument useful in this respect.

Therefore a limited study of consecutive RA patients was undertaken using both the HAQ and NHP (the SIP was rejected on the grounds that it was too long and produces results very similar to the NHP in the earlier study). The ward clerk was responsible for identifying the relevant patients and handing out and collecting the questionnaires. Completed questionnaires were located centrally in the ward. Some concerns had been expressed about the patients views on the confidentiality of the information. Therefore an initial pilot study of ten patients was undertaken with the research team interviewing patients after completion of the questionnaire. There were no cases when patients expressed concern for the confidentiality of the data and they were all happy for the data to be collected and stored centrally.

In addition to the questionnaire clinical details of the patients were drawn from the local computer data bases to examine the effects of diagnosis, age and sex on the observed scores.

Recruitment of patients was stopped after 100 cases. A follow-up questionnaire was sent to patients 3 months after discharge. and this data is now being returned. The results have not as yet been discussed with clinical staff.

Early Results

With regard to the use of the NHP by nursing staff, it appears that, unlike in orthopaedics, there is little if any use made of the questionnaires.

The distributions of the NHP scores are shown in Fig 8.5. It is noticeable that these distributions were not skewed for four of the six dimensions - with the exceptions of 'Emotional Reactions' and 'Social Isolation'. The early results showed high correlations

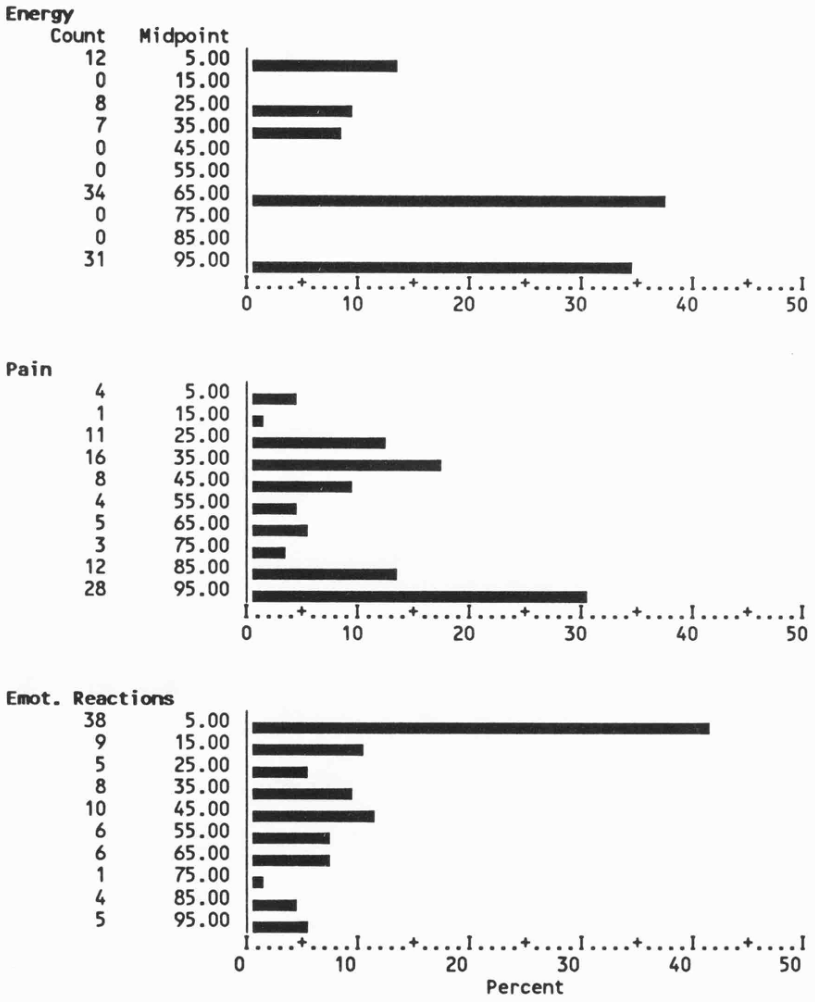
between the HAQ scores and the NHP dimensions when comparisons were made between the initial scores, the follow-up scores and the changes in scores. The largest correlations were seen for the NHP dimension of 'Mobility' at admission and follow-up where correlation coefficients over 0.7 were observed and the scores for 'Energy' where correlations are greater than 0.4. The weakest correlation are for 'Social Isolation' and 'Emotional Reactions' as observed in the outpatient study. Though correlation coefficient between NHP and HAQ scores are lower when the changes in score are compared they are still highly significant and suggest that changes in one score are matched, by and large, by changes in the other.

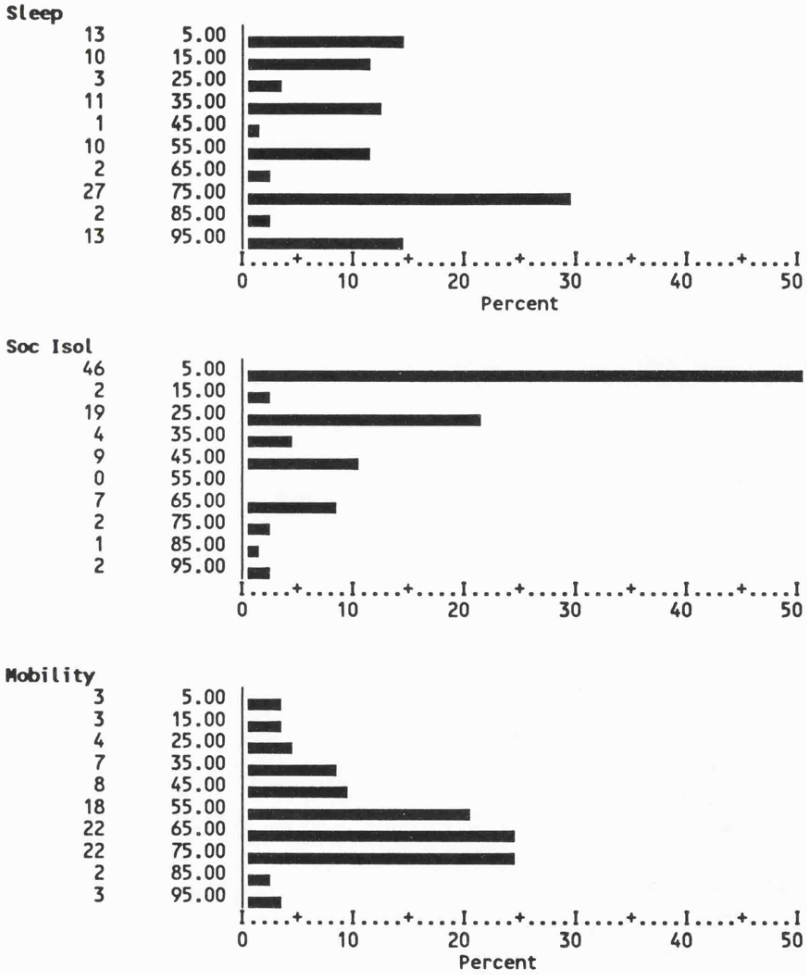
Table 8.5 Correlations between NHP, HAQ and changes between admission and three months. Significance levels shown underneath correlation coefficients (n=92).

	Initial NHP vs initial HAQ	Follow-up NHP vs follow-up HAQ	Change in NHP vs change in HAQ
Energy	.4192 p= .000	.4224 p= .000	.3526 p= .000
Pain	.2785 p= .004	.4802 p= .000	.3206 p= .001
Emot.Reacs	.2693 p= .005	.2371 p= .011	.3898 p= .000
Sleep	.2563 p= .007	.3359 p= .001	.1736 p= .049
SocIsol	.3078 p= .001	.2034 p= .026	.2604 p= .006
Mobil	.7477 p= .000	.7651 p= .000	.3551 p= .000
PartII	.1124 p= .143	.4152 p= .000	.3468 p= .000
Combined	.5015 p= .000	.5493 p= .000	.4546 p= .000

Both scores show the same picture with regard to the average change between admission and the three month follow-up, namely that there is no overall change in the population means. Further analysis of this data is being undertaken.

Fig 8.5 Distributions of NHP scores for In-patient Rheumatoid Arthritis. n=92





Conclusions - Rheumatology

The various pieces of work undertaken in Rheumatology have been largely explorative. The clinical team have a well established pattern of audit and this work has tried to help that in a number of ways. The studies of alternative health status measures have broadened the view of outcomes available to the team. Results with the NHP and SIP have shown dimensions where rheumatoid patients are scoring highly (indicating poorer health) yet are not covered in the functionally based measure of the HAQ. Comparisons of RA patients to other case types have also been of interest and have shown the high levels of disability among this patient group.

Finally the study has also demonstrated that in the right circumstances, the simplest piece of information from a patients follow-up, ie survival, can prompt a productive investigation into the quality and effectiveness of care provided.

C. Urology - Prostatectomies

Work with the urologists was once again prompted by their interest in the area of outcome measurement. From the start it was made clear that resources were very limited and a full scale investigation of outcomes would have to be resourced from elsewhere.

However two pieces of work were completed. One was an exploration of the type of dataset required to monitor outcomes of prostatectomies, and the second was a simple study of ways to monitor symptomatic improvement in patients.

Outcomes indicators for prostatectomy

Prostatectomy was chosen as the most suitable group for outcome monitoring in urology. It formed the most common patient group and represented a fairly standard treatment procedure. A number of other studies of outcome of prostatectomy have been undertaken

yet uncertainty still exists over the benefit that may be expected (Fowler et al 1988, Roos et al 1989). The available evidence suggests that the operation is successful in terms of symptom relief in 70-80% of cases (Neal et al 1989), though changes in the quality of life are less certain (Fowler et al 1988). Recently, outcomes of prostatectomy have been the subject of international collaborative studies (WHO 1988) and in the UK a large inter-regional study has just been completed. The discussion of outcome measures sought to exploit this earlier work.

In order to cope with large numbers of patients it was agreed that the data set should be kept as simple as possible. The outcome indicators would be based on:-

Symptomatic improvement - using a simple scale developed by Frimodt-Moller (Frimodt-Moller et al 1984) which assesses irritative and obstructive symptoms. A patient completed version of this scale was devised and tested.

Complications of procedure - a checklist of basic post-operative problems was developed (as in other studies) to be completed by the clinical staff at discharge.

Improvements in health status- The Nottingham Health Profile was selected as an appropriate instrument based on evidence of the other outcome studies on elective surgical cases in the Freeman Hospital, and from the N.W.Thames/Oxford region studies.

Major events after discharge - it was agreed to include questions covering adverse events after discharge including repeat operations and re-admissions in the review of patients at follow-up.

In addition key features of the patient on admission and of the process of care were considered important and added to the list of necessary data items. These included:
Age
Co-morbid conditions - by organ system
Indications for surgery - Acute/chronic retention, symptomatic, malignant prostate or other.

Procedure type - Trans-urethral, open, or bladder neck incision.

Though the routine collection of data for this outcome study was not undertaken by the research team the work was used as part of a successful bid for regional funds to audit prostatectomy outcomes in a number of sites. This work is now being undertaken by the Urologists. Plans to develop a multi-region audit of prostatectomy are also being developed in collaboration with the Royal College of Surgeons.

Assessing symptom status

One of the potential problems in the assessment of outcomes in urology was the requirement of clinical staff to score patient symptoms with a scale that was not part of the routine practice. Moreover repeating the score at follow-up would require an interview by the clinician. It was agreed that it would much simpler, and possibly more reliable or accurate, if patients could complete the scale themselves.

A simple patient questionnaire was therefore developed based on the Frimodt-Moller scoring system. The questions were worded so as to be as unambiguous as possible while retaining the ability to translate back to the original scoring structure. The new questionnaire was tested on fifty patients by being given before the operation and six months after and then comparing the scores derived from patients with the clinicians assessments. Full details are presented in an article submitted for publication (Bardsley et al 1991b).

The results suggest that this method produced scores that correlated sufficiently well with the clinician based assessment. The scales show significant improvement following the operation in the way that was predicted. The patient completed score also showed a better correlation with the the patients perception of the success of the operation, than the clinical scoring. It is suggested that the use of such a patient completed scale provided a much more practical way for developing symptomatic assessments of patient outcomes.

D. Conclusions - Other Specialties

The studies that have been undertaken in orthopaedics, rheumatology and urology have only been reported briefly in this report. Though the data that had emerged from these studies is interesting perhaps the main messages from this work concern the possibilities for outcomes monitoring in different settings and the interest of the clinicians involved.

These studies have been important in testing both the theoretical and practical work done in the main specialties. The resources devoted to data collection in these specialties has been less than for the main specialties yet useful results have been observed. The main input has been in the process of developing sets of outcome indicators and in the analysis of results. In one sense the research team was acting as a resource available to the hospital to help various clinical groups develop outcome measurement within their own specialties. It was therefore possible to draw on expertise in both the theory and practice of outcome measurement that developed during the course of the project. This type of role for an 'outcomes office' which can advise and assist specialties in their own development of outcomes would seem to be a useful organisational model for promoting such measures.

Chapter 9 An overview of the results

The previous chapters show the detailed results in the individual specialties and cover the choice of outcome indicators, the results obtained and brief descriptions of the ways these results were interpreted and used by clinicians.

In contrast this chapter brings together the results from a general perspective and with an emphasis on the lessons that might apply to other specialties and other hospitals wishing to undertake similar work. The results are presented in three parts. The first deals with general observations on the processes of **Data Collection** and its associated costs. The second section gives an overview of the similarities and differences in the **Outcome Indicators** selected and used across the different conditions. Finally the chapter summarises some general findings with regard to **Using Outcome Information.**

A. Data Collection Methods.

Work with seven different specialties during the course of this project has inevitably involved a wide variety of data collection methods. Before considering some of the detailed issues of which methods worked in which circumstances it is important to note some key findings. The first key point is that in all cases **some form of outcome assessment proved possible.** The resources devoted to various projects varied, as did the breadth of outcome measures developed but in all cases data were collected which enabled longitudinal comparisons in patients' health to be made in much more sophisticated ways than had been previously possible.

A second key point is that **outcome assessments can include patients' assessments of their own health.** In fact data collection for these instruments proved less of a problem than for some of the traditional clinically based measures. There are very few examples of situations where such measures have been used in routine monitoring of hospital caseload. The results of this study suggest enormous potential for the future.

A third point is that the **scale of data collection depends on the purpose of outcome measurement and the resources available.** There is no simple answer to the questions how should we measure outcomes and what data do we need to collect? The answers to these questions depend on why you want to measure outcomes and how much you are prepared to pay for the information. The type of data required in a routine monitoring system covering a whole hospital will be different from those required to judge the efficacy of two alternative treatments.

1. General Models of data collection

Though there were a variety of data collection methods used they did fall into two general models (outlined in Figs 9.1 and 9.2) one for specific hospital based intervention (typically acute care), the other for monitoring of chronic disease over longer time scales. These patterns are related.

Hospital Treatment Model (Fig 9.1)

The health benefits of a particular intervention can be expected to emerge within a fairly well defined time period. Thus data collection required a pre-treatment baseline assessment of the patient - preferably involving the patient themselves, with this basic data being supplemented by detail at discharge and most importantly at follow-up after an agreed time period, say three months and twelve months. Success was measured by the changes from baseline to follow-up. The data were collected from a combination of sources including patient completed questionnaires, medical notes or pro-forma from clinicians. Ideally data can be extracted from existing computer systems such as the PAS or case mix computer.

Patients who receive the same treatment twice - for example a repeat PTCA cause logistical problems for data storage. A repeat procedure will require the first episode to be stopped and a new one begun. This effectively means that the analysis of results compares interventions performed rather than patients receiving the treatment. For those few patients who do receive the intervention twice - the

fact is recorded as a presenting characteristic for the second episode.

Long term chronic disease (Fig 9.2)

With a chronic disease there may be no clear endpoints, the outcomes are therefore typically concerned with maintaining health and preventing deterioration and are expressed as marginal changes from one point to the next. The clearest example is the annual review process in diabetes where all clinic patients should be seen at the Freeman hospital at least once a year. The details of the patient's health are captured at each visit and outcomes expressed as longitudinal change between them. Determining the link between process and outcome can be especially problematic in the treatment of chronic diseases when the disease develops over long timescales. It may be that measurement at each point of contact (eg out-patient appointment) is not necessary and that comparison over defined time periods are necessary. It may be that a subset of an individual patient's visits may have to be selected as the appropriate observation points for outcome monitoring. The exploitation of local clinical micro-computer data bases is important in this respect. They can fulfil the role of maintaining the register of patients, recording some of the basic information (which may be integral to most consultations) and then enable longitudinal comparisons across groups of patients. In both diabetes and rheumatology there was some earlier experience in developing just such data bases.

In many cases patients with chronic disease will experience a stay in hospital. These should be regarded as events within the longer term course of treatment. It may be that the specific goals of the in-patient episode are studied separately - as for example in diabetes in-patient studies. Thus the hospital treatment model is more appropriate to examine such changes in the short term.

Fig 9.1 General Hospital Treatment (In-patient) data collection model

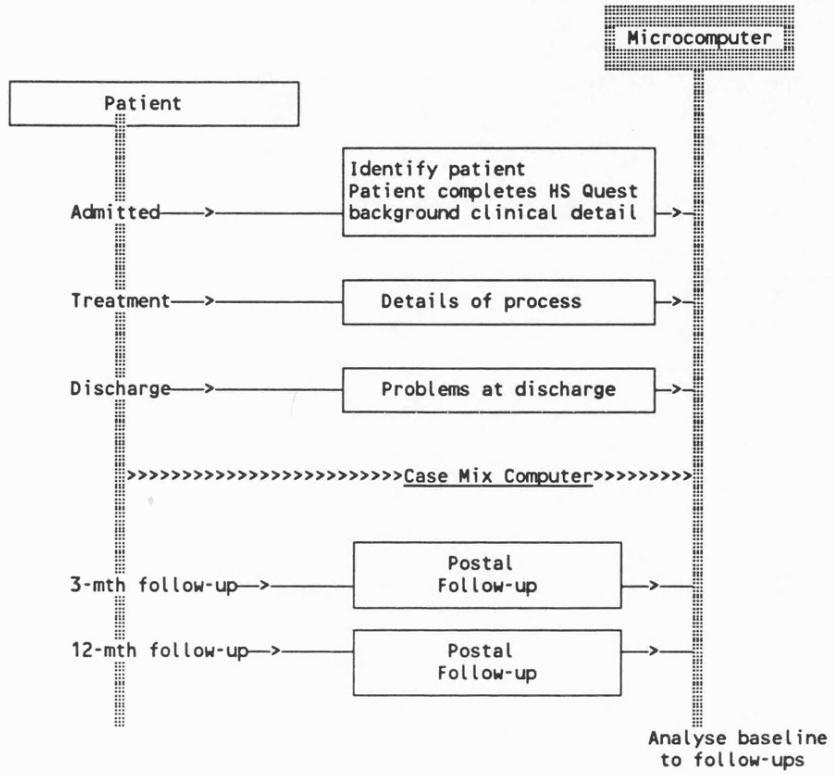
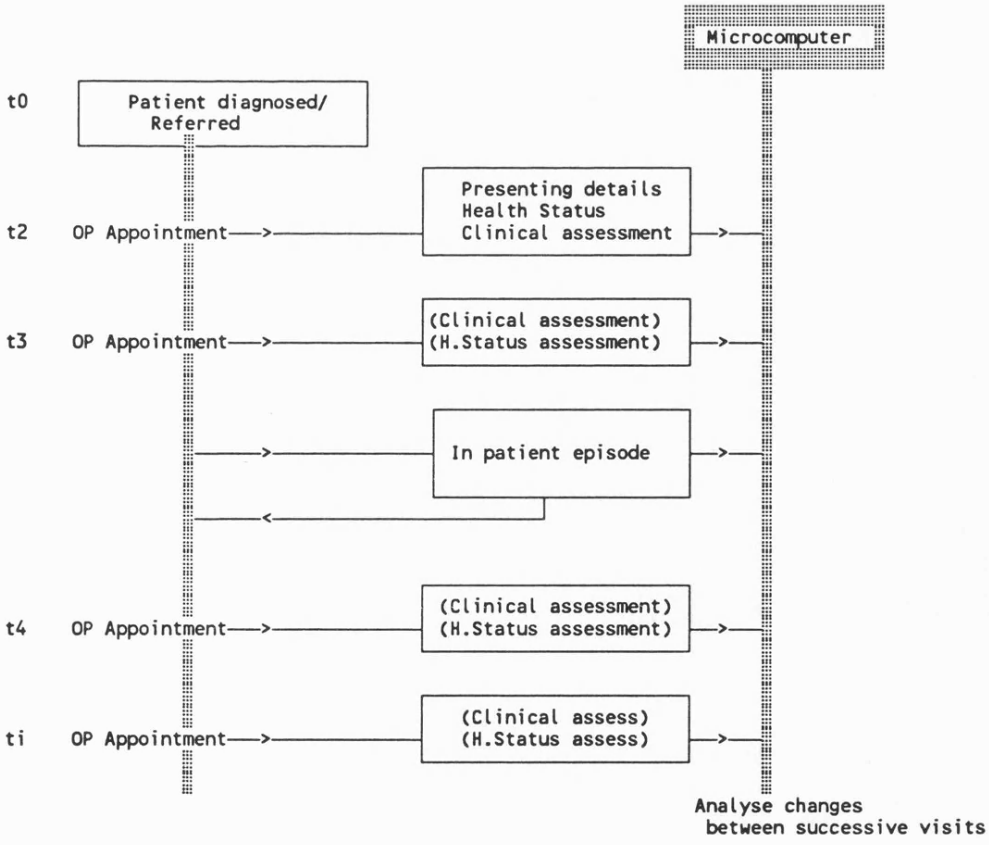


Fig 9.2 Long term /chronic disease data collection model



b. Specific Methods

The various data collection methods used in each specialty were described in detail in previous chapters. Table 9.1 summarises some of the tasks required for data collection. The methods of data collection were under constant review throughout the course of the project and it is difficult to generalise about which methods will be most reliable and least labour intensive in any one area. The choice of method clearly depends on the working patterns within one area, the personnel involved and the cost/resource commitment that can be made.

Throughout, the study has tried wherever possible to use hospital staff to perform the data collection - provided this did not disrupt their normal work patterns. As such it was reasonably successful in achieving this but it has become clear that outcomes measurement of this form will not come about without some resourcing specifically devoted to development of that aim. In particular the development of data collection systems for outcome monitoring required a process of co-ordination of the various data elements as well as the collection of data which was not part of previous assessments - for example health status measures from patients. The scale of costs of data collection are discussed later.

In addition to this general observation some more specific points about the processes of data collection that emerged from this study are:-

* Postal follow-ups of patients showed high response rates and were easy to administer even with fairly long questionnaires.

* Patient questionnaires presented fewer problems (in terms of their timely completion) than some of the forms completed by clinical staff. The major gaps in information coverage tended to be in the clinical information that should have been supplied by the doctors. Outcomes assessment based solely on information drawn from patients would be far easier in terms of data collection.

* Retrospective collection of data from case notes is practical if staff are available and providing the relevant information was actually in the notes. Straightforward items such as brief details of medical history and concurrent problems were recorded in some detail. More sophisticated issues - such as the indications for cholecystectomy or the severity of a patient's angina may not be present. Substantial amounts of time were spent searching for notes to fill in gaps in the information. If notes were used then the more recent they were (preferably while still on the ward/in the clinic) the less effort involved. There were obvious economies of scale in this area, it being far easier to abstract the relevant data from one pile of notes than chase notes around the hospital.

* The process of collecting data for outcomes monitoring overlapped with other areas of information collection. Problems in the current information systems tended to produce problems for the outcome review.

* Time sensitive information - typically capturing a patient's views before a procedure created problems and requires a relatively high labour input. The difficulty was that if patients were not to be missed, admissions to wards had to be checked fairly regularly. Attempts at contacting elective admissions before coming to the hospital were not successful -though for some hospitals and some case types they may well be. Our solution was to have a local contact in the department who could monitor all the admissions to the relevant wards and who would be on hand to see patients if necessary.

* The existing computer systems already contained some important aspects of the information required. In diabetes a substantial proportion of the information was already being collected on a micro-computer. In orthopaedics the proposed local clinical computer system aimed to collect much of the basic outcome dataset. The Patient Administration System (PAS) was essential for tracking patients as readmissions or through outpatient appointments. Finally basic information on diagnosis and demography was available from the resource management systems

(though collected independently), although further work would be needed to make this routine. It was shown that it was possible to link this data to the outcomes database. There are considerable potential benefits from this integration of local clinical data bases with the main hospital systems. Among the advantages are an improvement in the quality of data to the main systems; the reduction of the unnecessary re-inputting of data; and a consistency in data elements across systems. More sophisticated systems should also be capable of checking for re-admissions (and deaths) of patients treated in a particular specialty.

* In a number of cases, the outcomes study required clearer definitions on data items that were already being collected. For example in orthopaedics a pro-forma was used to formalise and document the assessments made by clinicians.

* The identification of whether patients had died between contacts caused problem in all specialties - as has been noted by others (Walters 1990). In some cases forms were sent out to patients who had died and this caused unnecessary distress to friends or relatives. Therefore a variety of methods were used to check on patients survival at follow-up:

  - Hospital PAS systems were checked. However it was realised that these were not always up to date. In fact project staff have been helping hospital information staff by notifying them of any deaths they discover.
  - All local health authorities were requested to circulate their notifications of deaths to the project office. The intention was to add this information to a special database in our office. The volume of cases, and the backlog however has meant that we were only able to deal with Newcastle residents. There would seem to considerable scope for improving the co-ordination, storage and dissemination of this information, probably at a regional level.
  - The name and telephone number of a patient's General Practitioner is now recorded. Where there is a likelihood that the patient may have died these were contacted before sending out questionnaires.
  - In cases of uncertainty the relevant clinical staff were asked if they had

any knowledge of particular patients.

These types of steps will be required in many outcome studies.

3. Costs of data collection

It is difficult to answer the question about what the costs of data collected for outcome studies will be since they are dependant on a number of factors.

Table 9.2 illustrates the typical expected time spent on data collection in each of the specialties. It is clear from this table that the costs vary quite considerably across systems. The key factors governing the cost appear to be:

- The extent to which data collection is additional to current practice or whether it exploits existing staff or procedures. For example the use of a standard pro-forma in the diabetes annual review clinic and orthopaedic knee assessments means that information collection is part of the standard clinical assessment that would go on anyway.

- The process of data collection can demonstrate considerable economies of scale. The more patients seen the less time per patient.

- The physical location and timing of data collection can affect the costs. Comparing cholecystectomy with angioplasty, the basic task of checking for patients in the cholecystectomy study was more complicated as operations were spread over a number of days and possibly wards. All the angioplasties were done on the same day and ward block.

- The most expensive items were those connected with identifying patients and chasing and abstracting information from notes. The rheumatology study was a simple comparison of health status measures in which the patients provide most of the information.

Table 9.1 Data Collection Tasks

Task		Possible Staff
Identify patients	Check admission to relevant ward or ward block. or Check out-patient appointment list.	R Nurse Ward staff Clinic staff
Give forms to patient and collect	Either hand forms+letter to patient directly or leave in out-patient department	Research Nurse Ward staff Clinic staff
Complete clinical pro-forma	Collect clinical details direct from patient	Clinician
Extract data from notes	Transcribe details from notes/lab results to pro-forma	Research Nurse
Chase forms (and completers)	Ensure all relevant forms have been completed and chase those that haven't.	Research Nurse
Chase notes	Ensure completeness of information from notes	Research Nurse
Data input	Entry to software	Secretary
Check details for postal follow up	Check hospital numbers addresses, deaths etc.	Research Nurse
Send out follow up	Print labels envelopes etc.	Research Nurse
Check Returns	Ensure all postal forms returned - re-send to non-responders	Research Nurse
Check database integrity	Ensure various entries on data base consistent and complete. (occasional validations)	Research Nurse

- Data inputting was not been a great problem. Efforts could have been made to improve the speed of inputting software but this did not prove to be necessary despite some large data-sets.

- The identification of patients in the hospital and making contact with them was expensive if the data was to be collected at a certain time during the stay (eg shortly before an operation) and for all patients.

Estimates for the inpatient projects with the largest datasets such as the angioplasty and cholecystectomy models (collecting pre-operative baseline data, process details and two follow-ups) are that with reasonable co-operation from other hospital staff data collection time per patient averaged out at around one hour. This hour would include contacting the patient and collecting pre-procedure data (25% of time), abstracting additional data from notes and co-ordinating forms (25%), inputting the data and doing two follow-ups at 3 and 12 months.

A team of one whole time equivalent researcher (say nurse) plus a half time input clerk could expect to cope with around 50 patients per week (2500 patients per year). The additional costs of just the marginal data collection for this most intensive of the approaches would therefore be up to £20,000 per year for roughly 2500 patients (around 10% of the hospital throughput). In addition some time for analysis of the data would be required as well as the initial set-up costs for software and computers. Though extrapolating from this limited base is a risky exercise and assuming economies of scale balance the overhead costs, it would appear that this most intensive form of data collection on all in-patients would cost up to £200,000 per year which for this hospital is less than 0.5% of annual recurrent expenditure.

Table 9.2 Typical resources required for data collection.

Task	Angio	Chole	Ortho	Geriats	Diab new	Diab AR	Rheumatol IP	Typical Hrs/pat
Max monthly throughput (No. of patients)	20	20	40	80	20	20	20	
Approximate Hours/month								
Identify patients Give forms to patient Complete clinical pro-forma	6 [1]	4 [1]	2 [2] {10}	[6] Nurses [24]Doctors	4 [6] {5}	{5}	4	.25
Extract data from notes Chase forms (and completers) Chase notes	10	8	8	-	8	-	-	.25
Data input	4	4	12	8	6	2	2	.25
Check details for FU Send out FU Check returns Check database integrity	6	6	12	16	8	-	6	.25
Total additional hrs/month	26	26	34	24	26	2	12	-
Hours/patient	1.3	1.3	0.85	0.30	1.3	0.1	0.6	1.0

[ ] - Non-project staff tasks additional to normal
{ } - Non-project staff - pre-existing tasks - not included in totals

Table 9.3 What information at what cost for what purpose? Suggested hierarchy of information sources and rough estimates of the numbers of patients that could be included from a fixed resource of one and a half wte.

Information	No. patients (cumulative information) [Extra tasks]	Comments
1. Postal follow-up to patient for readmissions, death	300/week	Crudest assessment, less than 10% of cases will show problems, of these maybe 1% will be 'interesting'. Tend to report as individual cases (exceptions)
2. Follow-up health status Q.	200/week [Extra data input from 1.]	With only follow-up need comparative data to see if scores good/bad etc. Will give finer grading of post-op health than 1. though interpretation may be the problem.
3. Basic presenting characteristics Age, sex, indications co-morbid conditions	150/week [Download from computers?] [Note chasing and abstraction?]	Important when aggregating results across patients. The indications for surgery are critical, other variables less so, but leaving them out seems unnecessarily risky.
4. Presenting health status (base line for follow up)	100/week [Trapping patients on admit]	Enables questions like - how many patients are 'better' and how many 'worse'? In 2. assumed all pre-op scores are the same. Interpretation more complex. Need to trap patients before the operation- may not get emergency admissions. Strongly correlates to clinically more complex cases.
5. Specific symptoms/problems at follow-up	100/week [Extra data input]	Opportunity to check on specific questions at follow-up eg abdo pain, vomiting etc. Best as simple yes/no with a clear idea about what the answer should be. Save costs by not having a baseline.
6. Post-op complications Use checklist. eg wound infections etc. yes/no	100 /week [Additional abstraction on discharge]	Fairly easy to collect retrospectively (assume major problems in notes). We found not too illuminating , a number of transient problems are seen which have gone by follow-up.
7. Process data Pre & Post-op LOS Theatre time, specific tests etc.	50 /week [Additional detail abstracted from notes]	Not outcomes, but good for discussions. Address specific questions of variation in resource use vs outcomes
8.Peri-operative complications	50 /week [Clinician completed form]	Opportunity for surgeons to report things going wrong in theatres (predictive of immediate post-op problems). May also want surgeons assessment of risk.
9. Specific symptoms/problems on admission	50 /week [Minor addition to patient questionnaire]	Baseline values for 5. may not be needed if the items chosen are fairly clear cut,. In our studies though we have found it dangerous to assume patients will have certain problems on admission

These costs could be dramatically reduced (or the volume of cases increased) by reducing the size of the data set, using more pro-forma completed by the clinicians as they see the patient, or more reliance on patient based data. A study using just follow-up data would be considerably cheaper. Table 9.3 gives rough estimates for the numbers of patients that could be included using different parts of the data collection process and gives some brief comments on how the analysis and type of question addressed will vary with the scale of data collection. At one extreme are simple follow-up studies using little or no baseline data. These would be feasible for large volumes of patients and enable some comparative analyses but the interpretation would be difficult. At the other extreme is the full data set used in this type of study.

B. The Outcome Measures

The study worked in parallel in a number of specialties and sought to develop outcome indicators that would be suitable for particular conditions and which might provide generalisable lessons about outcomes monitoring across specialties. There need not necessarily have been any common outcome measures across specialties yet certain common themes emerged.

Tables 9.4(a) and 9.4(b) put most of the measures into a common framework. There are many examples of common measures across specialties. The indicators have been classified into the following categories:-

- Deaths
- General Health Status
- Major adverse event (readmissions)
- Symptoms/problems relief
- Treatment complications
- Treatment Success

1. Commonality/specificity of data items

The study did not aim to develop measures which would be applicable across all disease types thus it is not surprising that the data-sets collected in each specialty show some differences. In a comprehensive analysis of outcomes such differentiation is inevitable and probably desirable (Patrick & Deyo 1989). There will always be the need for some measures to describe specific outcomes associated with specific patient types and possibly even processes of care.

Nevertheless some common data elements did emerge, for example:-

Basic administrative, demographic information

Age,sex,date of admission, hospital identifier, address etc.

Recording of co-morbid conditions for acute interventions

Recording peri-operative and post-operative complications for surgical procedures

Deaths - during or after treatment but before follow-up

Major adverse events after discharge before follow-up - re-admissions

General Health Status measures

2. Comparisons of health status scores across conditions

A number of different instruments were used to assess health status during the course of this study. Different tools were used in the belief that they would be more specific and more sensitive to particular conditions. However by the end of the study it became clear that there was a considerable interest in using the same instrument across conditions and data became available for the use of one instrument (NHP) in almost all case types studied (the exception being care of the elderly where administration of the questionnaire proved a problem). There are distinct advantages in having one instrument for a range of conditions. In practical terms standard methods of administration, scoring and interpretation make the collection and analysis of data easier. There is also the ability to compare between case types and treatments in a common context - though the analysis is potentially difficult.

Table 9.4(a) Summary of Key Outcome Indicators For Three procedures.

	Cholecystectomy	PTCA	Knee Replacements
Successful care or treatment	Surgeon pro-forma	Cardiologist pro-forma Specifically for vessels dilated	Surgeon pro-forma
Treatment complication	Pro-forma on discharge	Pro-forma on discharge	Proforma on discharge (in hospital)
	[Commonly include.. DVT, wound infection, urine retention, MI, etc.....]		
Symptom/problem	Patient completed questions before and at follow up on pain,vomiting, bowel function, flatulence, abdom. distension. Postal follow-up.	Patient completed questions on pain, walking distance. Angina score from notes. Postal follow up.	Knee scoring system follow up (3 & 12 mnths) Out-patient follow-up.
General Well Being.	Improvement in Nottingham Health Profile	Improvement in Nottingham Health Profile	Improvement in Nottingham Heath Profile
Readmissions	Patient questionnaire and hospital computer system.	Patient questionnaire,hospital and clinical computers Specifically repeat PTCA's.	
Death	DHA notification, hospital computer, postal.	DHA notification, hospital and clinical computer, post.	DHA notification, hospital computer.
Other indicators:	No visits to A&E	Reduce medication	No loosening prosthesis

Table 9.4(b) Summary of Key Outcome Indicators For Diabetes and Care of the Elderly

	Diabetes	Geriatrics
Successful care treatment	Metabolic control via HbA1, Body Mass Index, Cholesterol etc. Short-term complications, Angina, Blood Pressure.	Alleviation of acute medical condition
Treatment complications	No admissions for hypo/hyper-glycaemia	
Symptom/problems	Minimal long term complications from consultation pro-forma/ existing database eg retinopathy, ulcers, nephropathy.	Improvement in functional Barthel score against individual patient target Improved mental test score
General Well Being.	Improvements SIP completed by patient in clinic.	
No admissions	Patient asked in clinic and hospital computer system.	
Death		DHA notification, hospital and clinical computer
Other indicators	Patient education/compliance tested before/after education program.	Minimise burden on informal carers (strain scale) Maintain independence at home.
Important discriminatory	Age, sex, diabetes type	Medical condition, admission

Figs 9.3(a) to 9.3(f) present NHP results across the variety of case types we have examined. These graphs show the average values on the six dimensions of the NHP. The higher values on the scale indicate poorer states of health. Such comparisons within the hospital - or to values from the literature were received with great interest by the various clinicians concerned. Similar comparisons have been made using a different instrument by the medical outcomes study (Tarlov et al 1989).

By and large these results confirm the expected picture and help to re-assure us of those dimensions of health measured. In cholecystectomy the dimensions covering energy, 'Pain' and 'Sleep' were highest (ie indicating poorer states of health) before the operation and show the greatest change. This can be compared to knee replacements where 'Pain', 'Sleep' and 'Mobility' showed the greatest changes to 3 months. Both the initial scores and the residual scores at three months were greater in the knee replacement patients as one might expect for patients with chronic problems.

In Figure 9.3(d) the presenting NHP profiles are compared (using mean values) for cholecystectomy, knee replacements and angioplasties. Patients about to receive a knee replacement showed markedly higher scores on 'Pain', 'Sleep' and 'Mobility' with PTCA patients being slightly worse on 'Energy'. 'Social isolation' consistently scored as the lowest of all dimensions.

Figs 9.3(e) and 9.3(f) contrast the NHP profiles for types with lowest and highest average NHP scores from the cases studied. In one the high scores of rheumatoid patients were clear - with inpatients only slightly higher than out-patients. In contrast the mean scores in diabetes (in this case the distributions are very skewed, many score 0) are markedly lower. Nevertheless the scores of Type II patients (who are typically older with more diabetes complications) showed up higher than those of the sample of patients with type I diabetes. A closer examination of these NHP distributions in a population of patients with diabetes showed that the scores are very skewed with a high proportion of cases scoring zero, it has been found that the scale is able to differentiate between patients with and without specific clinical problems (Bardsley et al 1991a).

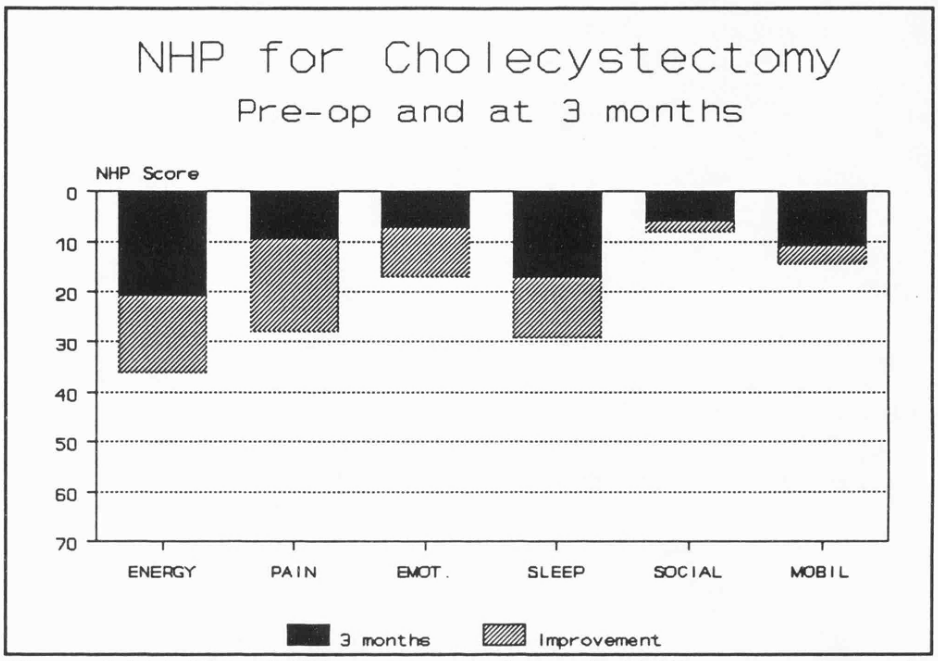


Figure 9.3(a) Mean NHP scores and change in NHP score to 3 months after cholecystectomy.

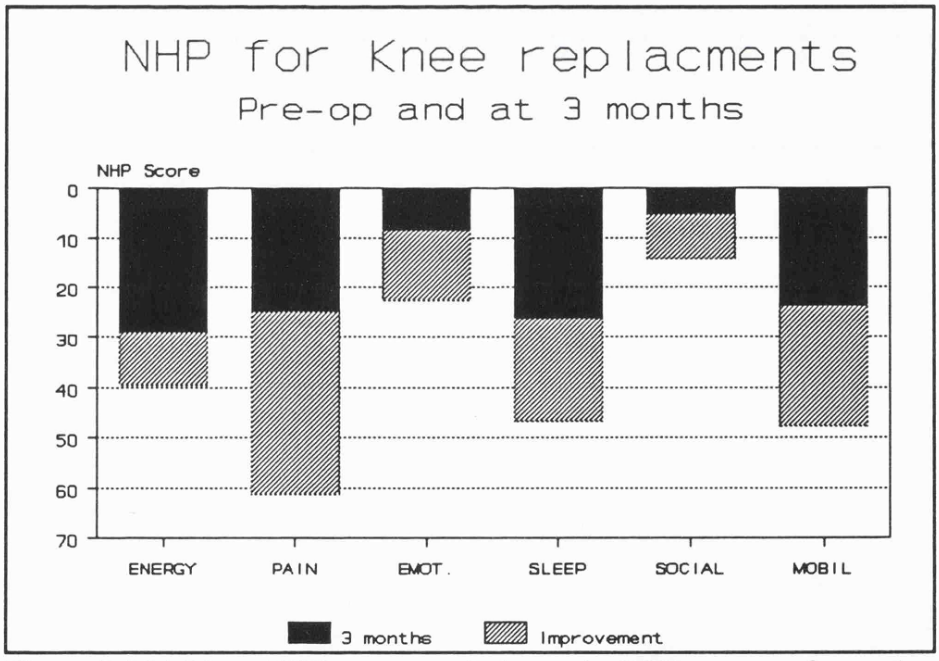


Figure 9.3(b) Mean NHP scores and change in NHP score to 3 month a fter knee replacement

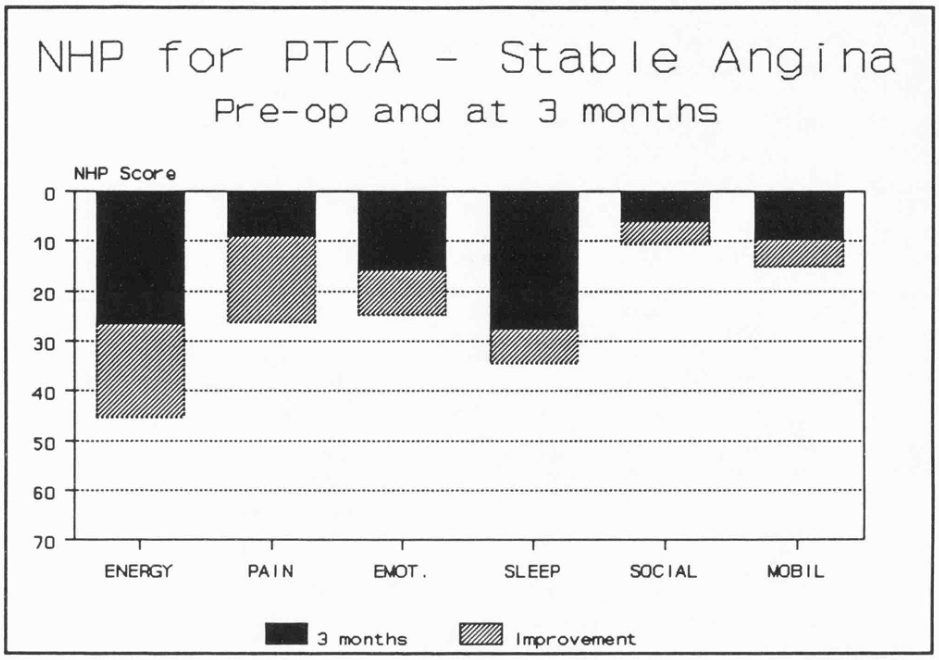


Fig 9.3(c) Mean NHP score and change in NHP score to 3 months after PTCA (stable angina patients)

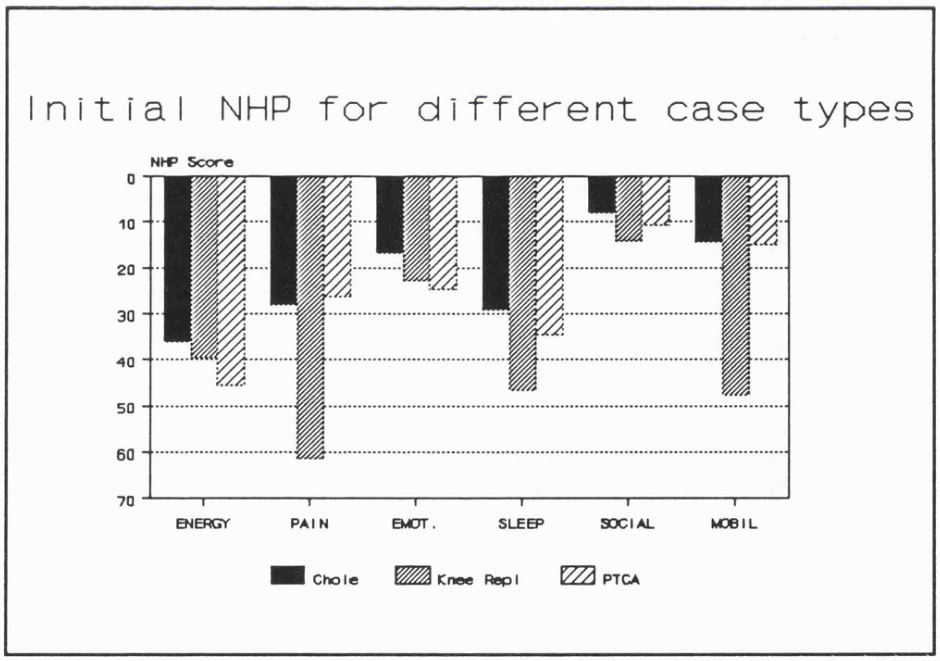


Figure 9.3(d) Mean NHP scores on admission to hospital for three procedures, Cholecystectomy, Knee Replacment and PTCA (Stable angina cases)

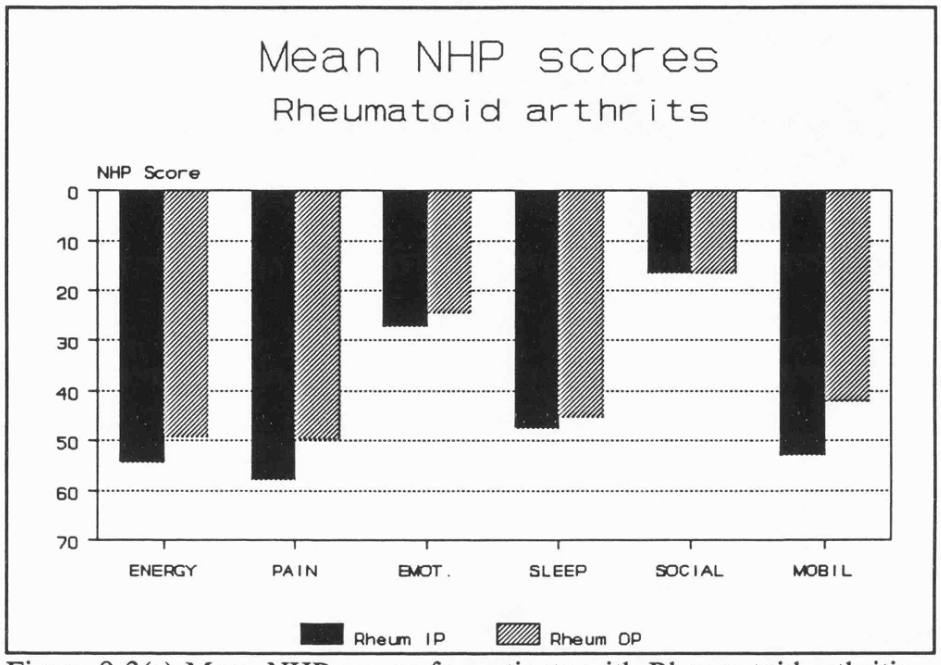


Figure 9.3(e) Mean NHP scores for patients with Rheumatoid arthritis (Inpatient and Outpatients)

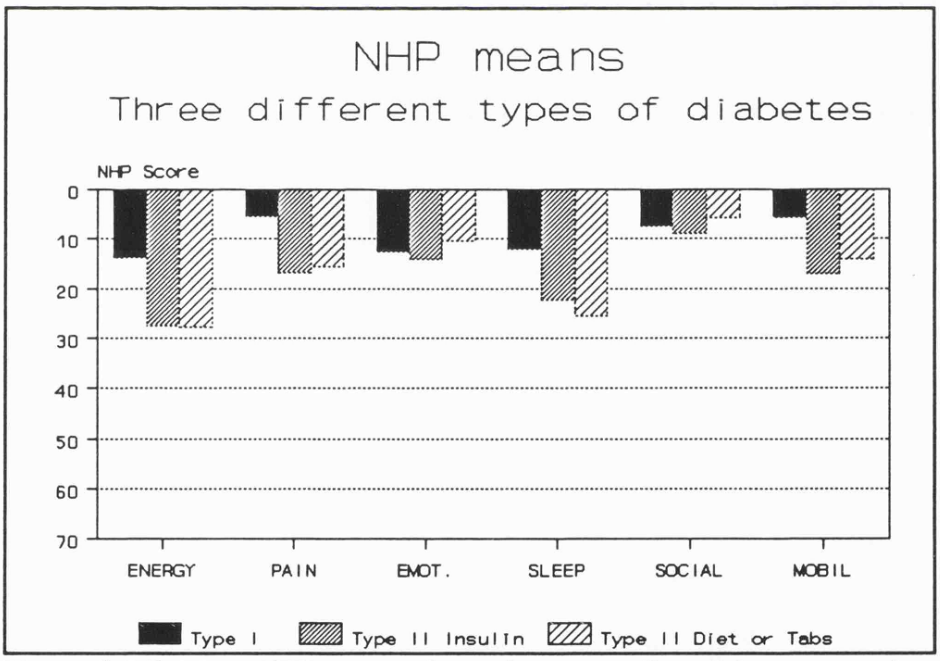


Figure 9.3(f) Mean NHP scores for patients attending diabetes outpatient clinic.

Thus for example the patients receiving knee replacements probably scored highest across all dimensions but in particular on 'Pain' and 'Mobility'. In fact they scored just slightly higher than an out-patient sample of cases with rheumatoid arthritis (suggest more severe patients are receiving the knee replacements). The knee replacements continued to score highly after the procedure (though there is a significant improvement) which is as one would expect from a population with severe chronic disease.

3. Timing of measurements

Ideally, relative outcome measurement should be a cumulative function of health over time. However continuous measurement over time is not practical and the assessment of outcome has to be based on changes observed between two points in time. The choice of appropriate time periods for assessing outcomes can be as important as the choice of outcome measures for valid assessments. Results in this study have suggested that the best times for measurement vary by condition. Assessing the best time for measurement requires balancing when the effects (if any) of the intervention are manifested (beneficial and harmful) against changes in health that may occur that are not related to the intervention. The longer one waits to assess the effects of the intervention, the more uncertainty is created by unrelated events. To answer the question fully requires an assessment of the relative sensitivity and specificity of different time periods such as by the use of receiver-operator characteristic (ROC) curves (Deyo & Centor 1986). In this study the use of 2-3 different time periods does give us some ideas about what may be the best time periods for different conditions.

In cholecystectomy the greatest benefit seemed to be observed by three months after discharge. Though it was not tested, surgeons felt that observations earlier than this would see patients still suffering from minor self-limiting problems following the operation. By twelve months some patients will show the effects of health problems which are largely unrelated to the operation thus serving to make the measures less specific.

In angioplasty, where the patient has long-term chronic disease, though most of the

benefits were shown at three months, during the following nine months a number of adverse events, connected with the condition, rose and substantially altered the picture of success. It is possible that such events will continue beyond a year.

In geriatrics, most of the benefit in terms of improvement in functional status occurred during the inpatient stay. Significant improvements were observed in patients during this time and no significant change in function (for better or worse) was seen after discharge. Though this is probably not an appropriate interval for all the outcome indicators - for functional status it may be the best way to monitor performance.

Finally in diabetes it was clear that there was a period of beneficial changes in newly diagnosed patients which could be monitored either at three months or a year. However after this early phase, established cases showed only slow overall changes over subsequent years. In this case the minimum of one observation per year after the initial diagnostic period seems appropriate to ensure that health status was being maintained.

4. Multi-dimensionality, clinical and patient perspectives

One common thread across work in all specialties was the multi-dimensional view of outcomes. Thus success was not measured against a single score but against a variety of measures. In this type of study such a multi-dimensional view of outcome is recommended. Different outcome measures can have different value and properties when it comes to interpreting results and relating outcome to the processes of care (Lohr 1989).

During the course of this study, outcome indicators could be categorised with respect to three key properties. Any one indicator can be mapped somewhere onto all three axes:-

1. From clinical perspective to the patient's perspective
2. Process/proxy measures or direct descriptions of health
3. Vaguely linked to processes of care to strongly linked to care

The ideal indicator was one that represented the patients' perspective (that is utility expressed by the patient rather than the clinical perception of that utility), was a direct description of the patient's health rather than a proxy measure, and could be closely linked to the process of care. No one measure satisfied all these criteria completely and different indicators tended to have different combinations of these properties.

Patient vs doctors' perspective.

As discussed earlier, the chosen indicators included different perspectives on outcome - most importantly a mixture of the clinical view and the patient's view. Though the incorporation of some type of assessment of how the patients feel was rather novel the principle was accepted in all specialties and in time the instruments used were accepted as valid.

It might be argued that the only perspective that matters is the patient's. However this would deny the value of a variety of clinical measures which either can be assumed to correspond to the patients perception (for example the judgement that amputations or pain is undesirable), or that were accepted as legitimate predictors of longer term health problems. In these circumstances it may not be possible for the patients to assess the relative value of different health benefits, or for the perceived benefits to be measured.

Process/proxy measures vs direct measures of health:

Similarly during the course of the project the distinction between process measures and outcome measures became blurred. At the start considerable efforts were made to concentrate on outcomes which were direct descriptions of patient health. A variety of strictly process measures eg length of stay were offered and rejected. However in many cases measures which were descriptions of process were accepted as proxy outcome measures in the belief that they were linked to real patient problems and so the benefits/disbenefits of treatment. For example re-admission following surgery, or HbA1 levels in diabetes. These measures tended to be easier to define and collect and perhaps most importantly easier to interpret. In some cases the process measures were claimed

to best define the goals of a particular service. They were administered in the belief that the process would lead to better patient health yet the evidence may have been flimsy. Thus for example in diabetes, HbA1 levels do not directly affect the patients every day life -yet they were very often the focus for determining the types of treatment offered.

Vaguely linked to treatment vs strongly linked to treatment.

The problems of determining the causal links between process and outcome inherent in Donabedian's definition were a common thread across specialties. Though a particular pattern of outcome was observed - explaining why it had occurred and how it could be improved was not necessarily easy. The causal links between certain processes and outcomes may be relatively unknown - which makes the identification of how to improve care difficult if not impossible. Thus the general health status measures were accepted as legitimate ways of describing a patient's health - but if a patient failed to improve they gave few clues as to the reasons why or how it could be avoided in future. For example in cholecystectomy poor performance on the NHP could be due to specific problems related to the surgery or other general health problems such as rheumatism or even in one case a road traffic accident. On the other hand the clinical measures tended to be more useful in this respect. For example a wound infection following surgery gives some indication of which areas of clinical practice need to be examined and hopefully improved. These differences between measures are important when the question of how to use the data is considered.

5. Availability of data

An early goal of the project was to exploit as much of the existing information as possible with the important addition of measures from the patient's perspective. The idea was that if the information was important as a clinical description of the patient then it would have been collected in some form somewhere. The ideal is when the data is already collected on a computer database. Local data from the case mix computer and

clinical micro-computers was used and such systems hold considerable potential value for outcomes information work.

In other cases data was extracted from clinical notes where it has to be admitted the level of recording was variable. In some cases the project required data in a standard format which was possibly a development and improvement on existing methods. For example in orthopaedics a standard assessment of knee function has been used for some time - the project merely required the information to be re-organised into a standard format.

C. Using the Outcome Information

This study has sought to examine how outcome information can be used in practice. It should be noted that the timescales required for collection and analysis have been very short - indeed some of the more complex analyses have only just been performed and are now being shared with the clinicians. It is therefore perhaps not surprising that the uses to which the information has been put are limited and the results rather mixed. Some of the important issues drawn from experiences at the Freeman Hospital are discussed below.

1. Clinical support

The results of the questionnaire sent to clinicians (Appendix 3) confirms the view developed by the project team that there has been a surprising commitment by clinical staff to the project and a desire to see the work continued. All responses suggested that the objectives of agreeing and collecting outcome measures have been met and clinicians would like, if possible, to continue the study in some form.

The commitment is manifested in resources devoted to the project either through the use of their own research staff for collecting and co-ordinating data to involvement in meetings to discuss and disseminate results. The conclusion is that in this particular

setting the clinicians are genuinely interested in assessing the outcomes of their care either at the level of individual patients or as groups of patients. Moreover they have proved willing to consider relatively sophisticated measurement instruments (though they may well be unused to them) as appropriate to measuring outcomes and have found them satisfactory. The interest shown by the clinicians in the results are fairly clearly shown in their responses to our questionnaire (Appendix 3) which included:

> *"We are now clear that the data can be collected, examined against standards and that the results can be presented in a useful form. All those who have seen the analysis found it useful"*

> *"Interesting as it showed elements of the patient's responses to treatment we would not normally have access to".*

> *"Enables detailed review after a busy clinic with highlighting of problems"*

At the start of the project the research team were aware of the potential sensitivity of the information and were correspondingly circumspect in drawing any conclusions. Though it is true that most of the results are generally positive, there were no problems when negative results have been presented and discussed. Similarly clinicians were willing to share their data to colleagues, managers and in some cases with clinicians from neighbouring hospitals.

There were some problems in this area - the most notable being a failure to complete the right forms at the right times but these have not proved insurmountable barriers.
On the question of whether clinicians would like to continue to monitor outcomes the responses were generally positive the reservations being about the costs of data collection:-

> *"Yes"* (3 times)

> *"Yes - difficult to know whether it should be seen as a 'project' type approach or 'routine' at this stage. Ultimately must be routine but the best measures and logistics (without extra resources) need to be sorted out"*

> *"We are doing. We are also extending the ideas to the Northern Region, and nationally and to GPs."*

"As far as is practicable but we do not have the resources in staff/record keeping to make a full continuing study practical"

2. Changes in information gathering

In most specialties the project has had some, hopefully beneficial, effects on the way information is gathered. The requirements for outcome measurement were for a greater degree of standardisation than existed before as well as the collection of some additional data which was agreed to be of value to practising clinicians. Examples include the recording of main care givers and addresses on discharge in geriatrics and the system of standardised knee scores in orthopaedics. In diabetes the ability to examine data longitudinally has identified some issues over the calibration of certain tests which would otherwise have probably gone unnoticed (Home et al 1991).

3. Changes in practice.

Despite the limited time available there are some examples of changes in practice that have resulted from this project.

In diabetes, the change in outcome indicators between successive annual reviews revealed results that were worse than expected and changes have been made as a result. The initial presentation of their own department's results came as something of a disappointment and there was a feeling that they were failing to meet their own standards. The study has changed the clinicians' perception of the benefit which they give to patients and they have identified areas of care in particular need of attention.

It is too early to assess the consequences of these changes but they include:

- new protocols of care for clinic staff
- introduction of formal targeting for individual patients
- increased emphasis on ensuring that clinic staff adhere to the basic recording practices

- replacement of a laboratory assay
- changes in the content of the educational programme

Perhaps more importantly there appears to be an acceptance that it is important to include the patients' perceptions in assessing care. For example the question of whether transferring patients to insulin yields the expected benefits in well-being, has been raised. This issue is about to be explored further by specifically studying changes in both blood glucose control and treatment satisfaction in patients as they are transferred to insulin treatment. The department is now committed to continuing a review of its outcomes in the widest sense.

In geriatrics there has been an increased attention given to the problems of informal carers and attempts to develop a support group. The care giver strain scale (discussed in Chapter 4) is being used as a simple screening tool for the nurses to highlight potential problem areas.

In orthopaedics, the results have changed the clinicians' views on the benefits of surgery for different groups. For example, the use of knee replacement is shown to provide almost as much benefit to rheumatoid patients as it does to those with osteoarthritis. The results have also suggested that younger patients appear to fare less well - the reason for this is not known and is being investigated further. Anecdotally there is the suggestion that some patients have higher expectation of outcome - an area that has not been explored but which may need addressing.

In cardiology there have been no significant changes in practice but one issue -the decision when to reduce medication is being considered. In rheumatology the use of audit data has helped to identify concerns within a particular patient group ie the rate of deaths from cervical myelopathy. Steps are underway to develop a more integrated approach to care of this particular problems. The value of all the measures in helping to shape the practice is also generally accepted.

In surgery, on the other hand no major changes have resulted. A number of important questions relating to the care have been raised but there are no clear indications about how to actually improve practice. There would seem to be a number of factors contributing to why this should be the case.

It is clear that different specialties have reacted in very different ways. Why? The remainder of this chapter attempts to outline some of the factors that in our experience are responsible for this variation and which might serve to help or handicap the exploitation of outcome information.

a. Interpretation and standard setting

As the indicators were developed so attempts were made to identify standards that would describe when observed results fell below the expected standards. In practice it was very difficult to identify explicit standards in anything but general terms as the expected outcomes of care were often unknown. Thus for example many of the agreed aims were couched in terms of improving the patients' condition from the baseline or having minimal adverse events. Of course in practice what emerged was that some patients did not improve and that adverse events were observed in a small but significant proportion of cases. In diabetes where explicit values were used - drawn from the literature - it became clear that expecting all patients to achieve these goals was too optimistic. A more realistic standard would be to say that x% of cases will achieve a given goal.

Comparisons to values in the literature were fraught with possible problems of differences in case types, context or process. Though they were used, they could not be considered as hard and fast bench-marks. Having studied outcomes for 1-2 years it now becomes possible to identify more appropriate norms.

The development of realistic standards therefore required a period of preliminary data collection and analysis - effectively one turn of the audit cycle. The novelty of the measures means that other projects in outcomes monitoring will experience similar problems. This has implications for studies where the goal is not continuous monitoring

but a snapshot of the outcomes achieved since a realistic assessment of whether the results are 'good' or 'bad' will be that much more difficult to make. The outcomes information seemed most effective when there was a surprise value - typically when results were far worse than expected. In this cases there was clearly some kind of implicit, or explicit, benchmark being used to assess results.

The availability of comparative data within and across hospitals can ease these problems - especially if sufficiently large variations in outcomes are is observed. By the end of the project it was clear that any comparable data would have been welcomed and could have helped identify areas where practice could be improved.

b. The value of different indicators - causality and information

The previous section discussed how different indicators have different merits with regard to their ability to reflect directly patient's health, the patient's perspective of their well-being and be associated with a process of care. Considering the discussions across all specialties it is clear that in order to prompt questions and suggestions about how improvements to care are to be made, then the issue of causality is important. Whilst the general principle that health status measurement reflects the care provided is accepted, these measures proved among the most difficult to interpret in terms of how clinical practice affects the results and there tended to more interest in indicators which were under closer control. For example short term levels of blood glucose control in patients with diabetes are felt to be easier to control than the long-term complications of diabetes.

Some of this may be due to the novelty of the tools used. Considerable efforts have been made to examine the relationships between the clinical measures and the more general health status measures and confidence in the latter has undoubtedly grown during the course of the study. We also have to admit to some initial scepticism about these tools. But it is now clear, for example, that if you wanted to measure anginal pain, a more reliable indication is obtained from a dimension of the NHP than the standard clinically accepted angina classification.

The least useful part of the outcome indicators tended to be those sections of health status measures dealing with psychological or social problems (though it is fair to say that generally these scores were lower). This may be just a reflection of the case types used in this study.

One of the consequences of using such general health status measures in patients with chronic or concurrent diseases was that the scores may reflect the effects of a disease process that is unrelated to the treatment process. For example a patient with arthritis might be expected to score highly (ie indicating 'poor health') after a cholecystectomy irrespective of the relative success of the surgery as their arthritis will remain a significant health problem. The positive side of this is that it potentially enables one to assess the effect of the intervention in the context of all the health problems that the patient faces. Patients undergoing knee replacements showed a marked improvement in NHP scores - but they also have high residual scores as a result of their arthritis. The negative side is that this may mask the beneficial effects of the treatment and clinicians may feel that the effect of these other disease processes are beyond their control.

On a number of occasions during the project discussions centred on the psychological profile of the patients and the subsequent benefit. One problem discussed in orthopaedics was the patient whose knee was working adequately yet they were unwilling to exercise it and improve their lifestyle. Similar discussions in cardiology concerned patients who were unnecessarily cautious about their condition and let it limit their lives far too much. The problems were seen to be in identifying such cases in advance and the ethics of possibly targeting care to those patients who would benefit most.

c. The search for truths

The limitations on the conclusions that can be drawn from this type of outcomes study were discussed in the first chapter. Throughout the project it has been clear that non-controlled studies which cannot assess 'what would have happened anyway' cannot identify the most effective model of treatment. What they can do is to identify where

goals are not being met or that variable x is associated with worse results, or to permit the comparison of two processes through post-hoc analysis. Such statements can act as the starting point for more definitive research or analysis or as a way of refuting currently held but misguided beliefs (eg that rheumatoids do not benefit from knee replacements; reducing length of stay affects patient outcome).

One of the implicit benefits of monitoring outcomes has been not just the identification of particularly poor outcomes that require improvement - but also the recognition of when a particular area of care is not felt to be a problem. In cholecystectomy the rates of post-operative complications were felt to be in line with those expected. Though some wound infections were observed it was not felt to require remedial action. One would hope that the effect with then be that, for example, future discussions on installing new laminar flow systems in theatres etc. will bear these results in mind. There are many areas in which a better knowledge of outcomes might support the status quo and be inferred as a positive result for this project.

A note of caution was required in many instances where it was felt that because we were looking at outcomes of care, we would necessarily identify the best treatment process. As mentioned earlier, a routine monitoring system may not enable such comparisons though if desired it is possible to formulate specific research questions with more specific methodologies - within the basic framework of outcomes monitoring. This would probably have helped in the analysis of results. In particular there seems scope to employ outcome measures in comparative studies over time to evaluate particular service changes that may be taking place.

d. Role of audit and management

It became increasingly clear that to make best use of the results there was a need for the sort of critical questioning environment that is to be found in good medical audit (Shaw 1989). There is the danger that outcomes information merely becomes 'interesting' without necessarily guiding improvements in the service. Similarly for the outcomes

information to be used as part of the management process, at the level of consultant firms or specialties, a link is needed between the local priorities and management issues being considered.

e. Uncertainty of efficacy or uncertainty of technique.

One general observation on the potential effectiveness of the outcome information was that it appeared more useful where there was recognised uncertainty over the value of a particular intervention or the choice between techniques.

It has been argued that clinicians will only be interested in outcomes measures if they look good. We would not support this view. Though it is undoubtedly easier to defend the NHP if it changes some of the results in the right direction, even negative results have been accepted.

f. Links between access, appropriateness and treatment

A recurrent theme in many of the outcomes review meetings was the importance of the patient's condition before the treatment began. In cholecystectomy, angioplasty and knee replacements, the key issues as perceived by the clinicians was about choosing which patients were to have the procedure rather than the technique used. In order to maximise outcomes it was recognised that the selection of patients was critical. The issues involved for each of these conditions are all slightly different. For cholecystectomy the decisions were about the additional dangers of surgery when judged against medical management of the symptoms in certain high risk patient groups ie elderly and those with complicating conditions. In orthopaedics the choice was who would benefit most from being taken off the waiting list to use a scarce resource. In angioplasty questions were whether the patient would benefit more from a PTCA, than from surgery or medical management.

g. Alternative methods of presentation - individuals vs groups vs exceptions

A number of different methods of presenting the information were possible. Specifically outcomes could be reported as :

- Individual patients.

  Data can be presented by individual patient either as a summary listing or as additional information in the clinical notes. Though this was used in some circumstances it was not done generally for all patients. However there is considerable potential if the approach could become part of the clinical information available say in the notes. There are a number of ways in which those could be made possible, though the resources to do it were beyond the scope of this study. It should however be noted that there are considerable similarities between local clinical databases and the outcome data used here and that in an ideal world a single system would feed both approaches.

- Changes in groups of patients

  Most of the results were presented as changes in groups of patients - subdivided by relevant key variables where necessary. Throughout the study there has been a need to reconcile the clinicians' tendency to think of individual patients, with the results for groups of patients. The groupings used for analysis are critical and considerable efforts have been put into exploring the relationships between the presenting characteristics of the patient - effectively the grouping criteria - and the observed outcomes. This is particularly important in developing indicators that are comparable across time and across sites. In general the relationships between these presenting characteristics and outcomes were not as strong as originally thought. In many cases variables that one might expect to have predicted worse outcomes - for example the level of co-morbidities - did not have a strong effect on the population of patients as a whole.

- Exception reports of patients with 'poor' outcomes.

  In some specialties patient-specific reports were collected on patients who had 'poor' outcomes and these were discussed individually. This exercise was useful in that it helped the understanding of the group results by having individual case histories. However it was not effective as a means of identifying generalisable conclusions either about those patients or about the way care was delivered. Explanations tended to be of the form that - 'that was a difficult or unusual patient anyway and there was little else we could do'.

h. Analysing the data

One of the reasons that this study extended beyond its original brief was that there was a genuine interest in exploiting the resources that the project had available. Apart from the growing experience in agreeing and collecting outcome indicators there were also some additional basic skills such as in computing and analysis and interpretation of data available.

In concluding this chapter, it should be noted that the timescales for developing the outcome indicator, collecting the necessary data and reporting back the results have been rather short (under three years). The necessity to develop a confidence in the behaviour of the outcome indicators requires both satisfactory sample sizes and some basic understanding of the natural variability of the measures. The study has shown a variety of methods of data collection and given some valuable lessons in how others might proceed towards outcome measurement if they wish. The ability to translate observed results, in terms of outcome, into practical steps to improve performance is clearly not simple yet an understanding of the issues is important if outcome information is ultimately to prove of value. Despite the limited timescales this project has provided some useful pointers and generated a genuine interest amongst the clinicians in applying outcome measures. The next chapter discusses some issues concerning the wider application of these results.

Chapter 10 Conclusions

This study has shown that through working closely with clinical staff in a variety of specialties it is possible to develop sets of outcome measures which are accepted as legitimate indicators of performance. Moreover data for the monitoring of these indicators has been collected and the results usefully discussed with clinicians and, to a lesser extent, managers. The project has also shown that in some circumstances the presence of the outcome information has prompted changes in the way care is delivered and the quality of the service assessed. In the longer term there is considerable potential for exploiting outcome information at a local level in helping with a variety of issues in both clinical practice and in the operational and strategic management of health services. Though there are a number of practical and theoretical problems in the effective use of outcome information but they do not constitute insurmountable barriers.

The outcome indicators used at the Freeman Hospital are relatively sophisticated and contain a certain amount of redundancy of information. More streamlined data sets can be envisaged as a result of this developmental work. Specific issues will require subsets of the indicators used in this study. The purpose of outcome measurement and the resources available for measurement will largely determine the indicators chosen.

A. Relevant dimensions of outcome in secondary care

It is possible that the outcome indicators developed at the Freeman Hospital may not be accepted by clinicians in the same specialty elsewhere. However, experience to date suggest that others would produce broadly similar indicators. One example of a wider consensus approach is that being adopted by the Royal College of Physicians working

party on functional assessments in the elderly which is in the process of identifying outcome indicators which will hopefully be acceptable to the profession as a whole. The problems this group faced were similar to the ones experienced in Newcastle, and the conclusions largely the same. They agree on the use of the Barthel test, the abbreviated mental test and some measures of accommodation status. The issue of a general quality of life measure, appropriate for this group of patients, has yet to be resolved.

Similarly the basic indicators in diabetes seem to be largely accepted in a number of different settings. Work in Newcastle has shown how they can be applied to GP care of people with diabetes. In this study the basic indicators of metabolic control and clinical complications are largely the same as those used with the Freeman Hospital outpatient population.

One further example of agreement on approaches to outcome measurement, is the issue of the use of generic health status measures. Here again the consensus seems to be that if possible these measures should be included. The most popular general tool in the UK still seems to be the Nottingham Health Profile - despite the debates over its derivation and weighting. In the US a similar instrument, the SF-36, is gaining widespread support and is being applied in a variety of settings. In fact there exists an organisation to promote the use of such outcome tools and to encourage the sharing of data (Inter-study 1990).

Though this project has purposely not used measures of patient satisfaction, it is being increasingly accepted that some measure of patient satisfaction is appropriate for a comprehensive evaluation of quality (Maxwell 1984; Cleary & McNeil 1988; Kerruish,Wickings & Tarrant 1988). The rationale behind this is that the objectives of care should include the aim of maximising patient satisfaction - independently from the health outcomes or technical quality of care offered. Beyond this lie the largely uncharted territory of issues concerning the relationships between the technical outcomes, patient expectations and patient satisfaction.

Fries (Fries 1983) described outcomes in terms of the five D's -death, disability, distress, drug side effects, and dollars. The classification that has emerged from this project is not surprisingly similar but its focus on the effects of interventions in a hospital give it a more specific flavour. (Dollars are not relevant in the UK context as for the most part patients do not incur direct financial loss as a result of care).

Most of the range of outcome measures used in this study can be classified under the following headings:

- *Deaths and survival*

  Though the observed rate at which patients die varies enormously across conditions, in most settings survival is a critical outcome measure.

- *General Health Status*

  The benefits of care should result in a marginal improvement in health status as perceived by the patient. The improvement may be expressed relative to a baseline value before the intervention or to the expected change that would occur without intervention. In some cases good outcomes may be about maintaining health status in the face of a chronic disease process.

- *Major adverse events after discharge*

  Following treatment there are a number of possibly serious health related events - typically indicators of worsening patient health - which can be used as realistic proxies of outcome. The obvious example is re-admissions (which as this study has confirmed is associated with worse health) Others may include visits to accident and emergency suites, other treatments, for example CABG surgery following angioplasty in stable angina patients.

- *Symptoms/problems relief*

  Some outcomes may be expressed in terms of the relief of specific patient symptoms or problems, for example anginal pain following PTCA, flatulence following cholecystectomy.

- *Treatment complications*

  Treatment may produce adverse consequences over a short time period. The lower the incidence of these problems the better the outcome. For example wound

infections following surgery, drug reactions etc.

- *Treatment/technical Success*

  In some circumstances the successful administration of the treatment itself can be considered an outcome - though it is also a process measure. The best example from this study is dilation of vessels during angioplasty.

These categories overlap slightly but they are a useful framework for considering the variety of outcome measures in the different specialties.

B. Possible roles for outcome measurement

Chapter 1 discussed some of the different applications of outcome measurement within the health service. The relevance of this work and extensions to it can be considered in a variety of contexts. The following sections consider these different approaches including the use of outcomes measurement in terms of the care of individual patients, in audit, in resource management, hospital management and in the contracting process.

1. Outcomes in management of individual patients

Clinical measures

The study has tried to build on a variety of measures that are currently used to assess patients in everyday medical practice. As such it could be said that some aspects of outcome measurement is already part of the way medicine is practised. However it is clear that a formal recognition of these indicators and their measurement may enhance the assessment of patients on an individual basis. One example form this study is the identifications targets for HbA1 in diabetes, the behaviour of measures across the clinic population and value of showing longitudinal changes in the individual. One of the potential improvements to the diabetes project is to reinforce this aspect of the study by presenting the collected information on individual patients perhap as a summary patient

history in a simple form which is presented to the doctor when confronted with a returning patient in a clinic. Such developments require local clinical databases to handle and format the necessary information longitudinally and will require a consistency in the way data is recorded.

For acute interventions, the potential value of the additional information collected in this study is to act as a monitoring screen on patients. In some cases follow-ups beyond the normal outpatient visit can help identify when a particular patient problem requires more attention. This is not to suggest that patients should never be discharged from a hospital consultants' care, but rather that outcome monitoring could potentially act as a way of screening for post-treatment problems, and in some circumstances one could speculate that this could be an acceptable alternative to an outpatient visit to hospital.

Using General Health Status Measures

The use of individual health status instruments as part of medical assessments of patients has been recommended in some quarters (McEwen 1988) though there is limited practical experience.

One study has developed a series of charts designed to assess the health status of the patient whilst present in the physician's office (Nelson et al 1990). One particularly ingenious approach to data collection has been to use a version of the questionnaire which can be completed by the patient whilst waiting, read by an optical mark reader machine and results given to the patient and clinician, and stored for longitudinal analysis (the technology is being tested in Boston by Ware and colleagues - pers. comm).

The appeal of such measures of general health status is that they can widen the clinician's appreciation of what will constitute a real benefit for the patients and avoid the danger that short term clinical success is sought at the expense of long-term patient benefit (the problem of 'the operation was a success but the patient died'). Thus a standardised instrument is supplementing the subtle assessments that clinicians already make (or should make), or acting as a screening tool to highlight potential problems. Some have suggested

it possible that such measures become part of the process whereby patients themselves learn to assess the likely risks and benefits of treatment and take a more active part in evaluating proposed treatment plans (Wennberg 1990).

The use of standardised scales can also provide a more objective and reliable assessment than informal methods. The use of standard measures of patient function has been advocated in care of the elderly (Dickinson & Young 1989) and in this study there was a period when the use of such a scale was felt to be helping focus inter-disciplinary discussions in geriatrics.

Similarly the use of the Nottingham Health Profile of individual patients in orthopaedics generated interest amongst the nurses who found the questionnaire a useful way to assess the effects of the disease on the patient. Other than these examples the project has not explicitly sought to present changes in health status measures on individual patients - most analyses have been of groups of patients. When individual scores have been presented they have tended to be part of the summaries of patients whose outcomes were poor (in which role they were generally consonant with other data items).

There are considerable uncertainties about how far these types of measures should become part of everyday clinical practice. It may be that they are no better than the assessments already made by experienced doctors as Hopkins has suggested :-

> *"Experienced physicians predict with fair accuracy which patients will do well and which will do badly. The head-shaking that goes on over the coffee in sister's room presumably reflects the unconscious assessment of comparatively simple clinical variables and their relation to probable outcome that are more formally dissected in scales of functional status"*
> (Hopkins 1990 p51)

Our experience would make us less confident on this point especially for acute interventions - if only because the progress of patients beyond their follow-up appointments are largely unknown.

2. Outcomes in clinical audit - local and national

The findings of this study have great relevance in the area of the development of outcomes monitoring in clinical audit. Audit processes typically require similar conditions as outcome measurement, namely a local commitment of resources and local skills in storing and analysing information, and therefore lends itself to the development of outcome measurement. Indeed we would suggest that the addition of outcome measures can increase enthusiasm for the audit process itself and add considerable interest. It is also likely that a combination of process and outcome measures will provide the most fertile environment for improvements in the service. Both the Royal Colleges of Physicians and of Surgeons are encouraging audit and welcome the addition of outcome measures at local or national level (Hopkins & Costain 1990). This does not mean that just because a particular local audit process includes outcome measures it will necessarily be successful, it will still require the type of supportive environment outlined by Shaw & Costain (Shaw & Costain 1989).

The limitations on analysing uncontrolled comparisons of outcome must be borne in mind. Outcome evaluation at one site cannot determine what is the best treatment or what would have happened without treatment. The value of the information critically depends on the extent to which observed outcomes can be related to the desired standards of care and whether the information can be interpreted in such a way as to identify current problems in the service or potential for improvement.

3. Outcomes in resource/local management

Moving away from the purely clinical area, what prospects are there for the relatively sophisticated outcomes measures used in this study becoming an important part of local health service management? Such moves are seen as being generally desirable but not without problems.

A good relationship between the clinicians and other managers will be even more vital in the future NHS. Given the uncertainties over the choice of outcome measures and skills needed in their interpretation, there is an important role for the clinicians in helping to establish the local validity of any results. Our experience has been that they can contribute considerable enthusiasm to the process. A top-down imposition of the criteria for outcome assessment is likely to meet strong opposition from the clinicians or at least apathetic disinterest. The history of performance indicators, which include some outcome measures, points to important problems in this respect (CASPE 1988) in the uneven use of information within the service and the perceived value of that information.

An extension of the approach to outcome measurement that was developed during this project would seem to have much potential. Thus outcome indicators would grow by agreeing local indicators within specialties, formalising these within the resource/local management structure and using the indicators in concert with other information on costs to address issues broadly relating to the best use of resources.

It is possible to envisage a monitoring role for outcome measures in this context. In particular they could be used to examine the relationships between patient outcomes and changes in the organisation of the service. This can be in the form of ensuring that resource or organisational changes do not have negative consequences , as well as checking that changes deliver the health benefit that was envisaged. The availability of the technology for outcome measurement should permit more sophisticated evaluations of different options to be undertaken. For example the growth of day surgery may be accompanied by a program of outcomes monitoring to ensure that care is to the standard required as throughput increases. Similarly the effects of reducing lengths of stay could be monitored in the same way.

As with much other management information it may not be possible to identify hard and fast truths about the best way to deliver care but it should enable the quality of evidence presented and debated to be improved. In the theoretical sense it may also be that

outcome information can help clinicians and managers to identify common priorities and objectives between what are often disparate and competing factions within the service.

4. Hospital management

Beyond the local management of resources at the level of the specialty or clinical team, there is a broader aspect to outcome management that can be considered across the whole hospital. Is it possible to envisage outcomes management being used to address issues at the hospital level?

a. Monitoring individual services/specialties

It is possible to envisage an extension of outcome monitoring at the level of individual service or departments to a system that considers information at a more aggregate level across the whole hospital. The role of the hospital management in this respect would be to support the individual departments in operating at the standards of outcome they have determined for themselves.

Using the degree of detail employed in this study would not be feasible (too much data and probably too expensive) and there would have to be agreed mechanisms for summarising outcome indicators or presenting tracer indicators for tracer conditions. Systems might be based on high-level indicators with a wide coverage of case types collected all the time, leaving detailed information to be used in specific circumstances.

b. You can't have all of the data all of the time

It may be that different levels of information become available to different levels of management within the service. Comprehensive monitoring of all the indicators is likely to be too impractical and expensive. However a combination of general indicators applied across the whole hospital and more specific indicators for selected conditions could prove valuable. There are a variety of ways of selecting the areas for in-depth

analyses:

1. Specific indicator conditions or tracers could be used as examples within one specialty.
2. Poor outcomes , as seen from the general figures, could prompt investigations in specific areas.
3. Issues could be selected which were relevant and of major importance to resourcing decisions eg scale of risk, scale of costs.
4. Areas could be selected on the basis that there was a lack of knowledge of their outcomes.
5. Specialties could be part of a planned programme for particular scrutiny - maybe as part of a hospital wide plan to cover all the major case-types every few years.

c. Differential Resource Allocation.

The ethics and practicalities of differential resource allocation to specialties or particular programmes of care on the basis of their outcomes achieved raises difficult questions. Could and should hospital management decide to spend money on angioplasties because they yielded a greater net benefit than knee replacements? Such questions encompass the issues of the balance between the breadth of services offered as well as the marginal utility of investment in particular programmes of care.

Questions on the relative efficiency of health investments are dear to the heart of economists who have developed the concept of comparative costs per QALY to guide such decisions (Williams 1985, Gudex 1986).

The assessments used in this study have been multi-dimensional and have purposely not sought to distil the different indicators of outcome into a single unitary measure. Leaving aside the technical (Rosser 1990) and moral (Smith 1987) arguments over the QALY (including the question of whether benefit to the health service is directly equivalent to the linear sum of individual patient utilities) there are problems simply trading off outcomes for one condition against those of another. Where there are different indicators

in different specialties, comparing the relative improvements in a functional knee score after knee replacement with the rate of coronary artery bypass surgery after angioplasty tells little about the relative value of the two procedures. Some common measurements have been used, including the marginal change in health status. In some cases it may be inappropriate to consider the benefits of treatment solely in terms of an immediate improvement in health status - as for example in angioplasty where the avoidance of myocardial infarction in unstable angina patients is considered relevant.

One problem arises in assessing changes operating at the margin of current practice. Observing that one elective procedure yields greater benefit than another (on the same yard-stick such as the NHP) only reflects current practice in terms of the nature of the treated patients and the techniques used. It does not necessarily mean that a marginal increase in volume will necessarily deliver the same benefits. For example increasing the volume of angioplasties would change the nature of the presenting population of patients - possibly the procedure will be carried out on less severely ill patients. Will the average benefit per patient be the same with the new case mix receiving treatment - not if the general observation that the sicker you are at the start the more your 'benefit' holds true.

Perhaps more importantly it will not say whether reducing the rate at which an intervention is given will proportionately reduce the benefit or avoid disbenefits. To take an extreme example, stopping cholecystectomies would potentially lead to fatalities amongst patients who were not treated. Outcomes monitoring of the form used at the Freeman says nothing about what would happen to patients if the intervention is not given.

A second problem concerns debates around the weighting and scaling of health status measures that is contentious. Most of these scales were devised as cross sectional tools which weighted different health states according to patient preferences (by some method). This does not necessarily mean that a change in health status would be valued in the same way by others or by the population overall, nor that the relationships are linear across the

scale such that say an improvement in pain of 20% is twice as good as one of 10%.

These factors, and others, will certainly serve to blur simple comparisons of relative utility between treatments. In practice these types of questions are unlikely to be answered mechanistically in isolation from the myriad of other local economic and political issues which influence practice. Nevertheless it is likely that outcome measures can contribute in some way as part of the evidence that guide decisions about the relative benefits of types of care or treatment.

5. Contracting and marketing

In recent months there has been considerable speculation about the possible role of outcome measures in the new contracting process. The White Paper placed a greater emphasis on developing the quality of care and certainly poses difficult problems about how this is to be measured and achieved. District and regional health authorities are making plans and proposals are emerging. Not surprisingly the most likely candidates at the moment are those indicators based on data that is already collected on a large scale (Charlton,Bauer & Lakhani 1984; Middleton 1987).

Is there a potential role for the outcome measures developed in this study here to become part of the purchaser-provider contract? It is unlikely that all the indicators used would be practicable for this purpose but a subset could be envisaged. The types of outcome indicators used in this study might suggest a health authority specify clauses in a contract for a type of elective surgery such as:-

- less than 2% deaths within a year
- 95% of patients to show improvement in health status of x points
- no related re-admissions within 3 months?

Interestingly the Freeman Hospital, which has been given trust status, recognises that outcome measurement can be important in its quality assurance programme and in

attracting potential customers. From the providers point of view outcome measurement can be used in internal quality assurance and management as discussed earlier and also in attracting providers to place care there.

For hospitals who are the first to develop local outcome measures there is an advantage in information terms in that they will be able to say something about the outcomes they achieve. This will provide a marketing advantage in assuring purchasers that they take quality assurance seriously, and even if the outcomes are less then perfect there will be little or no comparable information from other sites. Beyond this stage, it seems possible that outcomes information may come to act as rather sophisticated, and potentially quite effective marketing tool.

For purchasers there is considerable potential in using some of the outcome measures to bring about improvements in quality, though once again problems exist in analysis and in interpretation cannot be underestimated. Purchasers face the difficult problems of specifying and defining outcomes in a form that can be monitored. At present, even using crude measures what constitutes 'desirable outcomes' are largely unknown. For example deaths following elective surgery are undesirable, yet they will inevitably exist in some small proportion through no fault of the quality of care given by providers. The resolution of what can be considered acceptable would be as problematic as it is in the hospital. In such circumstances the ability to compare information between competing providers would seem critical.

One of the most difficult problems will be relating the process of needs assessment to the resulting outcomes. It was observed earlier that for the acute interventions in particular the presenting characteristics can be critical in determining the outcomes achieved. In secondary care the admitting health status represents need but the problem for the purchaser is in determining the need in its resident population and then relating this to the outcomes of treatment that should be expected. This implies the purchaser having knowledge about situations where care is not being provided and might require providers to monitor additional features of patients attending for care.

C. Future research agenda

Outcomes measurement in the future faces two conflicting pressures. On the one hand is the surge in interest in outcomes measures and the requirements of the organisational change in the service. On the other hand are the uncertainties of which measurement to use and how to interpret observed outcomes. Roberts has recently produced an action plan for the development of outcomes in the UK (Roberts 1990) which is comprehensive in its coverage of the issues and provides a useful framework for discussions.

From the perspective of this study the most important issues would be seen as:

Agreement on the dimensions and boundaries of outcome measures

This project identified appropriate ranges of outcome measures in hospital. Do these apply elsewhere? Will different specialties generate indicators of a different nature? This study adopted one approach with a high level of local clinical participation. gained a high level of support but might be thought to have suffered from an over-rich demand for data. Other more top-down approaches would be possible but would certainly induce different responses and requirements.

Methodological issues

Despite the widespread agreement on the use of general health status measures which represent the patient's perspective, there is still relatively little experience of applying these measures in practice. The operating characteristics of these tools are unknown and considerable work needs to be done on their evaluation in a theoretical and practical sense. The scales used in this study may not be the best - but they have by and large done the job expected of them in the clinical case types examined. Would they prove as useful in the rest of secondary and primary care? The results from the Freeman Hospital are only a starting point in the assessment of how outcome indicators can be derived that will enable fair comparisons between centres. There is considerable scope for examining the applicability of these measures in other settings and such work if conducted properly could form

the basis for nationally agreed indicators.

Operationalising outcome measurement and management

This project was organised as an external research team operating within a consenting hospital. For the future there are questions about how outcomes measurement might fit into the organisation of health services. For example within a hospital, there would be advantages in having a central outcomes office to advise specialties and co-ordinate data collection and liaise between managers and clinical staff. Similarly the questions of how outcome information can be used in hospital-wide or purchaser quality assurance mechanisms need to be tackled.

Enabling change in clinical practice

The timescales for examining the effects of outcome information on clinical practice were necessarily limited in this study though some changes have been observed. This work could be enhanced and further research structured to address the important issues of how to turn information into action to improve services .

This research project has exploited techniques from a variety of disciplines - clinical medicine, public health, epidemiology, sociology, information science, statistics, health service management - in order to explore potential improvements to health service management and organisation. It is one of the few projects in the world that has attempted to develop outcome measurement within a real hospital setting across a range of conditions. As a pilot study it has yielded results which point to considerable potential to exploit these measures further within the British health service.

References

Abbott D et al 1988 The impact of diabetes on survival following myocardial infarction of men versus women. JAMA 260:3456-3460

Alberti K G M, Gries F A 1988 Management of Non-insulin dependent Diabetes Mellitus in Europe: A Consensus view. Diabetic Medicine 5:275-281

Anderson T A, Mooney G (Eds) 1990 The challenges of medical practice variations. Macmillan, London.

Applegate W B, Blass J P, Williams F T 1990 Instruments for the functional assessment of older patients. New Engl J Med 322:1207-1214

Baim D S, Ignatious E J 1988 Use of percutaneous transluminal coronary angioplasty: results of a current survey. Am J Cardiol 61:3G-8G

Bardsley M, Coles J, Jenkins L 1987 (Eds) DRGs and health care. The management of case mix. Kings Fund, London.

Bardsley M, Astelle S, McCallum A, Home P D. 1991a A comparison of three approaches to health status/well being in an outpatient diabetes population. Manuscript in preperation. Available from CASPE Research, 14 Palace Court, Bayswater, London W2 4HT.

Bardsley M, Venning P M, Cham C W, Hassan T, Neal D E 1991b A self administered questionnaire in the assessment of symptoms before and after prostatectomy. Submitted to British Journal of Urology.

Bateson M C 1984 Progress in gallstone disease. Brit Med J 289:1163-1164

Batstone G F 1990 Educational aspects of medical audit. Brit Med J 301:326-328

Baumgarten M 1989 The health of persons giving care to the demented elderly: A critical review of the literature. J. Clin. Epidemiol. 42:1137-1148

Bergner M 1989 Quality of life, health status and clinical research. Medical Care 27 No.3 Suppl:S148-S156.

Bergner M, Bobbitt R A, Carter W B 1981 The Sickness Impact Profile: development and final testing of a health status measure. Medical Care 19:787-805

Bergsma J, Engel G L 1988 Quality of life: does measurement help?. Health Policy 10:267-279

Betteridge D J 1989 Lipids, Diabetes, and Vascular Disease: The time to act. Diabetic Medicine 6:195-218

Blalock H M 1981 Social Statistics Revised 2nd edition. McGraw Hill, Singapore.

Bland J M, Altman D G 1986 Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307-310

Bombardier C, Ware J, Russell L J, Larson M, Chalmers A, Reid J L 1986 Auranofin therapy and quality of life of patients with rheumatoid arthritis. Amer J Med 81:565-577

Borch-Johnson K, Kreiner S 1987 Proteinurea - predictor of cardiovascular mortality in insulin dependent diabetes mellitus. Brit Med J 194:1651-1655

Boyle M H, Torrance G W, Sinclair J C, Horwood S P 1983 Economic evaluation of neonatal intensive care of very low-birth-weight infants. New Engl J Med 308:1330-1337

Bradbeer C 1989 Wide variations in surgical mortality. Standard definition needed for postoperative mortality. Brit Med J 298:344-345

Bradley C, Lewis K S 1990 Measures of psychological wein dependent diabetes mellitus. Brit Med J 194:1651-1655

Boyle M H, Torrance G W, Sinclair J C, Horwood S P 1983 Economic evaluation of neonatal intensive care of very low-birth-weight infants. New Engl J Med 308:1330-1337

Bradbeer C 1989 Wide variations in surgical mortality. Standard definition needed for postoperative mortality. Brit Med J 298:344-345

Bradley C, Lewis K S 1990 Measures of psychological wedicine 112:221-226

British Multicentre Study Group 1984 Photocoagulation for proliferative diabetic retinopathy: a randomised controlled trial using the xenon-arc. Diabetologia 1984 26:109-115

Brook R H, Kamberg C J 1987 General Health Status and Outcome measurement: A commentary on measuring functional status. J Chron Dis 40 Suppl:131S-136S

Brook R H, Park R E, Winslow C M, Kosecoff J B, Chassin M R, Hampton J R 1988 Diagnosis and treatment of coronary disease: comparison of doctors attitudes in the USA and UK. The Lancet 1:750-753

Brook R H, Ware J E, Davies-Avery A, et al 1979 Overview of adult health status measures fielded in RAND's health insurance study. Medical Care 17 Suppl:1-131

Brooks W B, Jordan J S, Divine G W, Smith K S, Neelon F A 1990 The impact of psychological factors on measurement of functional status. Assessment of the Sickness Impact Profile. Medical Care 28:793-804

Buck N, Devlin H B, Lunn J N 1987 Report of a confidential enquiry into perioperative deaths. Nuffield Provincial Hospitals Trust, London

Bull B S, Farr M, Meyer P J, Westengard J C, Bacon P A, Stuart J 1989 Efficacy of test used to monitor rheumatoid arthritis. The Lancet 2:965-967

Bulpitt C J, Shaw K M, Hodes C, Bloom A 1976 The symptom patterns of treated diabetic patients. J Chron Dis 29:571-583

Bunker J P 1970 A Comparison of Operations and Surgeons in the United States and in England and Wales. New Engl J Med 282:135-144

Burton K E, Wright V, Richards J 1979 Patients' expectations in relation to outcome of total hip replacement surgery. Annals of Rheumatic Disease 38:471-474

Buxton M J, Acheson R, Cain N, Gibson S, O'Brien B J 1985 The costs and benefits of the heart transplant programmes at Harefield and Papworth Hospitals. DHSS Research Report No.12., London, HMSO

Cageorge S M, Roos L L, Danziger R G 1981 Gall bladder operations: A Population Based Analysis. Medical Care 19:510-524

Cambeau L 1976 Grading of angina Pectoris. Circulation 54:522-523

CASPE 1987 Use and validity of performance indicators - a national survey. CASPE Research, 14 Palace Court, Bayswater, London

CASS 1983a A randomised trial coronary artery surgery. Circulation 68:939-950

CASS 1983b Coronary artery surgery study (CASS). A randomised trial of coronary artery bypass surgery. Quality of life in patients randomly assigned to treatment groups. Circulation 68:951-960

Cassileth B R, Kusk E J, Strouse T B, et al 1984 Pyschosocial status in chronic Illness. A comparative analysis of six diagnostic groups. New Engl J Med 311 No.8:506-511

Chambers M, Clarke A 1990 Measuring readmission rates. Br Med J 301:1134-1136

Charlton J, Hartley R, Silver R, Holland W 1983 Geographical variations in mortality from conditions amenable to medical intervention in England and Wales. The Lancet 1:691-696

Charlton J R H, Bauer R, Lakhani A 1984 Outcome measures for district and regional health care planners. Community Medicine 6:306-315

Chassin M R, Kosecoff J, Solomon D H, Brook R H 1987a How coronary angiography is used: clinical determinants of appropriateness. JAMA 258:2543-2547

Chassin M R, Kosecoff J, Park R E, et al 1987b Does inappropriate use explain geographic variations in the use of health care services?. JAMA 258:2533-2537

Chiverton S G, Inglis J A, Hudd C, Kellet M J, Russell R C G, Wickham J E A 1990 Percutaneous cholecystolithotomy: the first 60 patients. Brit Med J 300:1310-1312

Clarke A 1990 Are readmissions avoidable?. Brit Med J 301:1136-1138

Cleary P, McNeil B 1988 Patient satisfaction as an indicator of quality of care. Inquiry 25:25-36

Cochrane A 1972 Effectiveness and efficiency; random reflections on health services. Nuffield Provincial Hospitals Trust, London

Collin C, Wade D T, Davies S, Horne V 1988 The barthel ADL index: a reliability study. Int Disabil Studies 10:61-63

Craddick J W, Bader B 1983 Medical Management Analysis: A systematic approach to quality assurance and risk management. Auburn: California

Dasbach E J, Fryback D G, Newcomb P A, Klein R, Klein B 1991 Cost-effective strategies for detecting diabetic retinopathy. Medical Care 29 No.1:20-39

Dawes P T 1988 Radiological assessment of outcome in rheumatoid arthritis. British Journal of Rheumatology 27 Suppl I:21-36

Day J L, Spathis M 1988 District diabetes centres in the United Kingdom. Diabetic Medicine 5:372-380

DCCT Reasearch group Control and Complications Trial (DCCT) 1987 Results of a feasibility Study. Diabetes Care 10:1-19

DCCT Research Group 1988 Reliability and Validity of a Diabetes Quality of Life measure for the Diabetes Control of Complications Trial (DCCT). Diabetes Care 11 No. 9:725-732

Department of Health 1989 Working for patients. Department of Health, London

Detre K, Holubkov R, Kelsey S et al 1988 Restenosis after percutaneous transluminal coronary angioplasty in 1985-1986 and 1977-1981. N Engl J Med 318:265-270

Deyo R A 1988 Measuring the quality of life of patients with arthritis. Chapter 11 in Quality of Life: assessment and application. Walker S R & Rosser R M (eds), MTP Press, Lancaster

Deyo R A, Centor R M 1986 Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Disease 11:897-

Deyo R A, Patrick D L 1989 Barriers to the use of health status measures in clinical investigation, patient care and policy research. Medical Care 27 No.3 Suppl:S254-S268

Deyo R A, Unui T S 1984 Towards clinical applications of health status measures: Sensitivity of scales to clinically important changes. Health Services Research 19 No.3:275-289

Diabetes Drafting Group 1985 Prevalence of small vessel and large vessel disease in diabetic patients from 14 centres. Diabetologia 28:615-640

Dickinson E J, Young A 1990 Framework for medical assessment of functional performance. The Lancet 335:778-779

Diehl A K, Beral V 1981 Cholecystectomy and Changing Mortality From Gallbladder Cancer. The Lancet 2:187-189

Donabedian A 1985 Explorations in quality assessment and monitoring. Vols 1-3. Ann Arbor MI, Health Administration Press

Donaldson C, Wright K G, Maynard A, Hamill J D, Sutcliffe E 1987 Day hospitals for the elderly utilisation and performance . Community Medicine 9 No.1:55-61

Donaldson C, Wright K, Maynard A 1986 Determining value for money in day hospital care for the elderly. Age ageing 15:1-7

Dorman J S, Laporte R E, Kuller L H, et al 1984 The Pittsburgh Insulin Dependant Diabetes Mellitus (IDDM) morbidity and Mortality study. Mortality Results Diabetes 33:271-276

Drummond M, Teeling-Smith G, Wells N 1988 Economic evaluation in the development of medicines. Office of Health Economics, London

Drury P L, Tarn A 1985 Blood Pressure distributions WHO values. Diabetic Medicine 2:79-82

Eagles J M, Craig A, Rawlinson F, Restall D B, Beattie J A G, Besson J A O 1987 The psychological well-being of supporters of the demented elderly. British J of Psychiatry 150:293-298

Drummond M, Teeling-Smith G, Wells N 1988 Economic evaluation in the development of medicines. Office of Health Economics, London

Drury P L, Tarn A 1985 Blood Pressure distributions WHO values. Diabetic Medicine 2:79-82

Eagles J M, Craig A, Rawlinson F, Restall D B, Beattie J A G, Besson J A O 1987 The psychological well-being of supporters of the demented elderly. British J of Psychiatry 150:293-298

Ellwood P M 1988 Shattuck Lecture - Outcomes management. A technology of patient experience. N Engl J Med 318:1549-1556

Epstein A M 1990 The outcomes movement - will it get us where we want to go?. New Engl J Med 323:266-269

Feldt-Ramusson B, Mathieson E R, Deckert T 1986 Effect of two years of strict metabolic control on progression of incipient nephropathy in insulin dependent diabetes. The Lancet 2:1301-1304

Finlayson N 1989 Cholecystectomy for gallstones. A good thing if they cause symptoms. Brit Med J 298:133-134

Fitzpatrick R 1990 Measure of patient satisfaction. In Hopkins A, Costain D (Eds) Measuring the outcomes of medical care. London: Royal College of Physicians

Fitzpatrick R, Newman S, Lamb R, Shipley M 1989 A comparison of health status in rheumatoid arthritis. British Journal of Rheumatology 28:201-206

Fletcher A, Hunt B M, Bulpitt C J 1987 Evaluation of Quality of Life in Clinical Trials of cardiovascular disease. J Chron Dis. 40:557-566

Foulds W S, McCuish A, Barrie T et al 1983 Diabetic retinopathy in the west of Scotland: its detection and prevalence, and the cost effectiveness of a proposed screening programme. Health Bull (Edinb) 1983:41-318

Fowler F J, Wennberg J E, Timothy R P, Barry M J, Mulley A G, Hanley D 1988 Symptom status and quality of life following prostatectomy. JAMA 259:3018-3022

Freeman Hospital 1990 Application for Trust Status. Freeman Hospital, Freeman Road, Newcastle upon Tyne UK.

Fries J 1983 Towards an understanding of patient outcome measurement. Arthritis and Rheumatism 26:697-704

Fries J M, Spitz P, Kraines R G, Holman H R 1980 Measurement of patient outcome in arthritis. Arthritis & Rheumatism 23:137-145

Frimoldt-Moller P C, Jensen K M-E, Iversen P, et al 1984 Analysis of presenting symptoms of prostatectomy. J Urol 132:272-276

Gatling W, Mullee M, Hill R D 1986 The prevalence of diabetic nephropathy in a defined population and the projected prevalence in the United Kingdom. Diabetic Medicine 3:381A-382A

Gilhooly A 1984 The impact of care-giving on care givers: Factors associated with the psychological well-being of people supporting a dementing relative in the community. British J of Medical Psychology 57:35-44

Gilleard C J 1985 Predicting the outcome of Psychogeriatric day care. The Gerontologist 25:280-285

Gilleard C J, Belford H, Gilleard E, Whittick J E, Gledhill K 1984 Emotional distress amongst supporters of the elderly mentally infirm. British J of Psychiatry 145:172-177

Glazer J J, Williams M G, Madden S, Rickards A F 1990 Clinical Outcome following balloon angioplasty in 100 consecutive patients with multivessel coronary artery disease. J of Royal Col of Physic (London) 24:292-294

Goldberg V M, Figgie M P, Figgie H E, Heiple K G, Sobel M 1988 Use of a total condylar knee prosthesis for treatment of osteoarthritis and rheumatoid arthritis. Journal of Bone & Joint Surgery 70A No.6:802-811

Gray D, Hampton J R, Bernstein S J, Kosecoff J, Brook R H 1990 Audit of coronary angiography and bypass surgery. The Lancet 335:1317-1320

Griffiths R 1983 NHS Management Inquiry. Department of Health and Social Security, London

Grimley Evans J 1983 Integration of geriatrics with general medical services in Newcastle. The Lancet 1:1430-1433

Gruer R, Gunn A, Gordon D, Ruckley C 1986 Audit of surgical audit. The Lancet 1:23-25

Gruntzig A 1977 Transluminal dilation of coronary-artery stenosis. The Lancet 1:263-

Gudex C 1986 QALYs and their use by the National Health Service. Centre For Health Economics Discussion Paper No.20. University of York

Ham C (Ed) 1988 Health Care Variations: Assessing the Evidence. King's Fund Institute (Research Report No.2) London

Henderson R A, Bucknall C A, Timmis A D, Karani S, Ditsas A, Sowton E 1989 Clinical Outcome of Coronary Angioplasty for Single Vessel Disease. The Lancet 2:546-550

Holmes D R, Reeder G S, Vliestra R E 1988 Role of percutaneous transluminal angioplasty in multivessel disease. Am J Cardiology 61:9G-14G

Home P D, Bardsley M, Pullman G, Gibb I 1991 Glycosylated haemoglobin assays. An important lesson from continuing diabetes audit. BDA conference Brighton .

Hope R A, Longmore J M 1987 Oxford Handbook of clinical medicine. OUP Oxford

Hopkins A 1990 Measuring the quality of care. Royal College of Physicians London:-

Hopkins A, Costain D (Eds) 1990 Measuring the outcomes of medical care. Royal College of Physicians, London

Hunt S M, McEwen J, McKenna S P 1984 Perceived health: age and sex comparisons in a community. J Epid Comm Health 38:156-160

Hunt S M, McEwen J, McKenna S P 1986 Measuring Health Status. Croon Helm, Beckenham

Insall J N, Hood R W, Flawn L B, Sullivan D J 1976 A comparison of four models of total knee replacement. Journal of Bone & Joint Surgery 58A:754-765

Insall J, Scott W N, Ranawat C S 1979 The total condylar knee prosthesis. A report on two hundred and twenty cases. Journal of Bone & Joint Surgery 61A:173-180

Jenkinson C, Fitzpatrick R, Argyle M 1988 The Nottingham Health Profile : An analysis of its sensitivity in differentiating illness groups. Soc Sci Med 27:1411-1414

Kerruish A, Wickings H I, Tarrant P 1988 Empowering managers. Hospital & Health Services Review April:64-67

Kind P, Carr-Hill R 1987 The Nottingham Health Profile: A useful tool for epidemiologist?. Soc Sci Med 25:905-910

Kirwan J R, Reeback J S 1986 Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. British Journal of Rheumatology 25:101-104

Klein R, Klein B E K, Moss S E, Davis M D, DeMets D L 1984 The Wisconsin epidemiologic study of diabetic retinopathy. Arch Ophthalmol 102:520-532

Kohner E M, Barry P J 1984 Prevention of blindness in diabetic retinopathy. Diabetologia 26:173-179

Laing W, Williams R 1989 Diabetes. A model for health care management. Office of Health Economics, London

Lancet editorial 1990. Diabetic Neuropathy. The Lancet 1:1113-1114

Larsen M L, Horder M, Mogensen E F 1990 Effect of long-term monitoring of glycosylated haemoglobin levels in insulin dependent diabetes mellitus. N Engl J Med 323:15-1021

Leighton Read J, Quinn R J, Hoefer M A 1987 Measuring overall health: an evaluation of three important approaches. J Chron Dis 40 Suppl:7S-21S

Leimgruber P P, Roubin J S, Hollman J et al 1986 Restenosis after successful coronary angioplasty in patients with single vessel disease. Circulation 73:710-717

Lewis K S, Bradley C, Knight G, Boulton A J M, Ward J D 1988 A measure of treatment satisfaction designed specifically for people with insulin dependent diabetes. Diabetic Medicine 5:235-242

Lewis K S, Jennings A M, Ward J D, Bradley C 1989 Health belief scales developed specifically for people with tablet-treated type 2 diabetes. Diabetic Medicine 7:148-155

Liaing M H, Cullen K E, Larson M G, Thompson M S, Schwartz J A, Fossel A H, Roberts W N, Sledge C B 1986 Cost-effectiveness of total joint arthroplasty in osteoarthritis. Arthritis & Rheumatism 29:937-943

Liaing M H, Larson M G, Cullen K E, Schwartz J A 1985 Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis & Rheumatism 28:542-547

LoGorfo J et al 1979 Rates of Surgical Care in Prepaid Group Practices and the Independent Setting: What are the reasons for the differences?. Medical Care 17:1-10

Lohr K N 1988 Outcome Measurement: Concepts and questions. Inquiry 25:37-50

MacFarlane J P R, Collings T, Graham K, Macintosh J C 1979 Day hospitals in modern clinical practice - cost benefit. Age Ageing 8(suppl):80-86

Mahoney F I, Barthel D W 1965 Functional evaluation: the Barthel index. Maryland State Med Jnl 14:61-65

Mallya R, Mace B 1981 The assessment of disease activity in rheumatoid arthritis using multi-variate analysis. Rheumatol Rehab 20:14-17

Martin J, Meltzer H, Eliot D 1988 The prevalence of disability amongst adults. OPCS surveys of disability in Great Britain Report 1, HMSO, London

Maxwell R 1984 Quality Assessment in Health. Brit Med J 288:1470-1472

Maynard A 1987 A plea for measuring outcomes. Health & Social Service Journal 26th March 1987:365-365

Mayou R, Bryant B 1987 Quality of life after coronary artery surgery. Qtly J of Medicine 62:239-248

McCance D R, Atkinson A B, Hadden D R, Archer D B 1989 Long-term glycaemia control and diabetic retinopathy. The Lancet 2:824-827

McEwen J 1988 The Nottingham Health Profile. Chapter 5 in Quality of Life: assessment and application. Walker S R & Rosser R M (eds) MTP Press, Lancaster

McIver S, Carr-Hill R 1989 The NHS and its customers. Centre for Health Economics, York.

McKenna F 1988 Clinical and laboratory assessment of outcome in rheumatoid arthritis. British Journal of Rheumatology 27 Suppl:12-20

Meadows K A, Fromson B, Gillespie C, et al 1988 The development, validation and application of a computer linked knowledge questionnaire in diabetic education. Diabetic Medicine 5:61-67

Meenan R F, Anderson J J, Kazis L E, et al 1984 Outcome assessment in clinical trials. Arthritis & Rheumatism 27:1344-1352

Meenan R F, Gertman R M, Mason J H 1980 Measuring health status in arthritis. Arthritis & Rheumatism 23:146-152

Middleton J D 1987 A discussion paper on outcomes for a non-teaching district. Community Medicine 9:343-349

Milne R, Clarke A 1990 Can readmission rates be used as an outcome indicator?. Brit Med J 301:1139-1140

Milner P, Nicholl J, Westlake L et al 1989 The design of a randomized controlled trial for the evaluation of lithotripsy as a treatment for gallstones. J of Lithotripsy and stone disease 1:122-132

Mogensen C E 1988 Management of diabetic renal involvement and disease. The Lancet 1:867-870

Nabarro J D N 1988 Diabetes in the United Kingdom: some facts and figures. Diabetic Medicine 5:816-822

Nathan D M 1990 Haemoglobin $A1_c$. Infatuation or the real thing?. New Engl J Med 323:1062-1063

Nathan D M, Singer D E, Hurxthal K, Goodson J D 1984 The clinical information value of the glycosylated haemooglobin assay. N Engl J Med 310:314-316

Naylor C D, Baigrie R S, Goldman B S, Basinski A 1990 Assessment of priority for coronary revascularisation procedures . The Lancet 335:1070-1073

Neal D E, Ramsden P D, Sharples L, et al 1989 Outcome of elective prostatectomy. Brit Med J 229:762-767

Nelson E C, Berwick D M 1989 The measurement of health status in clinical practice. Medical Care 27 Suppl:S76-S89

Nelson E C, Landgraf J M, Hays R D, Wasson J H, Kirk J W 1990 The functional status of Patients: How can it be measured in Physicians' Offices?. Medical Care 28:1111-1126

Nouri F M, Lincoln N B 1987 An extended activities of daily living for stroke patients. Clinical Rehabilitation 1:301-305

O'Brien B J, Banner N R, Gibson S, Yacoub M H 1988 The nottingham Health Profile as a measure of quality of life following combined heart and lung transplantation. J Epidemiol Community Health 42:232-234

O'Brien B J, Buxton M J, Ferguson B A 1987 Measuring the effectiveness of heart transplant programmes: Quality of life data and their relationship to survival analysis. J Chron Dis 40 Suppl:137S-153S

Opit L J, Greenhill S 1974 Prevalence of gallstones in relation to differing treatment rates for biliary disease. British Journal of Prevention and Social Medicine 28:268-272

Panzram G 1987 Mortality and survival in Type 2 (non insulin dependant diabetes mellitus. Diabetologia 30:123-131

Park R E, Fink A, Brook R H et al, 1986. Physicians ratings of appropriate indications for six medical and surgical procedures. Am J Publ Health 76:766-786

Patrick D L, Deyo R A 1989 Generic and disease-specific measures in assessing health status and quality of life. Medical Care 27 Suppl:S217-S232

Pettigrew R A, Burns H J G, Carter D C 1987 Evaluating surgical risk:the importance of technical factors in determining outcome. Br J Surgery 74:791-794

Pirart J 1978 Diabetes Mellitus and its degenerative complications. Diabetes Care 1:252-263

Qureschi A N, Hodkinson H M 1974 Evaluation of a ten question mental test in the institutionalised elderly. Age Ageing 3:152-157

Rand L I, Krolewski A S, Aiello L M, Warram J H, Baker R S, Maki T 1985 Multiple factors in the prediction of risk of proliferative diabetic retinopathy. New Engl J Med 313:1433-1437

Ransahoff D F, Gracie W A, Wolfenson L B, Neuhauser D 1983 Prophylactic cholecystectomy or expectant management for silent gallstones. A Decision analysis to assess survival. Annals of Internal Medicine 99:199-204

Reifler B V, Larson E, Cox G, Featherstone H 1981 Treatment results at a multi-specialty clinic for the impaired elderly and their families. J of the American Geriatrics Society 29:579-582

Reisenberg D, Glass R M 1989 Editorial: The medical outcomes study. JAMA 262:943-944

Ritchie D, Boyle J, McInnes J, et al 1968 Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. Quart J Medicine 147:393-403

Roberts H 1990 Outcome and Performance in Health Care. Public Finance Foundation Discussion Paper No.33 ,London

Robinson B C 1983 Validation of a care giver strain index. J. Gerontol 38:344-348

Rohan T E, Frost C D, Wald N J 1989 Prevention of blindness in screening for diabetic retinopathy: a quantitative assessment. Brit Med J 299:1198-1201

Roos L L, Brazauskas R, Cohen M, Sharp S M, 1990 Variations in outcomes research. Chapter 3 in Anderson T F & Mooney G (Eds) The challenges of medical practice variations. Macmillan. London

Roos N P, Ramsay E W 1987 A population based study of prostatectomy: outcome associated with different surgical approaches. J Urol 137:1181-1188

Roos N P, Roos L L 1981 High and Low Surgical Rates: Risk Factors for Area Residents. American Journal of Public Health 71:591-600

Rosser R M, Watts V C 1978. The measurement of hospital output. Int J of Epidemiology 1:361-368.

Rosser R M, Kind P 1978. A scale of valuations of states of illness - Is there a social consensus? Int J of Epidemiology 7:359-361.

Rosser R M 1990 From health indicators to quality adjusted life years: technical and ethical issues. In Measuring the outcomes of medical care. Hopkins A & Costain D (Eds), London, Royal College of Physicians

Royal College of Physicians 1989 Medical Audit- a first report what, why and how?. Royal College of Physicians, London

Royal College of Physicians 1990 Working group on functional assessments in the elderly. Royal College of Physicians, London

Royal College of Surgeons 1989 Guidelines to clinical audit in surgical practice. Royal College of Surgeons, London

Rubenstein L Z, Josepson K R, Wieland G D et al 1984 Effectiveness of a geriatric evaluation unit. A randomised clinical trial. New Engl J Med 311:1664-1170

Rupprecht H J, Brennecke R, Kottmeyer M, Berhard G, Erbel T, Pop T, Meyer J 1990 Short- and long-term outcome after PTCA in patients with stable and unstable angina. European Heart Journal 11:964-973

Rutstein D D 1976 Measuring the quality of medical care. N Engl J Med 294:582-588

Salive M E, Mayfield J A, Weissman N W 1990 Patient Outcomes Research Teams and the agency for health care policy and research. Health Services Research 25:5:695-708

Scott D L, Coulton B L, Symmons D P M, Popert A J 1987 Long-term outcome of treating rheumatoid arthritis: results after 20 years. The Lancet 1:1108-1110

Scott E A, Black N 1991 Appropriateness of cholecystectomy in the United Kingdom - A consensus panel approach. To be published in Gut

Serruys P W, Luitjen H E, Beatty K J et al 1988 Incidence of restenosis after successful coronary angioplasty: a time related phenomenon. A quantitative angiographic study in 342 consecutive patients at 1,2,3 and 4 months. Circulation 77:361-371

Shaw C D 1989a Clinical Outcome indicators. Health Trends 21:37-39

Shaw C 1989b Medical Audit. A hospital handbook. King's Fund, London.

Shaw C D, Costain D W 1989 Guidelines for medical audit: seven principles. Brit Med J 299:498-499

Sherwood S, Morris J N, Ruchlin H S 1986 Alternative paths to long term care:Nursing Home, Geriatric day hospital and domiciliary care options. Am J Public Health 76:38-44

Smith A 1987 Qualms about QALYs. The Lancet 1:1134-1136

Smith A 1990 Medical audit - closing the feedback loop is vital (Editorial). Brit Med J 300:65-65

Soloman D H, Brook R H, Fink A, Park R E, Kessey J 1986 Indications for selected medical and surgical procedures - A literature review and ratings of appropriateness. Cholecystectomy. RAND Corporation, Santa Monica USA

Spiegel J S, Spiegel T M, Ward N B, Paulus H E, Leake B, Kane R L 1986 Rehabilitation for rheumatoid arthritis. A controlled trial. Arthritis & Rheumatism 29:628-637

Spilker B, Molinek F R, Johnston K A, Simpson R L, Tilson H H 1990 Quality of life bibliography and indexes. Medical Care 28 Suppl:DS1-DS77

Sprigings D C, Jackson G, Chambers J B e t a l 1988 Balloon dilatation of the aortic valve for inoperable aortic stenosis. Brit Med J 297:1007-1010

Steinbrocker O, Traeger C H, Batterman R C 1949 Therapeutic criteria in rheumatoid arthritis. JAMA 140:659-662

Stewart A L, Greenfield S, Hays R D, Wells K, Rogers W, Berry S, McGlyn E A, Ware J E 1989 Functional Status and Well being of patients with chronic conditions. Results from the medical outcomes study. JAMA 262:907-913

Stewart A L, Hays R D, Ware J E 1988 The MOS short-form General Health Survey: reliability and validity in a patient population. Medical Care 26:714-735

Stickland M H, Paton R C, Wales J K 1984 Haemoglobin $A1_c$ concentrations in men and women with diabetes. Brit Med Journal 289:733-733

Sugrue D D, Vliestra R E, Hammes L N, Holmes D R 1987 Repeat Balloon coronary angioplasty for symptomatic restenosis: a note of caution. European Heart J 8:697-701

Symmons D P M 1988 Mortality in rheumatoid arthritis. British Journal of Rheumatology 27:44-54

Tarlov A R, Ware J E, Greenfield S, Nelson E C, Perrin E, Zubkoff M 1989 The medical outcomes study: An application of methods for monitoring the results of medical care. JAMA 262:925-930

Thompson P W 1988 Functional Outcome in rheumatoid arthritis. British Journal of Rheumatology 27 Suppl:37-43

Treasure T 1990 Angioplasty: where's the proof? British Journal of Hospital Medicine 43:99-100

Tunbridge F T 1991. Care of non-insulin dependent diabetic patients in general practice. Unpublished. Department of Diabetes Medicine, Freeman Hospital, Freeman Road, Newcastle upon Tyne.

Turner R C 1985 UK Prospective Diabetes Study III. Prevalence of hypertension and hypotensive therapy in newly presenting diabetic patients . Hypertension 7 (Suppl 2):8-13

Van Swieten J C, Koudstaal P J, Visser M C, Schouten H J A, Van Gijn J 1988 Interobserver agreement for the assessment of handicap in stroke patients. Stroke 19:605-607

Vayda E A 1973 Comparison of Surgical Rates in Canada and in England and Wales . New Engl J Med 289:1224-1229

Wade D T, Collin C 1988 The barthel ADL index: a standard measure of physical disability. Int Disabil Studies 10:64-67

Wallwork J, Caine N 1985 A comparison of the quality of life of cardiac Transplant patients and coronary artery bypass graft patients before and after surgery. Quality of Life and Cardiovascular Care 1:317-331

Ware J 1987 Standards for validating health measures: Definition and content. J Chron Disease 40:473-480

Weinstein M, et al 1987 Forecasting coronary heart disease incidence mortality and cost: The coronary heart disease policy model. American Journal of Public Health 77:1417-1426

Wells K B, Stewart A L, Hays R D, Burman A, Rogers W, Daniels M, Berry S, Greenfield S, Ware J E 1989 The functioning and well-being of depressed patients. Results from the medical outcomes study. JAMA 262:914-919

Wennberg J E, Bunker J P and Barnes 1980 The Need for assessing the outcome of common medical practices. Annual Review of Public Health 1:277-295

Wennberg J E, Mulley A G, Hamley D, et al 1988 An assessment of prostatectomy for benign urinary tract obstruction: geographic variations and the evaluation of medical care outcomes. JAMA 259:3027-3030

WHO 1988 The International Newsletter on Regional Variations in Health Care. Copenhagen Collaborating Centre 6, World Health Organisation.

Wilkinson G, Borsey D Q, Leslie P, Newton R W, Lind C, Ballinger C B 1988 Pyschiatric morbidity and social problems in patients with insulin dependent diabetes mellitus. Br J Psychiat 153:38-43

Williams A 1985 The economics of coronary artery bypass grafting. Brit Med J 291:736-737

Williams M E 1990 Why screen for functional disability in elderly patients. Annals of Internal Medicine 112:639-639

Winocour P H, Laker K F 1990 Drug therapy for diabetic dyslipoproteinaemia: A practical approach. Diabetic Medicine 7:292-298

Wood J A 1990 Review of diabetes care initiative in primary care settings. Health Trends 19:39-43

Woolf A D 1988 Setting the scene and posing the questions. British Journal of Rheumatology 27 Suppl I:1-4

Acknowledgements

I would like to express my thanks to my supervisor, Brian Abel-Smith and all members of the Steering Group and all staff at the Freeman Hospital who have helped during the course of the project.

In particular I would like to thank clinical staff who have been important in in data collection and discussion of outcome indicators including Joan Watson, Neil Wrightson, Felicity Tunbridge, Angela Skinner, Phillip Home, Judy Wilbourne, Chris Venables, Peter Wright, Douglas Reid, Ian Pinder, Steve Smith, Ian Griffiths, Catherine Streit and Stewart Parker.

I owe a special debt of gratitude to colleauges in the CASPE team, both past and present, including Paul Beveridge, Geraldine Byrne, James Coles, Lynda Dearden, Julie Goodfellow, Donna Swinden and Pam Venning.

Appendix 1

STEERING GROUP MEMBERS

Dr. G. Pledger	(Chairman) Director of Public Health (NHA)
Dr. G. Sanders	Deputy Director of Public Health (NHA)
Mr. C. Marshall	District General Manager
Dr. I. Wickings	Director CASPE (London)
Mr. J. Coles	Assistant Director CASPE (London)
Mr. M. Bardsley	Research Manager CASPE (London/Newcastle)
Dr. A. Lakhani	Dept of Health (London)
Mr. A. Beeston	Director of Nursing
Mr. I. Tarbit	Admin. Freeman Hospital
Mr. L. Fenwick	Chief Executive Freeman Hospital
Miss M. Davis	Admin. Freeman Hospital
Dr. P. Home	Consultant Physician (Diabetes)
Prof. K. Alberti	Dept. of Medicine
Mr. I. Pinder	Orthopaedics
Mr. S. Smith	Orthopaedics
Dr. I. Griffiths	Rheumatology
Dr. P. Platt	Rheumatology
Mr. C. Venables	General Surgery
Mr. P. Wright	General Surgery
Dr. D. Reid	Cardiology
Prof. O. James	Dept. Geriatrics
Dr. R. Cooper	Dept. Geriatrics
Ms. J. Goodfellow	Senior Research Assistant CASPE (Newcastle)
Ms. D. Swinden	Research Assistant CASPE (Newcastle)

Appendix 2(a) SICKNESS IMPACT PROFILE - SIP (Subset of complete instrument)

PLEASE TICK YES FOR EACH STATEMENT THAT YOU ARE SURE DESCRIBES YOU AT THE MOMENT AND IS RELATED TO YOUR STATE OF HEALTH.

This set of statements describes walking and use of stairs.	Yes
I walk shorter distances or often stop for a rest.	[ ]
I do not walk up or down hills.	[ ]
I only use stairs with a physical aid - for example, a handrail, stick or crutches.	[ ]
I only go up and down stairs with assistance from someone else.	[ ]
I get about in a wheelchair.	[ ]
I do not walk at all.	[ ]
I walk by myself but with some difficulty - for example, I limp, wobble, stumble or I have a stiff leg.	[ ]
I do not use stairs at all.	[ ]
I get about only by using a walking frame, crutches, stick, walls, or hold onto furniture.	[ ]
I walk more slowly.	[ ]

The following statements describe the activities you usually do in your spare time - for relaxation, entertainment, or just to pass the time.

I spend shorter periods of time on my hobbies and recreation.	[ ]
I go out to enjoy myself less often.	[ ]
I am cutting down on SOME of my usual inactive pastimes - for example, I watch less TV, play cards less or read less.	[ ]
I am not doing ANY of my usual inactive pastimes - for example, I do not watch TV, play cards or read.	[ ]
I am doing more inactive pastimes instead of my other usual activities.	[ ]
I take part in fewer community activities.	[ ]
I am cutting down on some of my usual physical recreation or more active pastimes.	[ ]
I am not doing ANY of my usual physical recreation or more active pastimes.	[ ]

Please Turn Over

The following statements describe your contact with your family and friends.

	YES
I go out to visit people less often.	[ ]
I do not go out to visit people at all.	[ ]
I show less interest in other people's problems - for example, I don't listen when they tell me about their problems,	[ ]
I don't offer help.	[ ]
I show less affection.	[ ]
I take part in fewer social activities than I used to - for example, I go to fewer parties or social events.	[ ]
I am cutting down the length of visits with friends.	[ ]
I avoid having visitors.	[ ]
My sexual activity is decreased.	[ ]
I often express concern over what might be happening to my health.	[ ]
I talk less with other people.	[ ]
I make many demands on other people - for example, I insist that they do things for me or tell them how to do things.	[ ]
I stay alone much of the time.	[ ]
I am disagreeable with my family - for example, I act stubbornly or spitefully.	[ ]
I frequently get angry with my family - for example, I hit them, scream or throw things at them.	[ ]
I isolate myself as much as I can from the rest of my family	[ ]
I pay less attention to the children.	[ ]
I refuse contact with my family - for example, I turn away from them.	[ ]
I do not look after my children or family as well as I usually do.	[ ]
I do not joke with members of my family as much as I usually do	[ ]

Please Turn Over

This group of statements covers work that you do on a regular basis.

	YES
I am retired.	[ ]
My work has not been affected by my health.	[ ]
I do not work at all because of my health.	[ ]
I only do housework or work around the house for short periods of time or I rest often.	[ ]
I do less of the daily household chores than I would usually do.	[ ]
I do not do any of the daily household chores that I would usually do.	[ ]
I do not do any of the maintenance or repair work that I would usually do in my home or garden.	[ ]
I do not do any of the shopping that I would usually do.	[ ]
I do not do any of the cleaning that I would usually do.	[ ]
I have difficulty using my hands - for example, turning taps, using kitchen gadgets, sewing or doing repairs.	[ ]
I do not do any of the clothes washing that I would normally do	[ ]
I do not do heavy work around the house.	[ ]
I have given up taking care of personal or household business affairs - for example, paying bills, banking and doing household accounts.	[ ]

Are there any comments you would like to make?

Thank you for your help.

Appendix 2(b)

NOTTINGHAM HEALTH PROFILE

Listed below are some problems people may have in their daily life. Look down the list and put a tick in the box under **YES** for any problem you have at the moment. Tick the box under **NO** for any problem you do not have.

Please answer every question. If you are not sure whether to say Yes or No, tick whichever answer you think is more true at the moment.

	YES	NO
I'm tired all the time	[ ]	[ ]
I have pain at night	[ ]	[ ]
Things are getting me down	[ ]	[ ]
I have unbearable pain	[ ]	[ ]
I take tablets to help me sleep	[ ]	[ ]
I've forgotten what it's like to enjoy myself	[ ]	[ ]
I'm feeling on edge	[ ]	[ ]
I find it painful to change position	[ ]	[ ]
I feel lonely	[ ]	[ ]
I can only walk about indoors	[ ]	[ ]
I find it hard to bend	[ ]	[ ]
Everything is an effort	[ ]	[ ]
I'm waking up in the early hours of the morning.	[ ]	[ ]
I'm unable to walk at all	[ ]	[ ]
I'm finding it hard to make contact with people	[ ]	[ ]
The days seem to drag	[ ]	[ ]
I have trouble getting up & down stairs or steps	[ ]	[ ]
I find it hard to reach for things	[ ]	[ ]

Remember, if you are not sure whether to answer yes or no to a problem, tick whichever you think is **more true** at the moment.

	YES	NO
I'm in pain when I walk	[ ]	[ ]
I lose my temper easily these days	[ ]	[ ]
I feel there is nobody I am close to	[ ]	[ ]
I lie awake for most of the night	[ ]	[ ]
I feel as if I'm losing control	[ ]	[ ]
I'm in pain when I'm standing	[ ]	[ ]
I find it hard to dress myself	[ ]	[ ]
I soon run out of energy	[ ]	[ ]
I find it hard to stand for long (eg.at the kitchen sink, bus queue)	[ ]	[ ]
I'm in constant pain	[ ]	[ ]
It takes me a long time to get to sleep	[ ]	[ ]
I feel I am a burden to people	[ ]	[ ]
Worry is keeping me awake at night	[ ]	[ ]
I feel that life is not worth living	[ ]	[ ]
I sleep badly at night	[ ]	[ ]
I'm finding it hard to get on with people	[ ]	[ ]
I need help to walk about outside (eg.a walking aid,someone to support)	[ ]	[ ]
I'm in pain when going up & down stairs or steps	[ ]	[ ]
I wake up feeling depressed	[ ]	[ ]
I'm in pain when I'm sitting	[ ]	[ ]

Now please go back to the beginning and make sure you have answered every question.

PART II

Now we would like you to think about the activities in your life which may be affected by health problems. In the list below, tick **YES** for each activity in your life which is being affected by the state of your health. Tick **NO** for each activity which is not being affected, **or which does not apply to you.**

Is your present state of health causing problems with your....

	YES	NO
Job of work (That is, paid employment)	[ ]	[ ]
Looking after the house (eg.cleaning & cooking, repairs & odd jobs around the house)	[ ]	[ ]
Social life (eg. going out, seeing friends, going to the pub)	[ ]	[ ]
Home life (That is: relationships with other people in your home)	[ ]	[ ]
Sex life	[ ]	[ ]
Interests and hobbies (eg.sports,crafts & arts)	[ ]	[ ]
Holidays (eg.summer/winter holidays, weekends away)	[ ]	[ ]

Thank you for your help.

Appendix 2(c) HEALTH ASSESSMENT QUESTIONNAIRE

WE ARE INTERESTED IN LEARNING HOW YOUR ILLNESS AFFECTS YOUR ABILITY TO FUNCTION IN DAILY LIFE. PLEASE FEEL FREE TO ADD ANY COMMENTS AT THE END OF THIS FORM.

PLEASE TICK THE ONE RESPONSE WHICH BEST DESCRIBES YOUR USUAL ABILITIES

	Without ANY difficulty	With SOME difficulty	With MUCH difficulty	Unable to do

1.DRESSING AND GROOMING
Are you able to:
- Dress yourself, including tying shoelaces and doing buttons?

- Shampoo you hair?

2.RISING
Are you able to:
- Stand up from an armless straight chair?

- Get in and out of bed?

3.EATING
Are you able to:
- Cut your meat?

- Lift a full cup or glass to your mouth?

- Open a new carton of milk (or soap powder)?

4.WALKING
Are you able to:
- Walk outdoors on flat ground?

- Climb up five steps?

PLEASE TICK ANY AIDS OR DEVICES THAT YOU USUALLY USE FOR ANY OF THESE ACTIVITIES:

.....Cane
.....Walking frame
.....Crutches
.....Wheelchair
.....Devices used for dressing(button hook, zipper pull, long handled shoe horn etc)
.....Built-up or special utensils
.....Special built-up chair

Other(specify).....

PLEASE TICK ANY CATEGORIES FOR WHICH YOU USUALLY NEED HELP FROM ANOTHER PERSON:

.....Dressing and Grooming
.....Rising
.....Eating
.....Walking

PLEASE TURN OVER

PLEASE TICK THE ONE RESPONSE WHICH BEST DESCRIBES YOUR USUAL ABILITIES OVER THE PAST WEEK

	Without ANY difficulty	With SOME difficulty	With MUCH difficulty	Unable to do

5.HYGIENE
Are you able to:
- Wash and dry your entire body?

- Take a bath?

- Get on and off the toilet?

6.REACH
Are you able to:
- Reach and get down a 5lb object (eg. a bag of potatoes) from just above your head?

- Bend down to pick up clothing from the floor?

7.GRIP
Are you able to:

- Open car doors?

- Open jars which have been previously opened?

- Turn taps on and off?

8.ACTIVITIES
Are you able to:

- Run errands and shop?

- Get in and out of a car?

- Do chores such as vacuuming, housework or light gardening?

PLEASE TICK ANY AIDS OR DEVICES THAT YOU USUALLY USE FOR ANY OF THESE ACTIVITIES:

.....Raised toilet seat .....Bath rail
.....Bath seat .....Long handled appliances for reach
.....Jar opener(for jars previously opened)

Other(specify).....

PLEASE TICK ANY CATEGORIES FOR WHICH YOU USUALLY NEED HELP FROM ANOTHER PERSON:

.....Hygiene .....Gripping and opening things
.....Reach .....Errands and housework

Thank You

Appendix 2(d) VALIDATION PROFORMA FOR ANGIOPLASTY PATIENTS

Name: JG.NW
Date:

BEFORE ANGIOPLASTY

How long have you suffered from anginal pain?
HISTORY OF ?:

MI	[ ]	CCF	[ ]
prev. CABG	[ ]	PBP	[ ]
prev. PTCA	[ ]	Smoker	[ ]
chron.resp.dis.	[ ]	Hyper.lip.	[ ]

How far could you walk each day before operation?

What limited exercise? angina [ ] sob [ ] fatigue [ ] other [ ]

What activities brought on your angina?

Unusual strenuous activity.
Walking..... how far.....
Stress,cold or windy, meals.
Climbing stairs (one flight)
Washing,dressing.

AFTER ANGIOPLASTY

Since Operation:
Have you been seen at A/E?

Have you been readmitted to hospital?

What medication?

How far can you now walk each day?

What limits your exercise? Angina [ ] SOB [ ] Fatigue [ ] Other [ ]

Are you more or less active than when you first suffered from angina?

More [ ] Less [ ]

Still suffer yes [ ] no [ ]

If yes -

What activities bring on angina?	YES	NO
Unusual strenuous activity	[ ]	[ ]
Walking	[ ]	[ ]
How far		
Emotion, cold, windy, meals	[ ]	[ ]
One flight stairs	[ ]	[ ]
Dressing, washing etc.	[ ]	[ ]

Appendix 2(e) VALIDATION PROFORMA FOR CHOLECYSTECTOMY PATIENTS

Name: JG JW

Date:

- Introduce yourself....
- What the project is....
- Purpose of visit: - to check people understand questions,
  - to see if anything has been missed,
  - to find ways to improve project.

- Interview takes approx. 30 mins.

1/BEFORE THE OPERATION

- When did the symptoms first appear?

- When was gall bladder problem first diagnosed?

- What symptoms were there?
  (pain,flat.,dist.,vomit,appet.,fatty,bowels*)
  * Probe bowels if reported as a problem.

- How would you rate your overall health?
  poor[ ] fair[ ] good[ ] v.good[ ]

2/AT PRESENT

- Did all the symptoms disappear after the operation?

- If not ..... - which are still present? *
  * Probe bowels if reported as a problem.

- How would you rate your overall health?
  poor[ ] fair[ ] good[ ] v.good[ ]

3/NEW PROBLEMS

- Have you seen your GP with a complication?

- Have you been readmitted to hospital/seen at A+E since your operation?

- Do you have any health problems that you did not have before?

Appendix 2(f) VALIDATION PROFORMA FOR ORTHOPAEDIC PATIENTS

NAME:

DATE:

Introduction .....
What the project is .....
Purpose of visit:- to check people understand questions.
to see if anything has been missed.
to find way to improve project

Interview takes approximately 30 minutes.

BEFORE THE OPERATION

What kind of surgery did you have?

new repacement
revision
bilateral

Have you had any other replacement surgery?

What was the extent of your symptons (arthritis) before the operation?

(How many joints affected)

Did you suffer from any other problems?

How would you rate your overall health?

poor [ ] fair [ ] good [ ] v good [ ]

KNEE SCORES

	BEFORE	AFTER	BETTER	SAME	WORSE
Walking pain					
Rest pain					
Climb stairs					
Transfer					
Walk/Stand					
Assisted by					

AT PRESENT

Any problems after the operation?

(how many follow-ups)

Any other health problems?

How would you rate your overall health?
poor [ ] fair [ ] good [ ] v good [ ]

Appendix 2(g) VALIDATION PROFORMA FOR RHEUMATOLOGY PATIENTS

NAME:
HOSPITAL NUMBER:

DATE:

Introduction:
Explain purpose of study:
Explain purpose of interview:
To identify any problems with questionnaires
To check people understand the questions
To find way to improve it, if necessary

Interview will take about 15 - 20 mins.

SECTION ONE - general

1. Do you remember completing the forms?

2. When was it given?

   Who by?

3. Do you know where forms are kept on completion?

4. Did this in any way affect your replies?

5. The form had your name on it, did this affect your replies at all?

6. Were there any questions you did not feel happy about answering?

7. Did you have any problems completing the forms?

(+ NHP eg. DIMENSION 1 - 7)

Appendix 2(h) PROFORMA FOR VALIDATION OF NHP IN:

- Cholecystectomy
- angioplasty
- orthopaedics

There are 6 dimensions and a general health section in the NHP. Each dimension should be assessed individually, on a continuum of: "none" to "high" depending on the degree to which the individual is affected. Use the prompts to structure the interview, if necessary.

Dimension 1 - PAIN

Prompt - Do you suffer from pain?
none[ ] low[ ] medium[ ] high[ ]
Comments:

Dimension 2 - MOBILITY

Prompt - Is your mobility restricted in any way?
none[ ] low[ ] medium[ ] high[ ]
Comments:

Dimension 3 - ENERGY

Prompt - Do you often feel tired?
none[ ] low[ ] medium[ ] high[ ]
Comments:

Dimension 4 - SLEEP

Prompt - Do you have any problems with sleeping?
none[ ] low[ ] medium[ ] high[ ]
Comments:

Dimension 5 - EMOTION

Prompt - How have you been feeling lately (eg happy, depressed)?
none[ ] low[ ] medium[ ] high[ ]
Comments:

Dimension 6 - SOCIAL ISOLATION

Prompt - Do you ever feel lonely/isolated?
none[ ] low[ ] medium[ ] high[ ]
Comments:

GENERAL HEALTH

Prompt - Are there any aspects of your life that you feel are affected by your health?

QUESTIONS ABOUT THE FORMS

Do you remember completing the forms?
Did you have any problems completing the forms?
Do you have any other comments about the forms/interview?

Thank you for your time and cooperation.

Appendix 3 - Response to questionnaire circulated to clinicians

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

   Yes. We are now clear that the data can be collected, examined against standards, and that the result can be present in useful form. All those who have seen the analysis have found it useful.

2. Did you find the outcome information and/or interesting? In what way?

   1. We have a much greater idea of the degree to which we were failing to meet targets.
   2. We have identified areas of care in need of particular attention (eg. BP/lipids).
   3. We have identified a major problem in one of the most important biochemical tests.

3. a) Has the information influenced the way you see the service or provide care?
   b) Have there been any changes as a result of this study?
   If so, in what way?
   If not, why not?

   a. We have devised and implemented new protocols of care as a result.
   b. We have introduced formal targetting.
   c. There is clearly a need for more attention to agreed protocols of care by clinic staff.
   d. We are to replace the present glycosylated haemoglobin assay.

4. In what ways could the study have been improved?

   We could have used more time in diabetes alone - eg for inpatients, new insulin starters. These projects have only just got going.

5. Do you intend to continue outcomes monitoring in your speciality?

   We are doing. We are also extending the ideas to the Northern Region and nationally and to GP's.

6. Has the study helped you in any other ways?

   Given a large boost to our general audit activities and expanded our views of what it might achieve.

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

Yes - within the time scale.

2. Did you find the outcome information useful and/or interesting? In what way?

Interesting as it showed elements of the patient's response to treatment he would not normally have access to.

3. a) Has the information influenced the way you see the service or provide care?

No

b) Have there been any changes as a result of this study?
If so, in what way?
If not, why not?

No

I would not expect this to occur with the time scale or without a comparative group to assess our results against.

4.In what ways could the study have been improved?

By a "comparative" study using a similar format.

5. Do you intend to continue outcomes monitoring in your specialty?

As far as is practicable but we do not have the resources in staff/record keeping to make a full continuing study practical.

6. Has the study helped you in any other ways?

Useful in identifying what is happening and surprising in the apparent value of the Nottingham Health Profile.

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

   Yes

2. Did you find the outcome information useful and/or interesting? In what way?

   Yes.

   Enables detailed review after a busy clinic with high lighting of problems.

3. a) Has the information influenced the way you see the service or provide care?

   Yes

   b) Have there been any changes as a result of this study?
   If so, in what way?
   If not, why not?

   Marked reluctance to operate on younger patients.

4. In what ways could the study have been improved?

   Screening predecisions as suitable for surveying the psychological profile

5. Do you intend to continue outcomes monitoring in your specialty?

   Yes

6. Has the study helped you in any other ways?

   Collated the results without as much involvement.

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

   As of yet in Orthopaedics still not at our one year objective which will be of value. We have certainly learnt the value of measurement but require further analysis to assess our conclusions.

2. Did you find the outcome information useful and/or interesting?
   In what way?

   Useful yes furthers the analysis of result of knee replacement.

   Interesting yes. In our project we have had excellent feedback and initial results appear to isolate a subgroup of patients which deserve further study and analysis.

3. a) Has the information influenced the way you see the service or provide care?

   b) Have there been any changes as a result of this study?
      If so, in what way?
      If not, why not?

   Yes. See above.

   The group identified require further study the determine which particular features are common to this group of young patients having TKR.

4. In what ways could the study have been improved?

   Within the restrictions appeared to be reasonably designed as a preliminary project. I think improvements can be made <u>now</u> following our initial project.

5. Do you intend to continue outcomes monitoring in your specialty?

   Yes

6. Has the study helped you in any other ways?

   Improved our understanding of outcomes study and value.

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

Yes. Rheumatology project not yet completed but data available has prompted re-evaluation of service.

2. Did you find the outcome information useful and/or interesting? In what way?

Yes. (a) sets diseases in context if a 'general' health status (eg.NNP) is used.
(b) demonstrates deficiencies in treatment measures for the severely disabled
(c) will allow some rationalisation of use of resources.

3. a) Has the information influenced the way you see the service or provide care?

b) Have there been any changes as a result of this study?
If so, in what way?
If not, why not?

(a) Yes.

(b) developing a more integrated service for cervical myopathy (commonest single cause of death)

4. In what ways could the study have been improved?

Given the constraints on time (and resources) available, difficult to see another approach that could have been more profitable.

No shortage of ideas for further studies!

5. Do you intend to continue outcomes monitoring in your specialty?

Yes - difficult to know whether it should be seen as a "project" type approach or "routine" at this stage. Ultimately, must be routine but best measures and logistics (without extra resource) need to be sorted out.

6. Has the study helped you in any other ways?

Invaluable contribution from an external group (CASPE) in formulating design and analysis of outcome data.

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

   Yes.

2. Did you find the outcome information useful and/or interesting? In what way?

   Useful and interesting.
   30% of patients die from the day of admission to 3/12 after discharge. I should have guessed a lower percentage (as doctors tend to do) but I am not very surprised at the finding as it reflects the multi-system impairment and the frailty and age of many of our patients.

3. a) Has the information influenced the way you see the service or provide care?

   b) Have there been any changes as a result of this study?
   If so, in what way?
   If not, why not?

   Not so far.

   a. It will no doubt influence our management. eg. re-admissions to 3/12 seems very high but again may reflect the frailty and the additional reason for the last admission. The findings will make us look more closely at the reasons for re-admission, and attempt to improve them.

   The apparent difference in pattern of care seems to reflect degree of disability/illness on admission, and perhaps there is a difference in age.

4. In what ways could the study have been improved?

   I am disappointed by the low numbers and incomplete assessment of care-givers obtain. We need to improve our information on this aspect of discharge and maintenance.

5. Do you intend to continue outcomes monitoring in your specialty?

   Yes

6. Has the study helped you in any other ways?

   It has focussed our attention on what we are doing and will help us to think about it more, as audit our work more often and more carefully.

Freeman Outcome study - Questions for clinicians.

The study has been concerned with examining the feasibility of outcome measures and the usefulness of that information in informaing practice at all levels of the service.

1. Did you feel that the project achieved these objectives?

Yes. There has been some 'discomfort' in the use of the 'softer' outcome measures in such a chronic ongoing disease such as CAD.

2. Did you find the outcome information useful and/or interesting? In what way?

Both useful and interesting - to have a more complete follow up in cases who have had PTCA.

3. a) Has the information influenced the way you see the service or provide care?

Yes.

b) Have there been any changes as a result of this study?
If so, in what way?
If not, why not?

The major change will be withdrawal of antianginal drugs.

4.In what ways could the study have been improved?

1. The outcomes aims are probably slightly different in unstable angina and stable angina and maybe this should have been incorporated into the design eg. Thus in UAP the usual aim is to stabilise patient so that can leave hospital without AMI. Subsequent management could easily be CABG at some point and one would expect CABG rate to be higher. In stable angina relief of symptoms.

2. Classification of angina should have been by cardiologist, not retrospectively from notes/reports.

5. Do you intend to continue outcomes monitoring in your specialty?

Yes

6. Has the study helped you in any other ways?

It has highlighted something we were always aware of - the difficulties in monitoring outcomes in a chronic condition which is so heterogeneous (Thus sometimes angioplasty is done to allow another procedure, eg. renal transplant, to be done more safely - sometimes to stabilise symptoms, sometimes to relief symptoms, sometimes to hopefully 'protect' myocardium) as is the patients' response to their illness.

Appendix 4(a) Data Collection Proforma - Care of the Elderly (In-patients)
QUESTIONNAIRE 1

Page 1

ADMISSION ASSESSMENT:

Date admitted:
Date assessed:
Name:
Date discharge:
Address:
Date death:
Consultant: OJ RC

Hospital no.:
Age:
Marital status:

Main Care - Giver:
(relationship)

Spouse	[ ]
Child	[ ]
Sibling	[ ]
Informal	[ ]
Other relation	[ ]

Admitted from home [ ]
Transferred from ward [ ]
Transferred from another hospital [ ]

Family Relationships: (Are there any significant problems?)	Attention required	Not relevant
	[ ]	[ ]

Abbreviated Mental Test (AMT):
(Each question scores one mark)

		Adm.	Dis.
1.	Age	[ ]	[ ]
2.	Time (to nearest hour)	[ ]	[ ]
3.	Address for recall at end of test - this should be repeated to patient to ensure it has been heard correctly: 42 West St.	[ ]	[ ]
4.	Year	[ ]	[ ]
5.	Name of institution	[ ]	[ ]
6.	Recognition of 2 persons	[ ]	[ ]
7.	DOB (day & month sufficient)	[ ]	[ ]
8.	Year 2nd world war	[ ]	[ ]
9.	Name of present monarch	[ ]	[ ]
10.	Count backwards 20 - 1	[ ]	[ ]
Mental Test Score		[ ]	[ ]
Rating (S,M or N)		[ ]	[ ]
Cognitive Score Rating:	[0-3] Severe	[4-6] Moderate	[7-10] Normal

This relates to disease relevant to this admission to hospital.

Significant Medical Conditions:

	ACTIVE	INACTIVE	Specify
Cardiovascular	[ ]	[ ]	
Respiratory	[ ]	[ ]	
Gastro-intestinal	[ ]	[ ]	
Musculo-skeletal	[ ]	[ ]	
Renal	[ ]	[ ]	
Genito-urinary	[ ]	[ ]	
Metabolic	[ ]	[ ]	
Nervous system	[ ]	[ ]	
Psychiatric	[ ]	[ ]	
Skin	[ ]	[ ]	
Social	[ ]	[ ]	

DISCHARGE ASSESSMENT

	Better	**Same**	**Worse**	**Died**
General Status:	[ ]	[ ]	[ ]	[ ]
Family Relationships: (Have aspects of care-givers strain been attended to?)	[ ]	[ ]	[ ]	
Medical Condition: (In relation to active problems at admission)	[ ]	[ ]	[ ]	

Mental Test Score: Admission [ ]
Discharge [ ]

Have goals been achieved overall?

Yes[ ] No[ ] Uncertain[ ]

QUESTIONNAIRE 2

PATIENT NAME: **HOSPITAL NO.:**

BARTHEL ADL INDEX

		Adm.	Goal	Dis.
Bowels	Incontinent	[ ]	[ ]	[ ]
	Occasional accident	[ ]	[ ]	[ ]
	Continent	[ ]	[ ]	[ ]
Bladder	Incont. / catheterised	[ ]	[ ]	[ ]
	Occasional accident	[ ]	[ ]	[ ]
	Continent(for over 7days)	[ ]	[ ]	[ ]
Grooming	Needs help	[ ]	[ ]	[ ]
	Independent, face/hair,etc	[ ]	[ ]	[ ]
Toilet use	Dependent	[ ]	[ ]	[ ]
	Needs some help	[ ]	[ ]	[ ]
	Independent	[ ]	[ ]	[ ]
Feeding	Unable	[ ]	[ ]	[ ]
	Needs help,eg-cutting	[ ]	[ ]	[ ]
	Independent	[ ]	[ ]	[ ]
Dressing	Dependent	[ ]	[ ]	[ ]
	Needs help(1/2 unaided)	[ ]	[ ]	[ ]
	Independent	[ ]	[ ]	[ ]
Bathing	Dependent	[ ]	[ ]	[ ]
	Independent	[ ]	[ ]	[ ]
Mobility	Immobile	[ ]	[ ]	[ ]
	Walks-help of 1(verb/phys)	[ ]	[ ]	[ ]
	Wheelchair indep.	[ ]	[ ]	[ ]
	Independent	[ ]	[ ]	[ ]
Transfer	Unable	[ ]	[ ]	[ ]
	Major help(1-4 people)	[ ]	[ ]	[ ]
	Minor help(verb/phys)	[ ]	[ ]	[ ]
	Independent	[ ]	[ ]	[ ]
Stairs	Unable	[ ]	[ ]	[ ]
	Needs help(verb/phys)	[ ]	[ ]	[ ]
	Independent up & down	[ ]	[ ]	[ ]
Home environment				
	Home alone	[ ]	[ ]	[ ]
	Home with soc.services	[ ]	[ ]	[ ]
	Home with nursing prof.	[ ]	[ ]	[ ]
	Home with other	[ ]	[ ]	[ ]
	Sheltered accom.	[ ]	[ ]	[ ]
	Residential care L.A.	[ ]	[ ]	[ ]
	" " Priv.	[ ]	[ ]	[ ]
	Nursing care H.A. Rehab.	[ ]	[ ]	[ ]
	" " Long-T	[ ]	[ ]	[ ]
	" " Private	[ ]	[ ]	[ ]

QUESTIONNAIRE 3 **QUESTIONNAIRE FOR CARERS**

Patient name: Hosp. No: Date:

Here is a list of things which other people have found difficult when caring for a relative /friend /neighbour. Can you say whether any of these apply to you? (examples are given just as a guide). Please tick either YES or NO for every item.

	YES	NO
1.**Sleep is disturbed.** (eg.because he/ she is in & out of bed at night)	[ ]	[ ]
2.**It is inconvenient.** (eg.because helping takes too much time/or his/her house is a long way away)	[ ]	[ ]
3.**It is a physical strain.** (eg.because he/she is hard to lift)	[ ]	[ ]
4.**It is restricting.** (eg.helping means no free time/can't go out much)	[ ]	[ ]
5.**There have been family disruptions.** (eg.there is no privacy/usual routine is disrupted)	[ ]	[ ]
6.**There have been changes in personal plans** (eg.holiday,had to turn down a job	[ ]	[ ]
7.**There have been other demands on my time** (eg.from the rest of my family)	[ ]	[ ]
8.**There have been emotional adjustments.** (eg.arguments,etc.)	[ ]	[ ]
9.**Some behaviour is upsetting.** (eg.incontinence,forgetfulness)	[ ]	[ ]
10.**It is upsetting that he/she has changed so much.** (eg.seems a different person since stroke).	[ ]	[ ]
11.**There have been work adjustments.** (eg.keep having to take time off work)	[ ]	[ ]
12.**It is a financial strain.**	[ ]	[ ]
13.**I feel completely overwhelmed.** (eg.worry of how you will cope).	[ ]	[ ]

Thank you for your help.

Appendix 4(b) Report on Outcomes - Care of The Elderly Feb 1991

1/PRESENTING CHARACTERISTICS

a/Sample: - 420 patients
- 241 Professor James' caseload (Oct.'89-December '90)- OJ.
- 179 Dr. Cooper's caseload (Jan.'90-December '90) -RC.

b/Sex distribution:

	male	female
OJ	93	148
RC	56	123
	35.5%	64.5%

Table 1.1

c/Age distribution:

Range: 61 - 98 years.
Mean: 81.9 years.

There were no significant age differences between the two consultants' caseloads.

d/Admitted from:

	home	ward	hosp.	total
OJ	66	132	8	206*
RC	75	95	8	178*

*Not all patients have this recorded.

e/Patient type:
- The patients were classified as:
  - Terminal (17 patients)
  - Investigation (17 patients)
  - Rehabilitation (382 patients)

These categories of patients have different expected outcome criteria and it was decided to analyse the rehabilitation group results only.

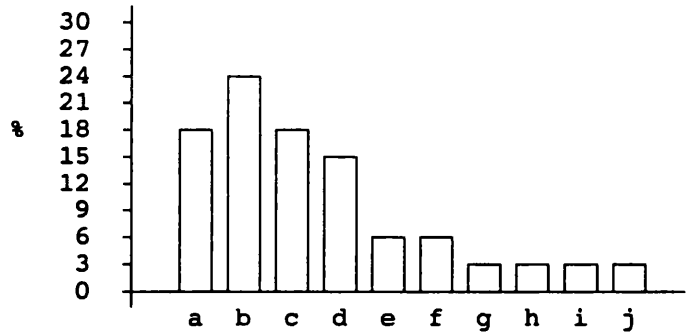
f/Mental Test Scores:

The shortened MTS has been done at admission on the majority of patients (62%):

	Severe 0-3	Moderate 4-6	Normal 7-10
OJ	18	40	95
RC	16	31	61
%Total	8%	17%	37%

Table 1.3

g/Length of stay:



a = < 1 week
b = 1 - 2 weeks
c = 2 - 3 weeks
d = 3 - 4 weeks
e = 4 - 5 weeks
f = 5 - 6 weeks
g = 6 - 7 weeks
h = 7 - 8 weeks
i = 71-100 days
j = >100 days

Figure 1.1

- 76% of patients stayed < 28 days.
- 5.5% of patients stayed > 56 days.
- 5 patients stayed > 100 days (2 OJ; 3 RC).

h/Presenting medical characteristics:

There is a high incidence of co-morbidities within this group. The presenting medical problems were classified as either active or inactive and these are presented in table-1.4

No. pats.=420 (382 rehab., 17 terminal, 17 invest.)

Medical problem	Rehab.- active	Terminal - active	Invest. - active	Inactive - total
Neurological	177	10	1	47
Psychiatric	151	5	5	16
Cardiovascular	131	5	3	87
Gastro-intest.	121	5	12	48
Musculo-skel.	104	2	2	47
Metabolic	101	5	6	37
Respiratory	95	5	2	26
Genito-urin.	91	8	0	24
Skin	38	0	0	8
Renal	21	2	0	8

- The most common active problems of patients who are admitted for investigations, belong to the gastro-intestinal tract e.g. bleeding in the GIT, causing anaemia.
- The most common active problems of patients who are admitted for terminal care are neurological in origin e.g.CVAs. Genito-urinary problems are also recorded as prevalent in this group, urinary incontinence accounting for the majority
- For the study group of patients (i.e. those admitted for rehabilitation), the range of active medical problems is listed in order of prevalence in table-1.4.

i/Home environment on admission:

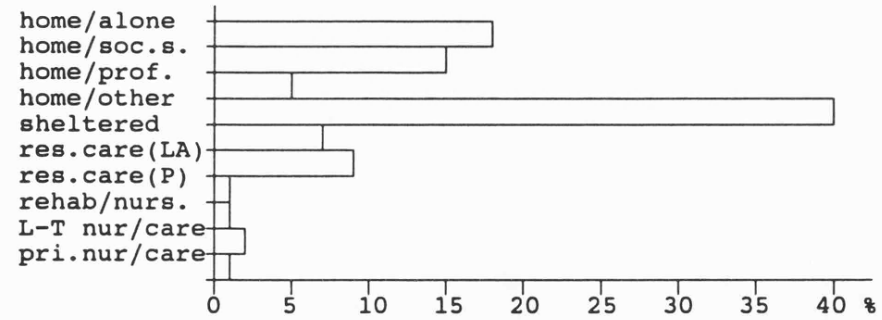


Figure 1.2

- 78% of patients are living in the community, either alone, or with support in their homes, on admission.

j/Care-givers' strain:

The care-givers' strain scale is scored from 0 to 13, with 0 being equivalent to no stress reported and 13 being equivalent to the highest level of stress reported. Despite problems with response (only about half of those sent a form have replied) a picture of which aspects of care are

found most stressful to carers is emerging from the data.

- 38% of the carers reported a low level of stress(<5).
- 42% reported a moderate level of stress(5-9).
- 20% reported a high level of stress(10-13).

The aspects of caring that are most frequently reported as stressful by the carers are:

	%tage respond +ve
- upsetting behaviour	71%
- physical strain	60%
- restrictions on self	57%
- upsetting as changed	56%
- feeling of overwhelm	51%
- sleep disturbance	47%
- emotional adjustments	46%
- changes in personal plans	45%
- inconvenience	44%
- family disruptions	40%
- demands on time	40%

Work adjustments and financial worries are reported as stressful in only about 25% of carers (this could be largely due to the fact that the carers tend to be older and largely female).

2/OUTCOMES:

Outcomes to 3/12 for rehabilitation group (for those who were discharged before December 1990 & therefore followed up to 3/12): - Table 2.1

	1	2	4	6	0	Tot
OJ	37	26	84	28	41	216
RC	32	19	70	18	27	166
Tot	69	45	112	46	68	382

Outcome codes: 1 = died in hosp.
2 = died to 3/12
4 = OK to 3/12
6 = readmitted to 3/12
0 = no follow-up

The outcomes can be summarised:

- 30% of patients die altogether (no difference between consultants).
- 19% of RC's patients die in hospital.
- 17% of OJ's patients die in hospital.
- a further 12% of OJ & 11.4% of RC die before 3/12 follow-up.
- 13% of OJ & 11% of RC patients are readmitted to Ward 14 by 3/12 follow-up (17% & 16% of those surviving to 3/12 follow-up respectively).
- 81% OJ patients are followed-up to 3/12.
- 84% RC patients are followed-up to 3/12.
- of those reaching 12/12 follow-up,46% of OJ patients & 38% of RC patients die. 31% of patients are readmitted to 12/12(However, the numbers are very small to date).

a/Readmissions:

The number of readmissions recorded refers to those patients who are readmitted to Ward 14 and not other wards, or hospitals. Of 46 readmissions to 3/12, 5 patients died during the hospital stay and 5 patients died in

the subsequent 3/12. The difference in functional capacity (Barthel score) between the first discharge and the readmission tends to be large and negative eg. for those who die in hospital on a readmission, the difference in Barthel scores is at least -7.

Some individuals are readmitted more than once - these include 2 patients who have been readmitted three times and 8 patients who have been readmitted twice. The presenting medical characteristics at subsequent readmissions tend to be related to those at the original admission. Five of these in-patient episodes are due to patients being referred to another specialty for treatment and then being re-transferred back to Ward 14 afterwards.

b/Barthel Scores:

The functional ability of the patients has been measured at admission(B1), at discharge(B3), at 3/12 follow-up(B21) and at 12/12 follow-up(B31). The expected outcome at discharge (ie. the goal) has been set at admission(B2).

The mean scores for the group of rehabilitation patients are seen in table 2.2:

CONS.	BARTHEL1	BARTHEL2	BARTHEL3	B3-B1	BARTHEL21
OJ	13	16	15	2	14
RC	10	14	14	4	12
BOTH	10	14	12	2	13

Table 2.2

The distributions of scores at both admission and discharge show a high proportion of scores over 17 - indicating relatively less disability. A ceiling effect occurs with these patients whereby their potential for improving scores is considerably reduced if the initial score is high.

		Better	Same	Worse
From admission to discharge: Improvements are significant (p<.001 using Wilcoxon)	OJ	60%	34%	6%
	RC	70%	21%	8%
Discharge vs. goal set:		**Better**	**Same**	**Worse**
	OJ	17%	56%	27%
	RC	19%	44%	36%
From discharge to 3/12:		**Better**	**Same**	**Worse**
	OJ	28%	21%	51%
	RC	27%	15%	58%
From admission to 3/12:		**Better**	**Same**	**Worse**
The follow-up scores are sig. better than at admission. p=.0061.	OJ	46%	11%	43%
	RC	66%	3%	31%

- The admitting pattern of the consultants appears to be different - RC tends to admit patients with a lower Barthel who subsequently show a better improvement than those patients admitted by OJ. The difference in the change in Barthel scores is statistically significant (p=.036).

- When the individual items of the Barthel are examined between admission and discharge, only 15-20% of patients are worse at

discharge. The majority are the same, with a varying number who are better at discharge. The best improvements are seen in mobility, dressing ability, transfer ability and grooming - the improvements being highly significant.

- However, as can be seen above, the overall Barthel scores are largely better at discharge than at admission - these improvements are statistically significant

- A larger proportion of OJ patients achieved their set goals than did RC patients.

- The best performances against goals are seen for items covering control of bowels and bladder and for bathing.

- The majority of patients' functional ability declined at 3/12 follow-up, compared to discharge. 15-20% maintain their discharge level.However there are a handful of quite disabled patients on discharge who have subsequently improved to 3/12 - the majority are in some other care facility.

- When admission Barthel scores are compared to 3/12 scores, there is a significant improvement to 3/12 & a marked difference between the consultants. 66% of RC patients who reach 3/12 follow-up improved their functional ability from admission, compared to 46% of OJ patients.

12/12 follow-up:

Only 11 of the rehabilitation group of patients have had a follow-up to 12/12. Of these:

	Better	Same	Worse
From 3/12 to 12/12:	5	2	4
From admission to 12/12:	8	0	3

Therefore overall, there appears to be an improvement in functional ability over time. However, this is as yet a very small sample.

c/Care-givers strain scale:

- Only 27 of the carers who successfully completed a strain scale during their dependants' admission have been followed up to 3/12. Of these, 10 report less stress than admission, 7 report the same degree of stress and 9 report a higher level of stress.

- The most frequently reported aspects of caring that are reported as stressful are similar to those at initial contact.

- Of those who report a lower level of stress at follow-up than at initial contact, two carers show a reduction of at least 7 points. Four out of the nine who report increased stress show an increase of 5 or more points.

- Only one carer has been followed up to 12 months. This is a young lady who cares for her grandmother and has reported a low level of stress (1) on each occasion.

d/Home environment:

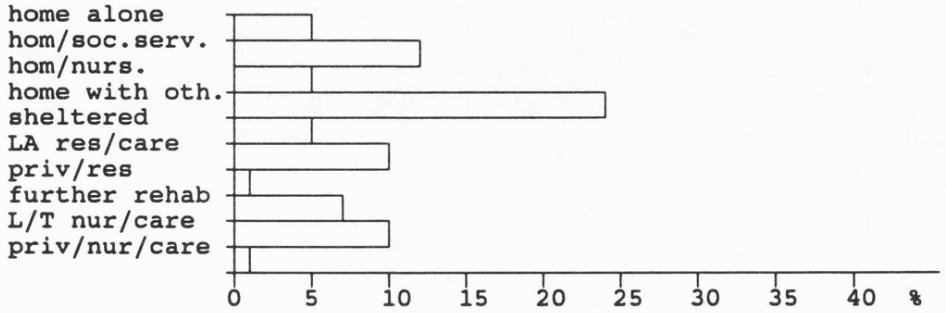


Figure 2.1

The proportion of the sample living independently (either supported or not) in the community has decreased from the time of presentation to hospital. There is a larger proportion of patients in hospital, either for further rehabilitation, or for long-term care. The proportion of patients in residential care is about the same.

12/12 follow-up:

Of the 11 patients that have been followed up to 12/12:

- 2 patients are at home alone
- 2 patients are at home with social services support
- 3 patients are at home with another person
- 2 patients are in res.care
- 2 patients are in private res.care

- 8 of these patients are in the same home environment as at admission 12/12 previously.
- None of them improve, that is , become more independent.
- However, there is little or no deterioration in independence over the 12/12 period for this small number of patients.

e/General ratings:

At discharge clinical staff were asked to assess whether goals had been achieved for each patient and whether medical problems had been resolved. They were also asked to assess whether or not the patient was generally better. For the majority of cases (around 80%) the answers to these questions were positive. Although it was felt originally that this was infact due to over confidence in the outcomes a more detailed analysis of the assessments, in conjunction with the accommodation on admission and discharge (as an indicator of degree of dependence) has revealed that these improvements are perhaps genuine -see table 2.3:

	Accommodation same		**Accommodation worse**
General: better	82%	p=.001	19%
same	63%		38%
worse	20%		80%
Medical: better	80%	p=.004	20%
same	66%		34%
worse	29%		71%
Goals: better	81%	p=.000	20%
same	74%		26%
worse	36%		65%

All relationships are highly significant (Chi-squared p=.0000)

Table 2.3

3/DISCUSSION:

There are strong and statistically significant links between the admitting Barthel scores and the accommodation status of patients. This is as one would expect - patients who have come from other institutions score on the whole lower than those who have been living at home.

These relationships are consistent in the discharge Barthel scores:

Accommodation	- on admission			- on discharge		
	admit	discharge	change	admit	discharge	change
Home	12.2	15.1	3.0	13.1	16.4	3.4
Residential care	8.0	10.1	2.6	9.9	13.6	3.7
Rehabilitation	7.5	9.0	1.5	8.7	10.5	1.9
LT care	6.2	8.4	2.2	8.5	7.2	-.61

Barthel scores against home environment. Table 3.1
All relationships are highly significant.

- There is a weak link between overall admitting Barthel scores and mental ability at admission. The patients with lower MTS show a lower functional ability than those with higher MTS. This relationship is to be expected, as patients with confusion will most probably be more more dependent.

- The functional abilities at admission, of those patients who die in hospital are lower than for the survivors - this relationship is highly significant (p=.0000).

- Patients under 70 years tend to do better in terms of survival, changes in functional status and changes in home environment compared to patients over 70 years. This may reflect a proportion of patients where the acute illness is more significant than chronic, longer term problems.

- The length of stay is over 4 weeks in 24% of cases and 3% patients stay over 10 weeks. Of those remaining in hospital over 4 weeks, 23% are discharged to residential care, who have been in the community previously. Is this an indication of a delay in arranging accommodation at discharge, or are these patients just slow to respond to rehabilitation? A further 47% of those whose LOS was over 4 weeks and who were previously living in the community, have been discharged to another hospital, either for further rehabilitation or for long-term care.

- Does this indicate a set of patients who are slow to respond, or a delay in arranging the transfers?

- Of those patients who die in hospital, but are not classified as terminal", 64% are transferred to Ward 14 from other wards within the hospital. Also, 30% of those who die, do so within 2 weeks of admission to the ward. Does this fit the expected pattern of admission to the ward?

- When the goals set for functional improvement are compared to the actual level attained at discharge (B2 and B3) a large proportion of patients do not attain their set goals (27% for OJ & 36% for RC patients). Why is this? Are the goals being set realistically? Or is there a problem with the rehabilitation of this group of patients? Would it be useful to know which patients have not achieved their set goals?

General comments:

- The measurement of outcomes in this type of Elderly Care setting have proved to be largely dependent on survival, readmissions and functional ability (as measured with the Barthel index). Although the numbers over this period are fairly large (420) there have been difficulties with collection of data at follow-up. About 20% of these patients have not been followed up to 3 months - the difficulty arising when they do not attend for an Out-patient appointment at around this period and they are not living in some kind of care facility where they are easily accessible.

- There are only a small number of this group of patients who are due to have a 12 month follow-up, but as numbers grow, the problem of successfully locating the patients will become more of a problem. It may necessitate a change in the strategy of data collection, with attempts being made to check survival, readmissions and home environment only, as opposed to collecting functional assessments on a routine basis.

Julie Goodfellow
Martin Bardsley

Feb 1991.

Appendix 5(a) Data collection forms - Cholecystectomy

Admission Date [/ /]

Discharge Date [/ /]
Op Date [/ /]

Freeman Hospital Cholecystectomy Study
PRE-OPERATIVE ASSESSMENT
(Obtained from the patient notes)

Consultant: CV PW
NJ JC

Hospital No.: [.....]

NAME: .....
ADDRESS: .....
.....
.....
.....

The indication for surgery is:

History of acute cholecystitis	[ ]	Chronic cholecystitis	[ ]
Acute cholecystitis on admission	[ ]	Common bile duct stone	[ ]
		Obstructive jaundice	[ ]
Asymptomatic gallstone	[ ]	Mucocoele	[ ]
Empyema	[ ]	Cholangitis	[ ]

Other, please specify .....

Significant current/chronic disease:

		Please Specify	Taking Medication
Cardiovascular	[ ]	.....	[ ]
Nervous system	[ ]	.....	[ ]
Respiratory	[ ]	.....	[ ]
Musculoskeletal	[ ]	.....	[ ]
Gastrointestinal	[ ]	.....	[ ]
Metabolic	[ ]	.....	[ ]
Renal/genito-urinary	[ ]	.....	[ ]
Other problems		.....	[ ]
		.....	

Other Details

Weight : [ ] Height : [ ]
Smoker : [ ] Non-smoker: [ ] Ex-smoker : [ ]
Heavy drinker: (Men > 21 units; Women > 14 units) [ ]

Previous abdominal surgery : [ ]

QUESTIONNAIRE 2 Date [/ /]

Freeman Hospital Cholecystectomy Study

Surname : [.....]

First name : [.....]

Hospital No: [.....]

Patient Pre-operative Self-administered Questionnaire

We would like to ask you some questions about your health and in particular about the problems related to your gall bladder.
Please tick the appropriate box.

	Yes	No
Have you had any abdominal pain?	[ ]	[ ]

If yes.....

How many times in the last year have you had abdominal pain?

Once	Twice	Three times	Four times or more
[ ]	[ ]	[ ]	[ ]

	MONTH	YEAR
When did your gall bladder symptoms first appear? Please give an approximate date.		
When were you first diagnosed (by your G.P. or hospital surgeon) as having gall bladder disease?		

Please tick 'Yes' if the following statements apply to you, otherwise tick 'No'.

	Yes	No
I take tablets/drugs for my abdominal pain	[ ]	[ ]
I suffer from flatulence a lot (belching/wind)	[ ]	[ ]
I suffer from abdominal distention (bloating)	[ ]	[ ]
I feel sick and/or vomit	[ ]	[ ]
My appetite is normal?	[ ]	[ ]
Eating fatty food bothers me	[ ]	[ ]
My bowel function is normal	[ ]	[ ]
I have constipation	[ ]	[ ]
I have diarrhoea	[ ]	[ ]

Overall how would you rate your health at present?

Poor [ ] Fair [ ] Good [ ] Very Good [ ]

QUESTIONNAIRE 3.

PRE-OPERATIVE FORM

PATIENT IDENTIFICATION: DATE:[/ /]

Name.....

Address.....
.....

Hospital No:.....

Operating Surgeon or Consultant [.....]

1. Indications for surgery:

History of acute cholecystitis	[ ]	Chronic cholecystitis	[ ]
Acute cholecystitis on admission	[ ]	Common bile duct stone	[ ]
		Obstructive jaundice	[ ]
Asymptomatic gallstone	[ ]	Mucocoele	[ ]
Empyema	[ ]	Cholangitis	[ ]

Other, please specify.....

2. How do you rate the risk to the patient of significant post-operative morbidity?

Low Risk ├──┤ High Risk

3. Are there complicating factors in this patient?

4. Operation type:

Elective	[ ]
Urgent	[ ]
Emergency	[ ]

Date of operation: [.....]

Were there any significant *complications during the operation?
If yes were they:

Surgical
Please specify
.....

Anaesthetic
Please specify
.....

* Significant = a condition that.....
- prolongs the length of stay
- requires extra clinical resources
eg. nursing diagnostic tests, extra time in theatre.

Completed by..... Date: [.....]

QUESTIONNAIRE 4
PATIENT IDENTIFICATION:

Name..... Date:[/ /]

Address.....
.....

Hospital No.....

Post-operative complications

Were there any significant post-operative complications?

	YES	NO
- Wound infection.....	[ ]	[ ]
- Intra-abdominal infection.....	[ ]	[ ]
- Post-op bleeding.....	[ ]	[ ]
- D.V.T./Pulmonary embolis.....	[ ]	[ ]
- Respiratory infection/complication.....	[ ]	[ ]
- Cardiovascular complication (eg. M.I.).....	[ ]	[ ]
- CNS complication (eg. stroke).....	[ ]	[ ]
- Urinary complication (eg. retention/infection).	[ ]	[ ]
- Septicaemia.....	[ ]	[ ]
- Renal failure.....	[ ]	[ ]

- Other, please specify

.....

.....

Significant = a condition that...

- prolongs the length of stay
- requires extra clinical resources
  eg. nursing, diagnostic tests, extra time in theatre.

Completed by..... Date: [.....]

QUESTIONNAIRE 5 Post-operative patient questionnaire

Three month/one year follow up:

Hospital No: [.....]

Some people continue to have problems for several months after their operation. Some people have no problems at all. We would like you to say whether you agree with the following statements.

Please tick 'Yes' if the following statements apply to you, otherwise tick 'No'.

SINCE THE OPERATION.....	Yes	No
I take tablets/drugs for my abdominal pain	[ ]	[ ]
I suffer from flatulence a lot (belching/wind)	[ ]	[ ]
I suffer from abdominal distention (bloating)	[ ]	[ ]
I feel sick and/or vomit	[ ]	[ ]
My appetite is normal	[ ]	[ ]
Eating fatty food bothers me	[ ]	[ ]
My bowel function is normal	[ ]	[ ]
I have constipation	[ ]	[ ]
I have diarrhoea	[ ]	[ ]
I have visited the GP with a complication related to the operation	[ ]	[ ]
I have health problems I did not have before	[ ]	[ ]
I feel no better than I did before	[ ]	[ ]

If so please specify

.....

Overall, how would you rate your health at present:

Poor [ ] Fair [ ] Good [ ] Very good [ ]

Please tick the correct box	Yes	No
Have you been seen at Accident & Emergency (A & E) since your operation?	[ ]	[ ]
Have you been re-admitted to hospital since your operation?	[ ]	[ ]

If yes which hospital?

Freeman	[ ]
Royal Victoria Infirmary (RVI)	[ ]
Newcastle General (Westgate Road)	[ ]
Other	[ ]

Please specify

.....

If yes, what were you admitted for?

.....
.....

Repeat Nottingham Health Profile

Thank you for your help. Please return these forms in the stamped addressed envelope provided.

Appendix 5(b) Caspe/Freeman Outcome Study Report Aug 1990

Cholecystectomy Outcomes by Quarter.

Q1=J-F-M 1989 Q3=J-A-S 1989 Q5=J-F-M 1990
Q2=A-M-J 1989 Q4=O-N-D 1989 Q6=A-M-J 1990

1. Patient characteristics and events in hospital.

		Follow-ups			Age distribution			
Q.	Tot.	3mnth	12mnth	%Male	Mean	<60	60-70	>70
1	27	85%	81	41%	65.5yrs	28%	28%	44%
2	22	100%	77	27%	57.9yrs	53%	32%	16%
3	27	100%	0	26%	59.8yrs	42%	31%	27%
4	16	94%	0	38%	56.7yrs	44%	44%	13%
5	38	89%	0	34%	59.0yrs	53%	17%	31%
6	14	43%	0	21%	56.3yrs	40%	40%	20%
Total	144	88%	27	32%	59.8yrs	44%	29%	27%

	Indications for surgery:					
	%Emergency	%Asymp.	%Chronic	%History	%Acute	%Other
1 JFM	19%	19%	11%	52%	4%	15%
2 AMJ	0%	14%	27%	41%	0%	18%
3 JAS	22%	19%	19%	33%	11%	19%
4 OND	13%	6%	31%	50%	6%	6%
5 JFM	26%	21%	26%	18%	18%	16%
6 AMJ	0%	29%	0%	36%	0%	36%
Total	16%	18%	20%	36%	8%	17%

		Mean length of stay					
	Nos.	Preop	Post	All	<1wk	1-2wks	>2wks
1 JFM	25	2.3	6.5	9.8	20%	52%	28%
2 AMJ	18	1.8	4.8	10.5	6%	83%	11%
3 JAS	22	1.5	6.9	9.0	18%	73%	9%
4 OND	16	2.4	6.5	9.6	6%	88%	6%
5 JFM	24	1.7	5.9	9.7	21%	71%	8%
6 AMJ	6	2.2	3.5	8.0	17%	83%	0%
Total	111	2.0	6.0	9.6	15%	72%	13%

2. Post-op complications and events after discharge

		Copmplications		To 3 months			To 12 months		
		Peri-op.	Postop	Readmit	AE	GP	Readmit	AE	GP
1	JFM	19%	26%	9%	9%	13%	18%	18%	18%
2	AMJ	27%	32%	5%	9%	32%	12%	24%	47%
3	JAS	7%	19%	7%	0%	19%	-	-	-
4	OND	13%	13%	7%	13%	33%	-	-	-
5	JFM	24%	26%	6%	12%	24%	-	-	-
6	AMJ	7%	0%	0%	0%	0%	-	-	-
Total		17%	22%	6%	8%	22%	15%	21%	31%

Post-op complications:

Urine retention	9 cases
Resp. Infection	2 cases
Cardiovasc. comp	2 cases
Abdom complic	1 case
Bleeding	1 case
DVT	1 case

Readmissions:
After three months: *(patient ID's removed)*
- -"Replace blocked arteries stomach"
- -"Infection in wound"
- -"Pain and sore around wound"
- -"Scar tissue turned septic"
- -"Pain and vomiting"
- -"Pain and coughing up blood"
- -"Heart attack"
- -"Road traffic accident"
- -"To convalescence"

From 3 to 12 months:
- -"Cyst in throat"
- -"Broken collar bone"
- -"Heart attack"

Patients died:
==========
4 patients have died (up to 12 months).
a -
b -
c -
d -

We know at least one other patient has died but not yet reached their 12 mnth f/u point.

3. Changes in Symptoms to three months

Before: % reported problems

SYMPTOMS	Tablet	Flat.	Dist Ab	Vomit	Bowel	Appetit	F.Food
1 J-F-M	26%	59%	52%	44%	22%	33%	63%
2 A-M-J	41%	59%	59%	68%	18%	32%	68%
3 J-A-S	48%	52%	52%	70%	37%	37%	74%
4 O-N-D	44%	81%	44%	56%	44%	31%	75%
5 J-F-M	37%	61%	53%	53%	45%	26%	71%
6 A-M-J	29%	43%	36%	57%	14%	21%	71%
Total	**38%**	**59%**	**51%**	**58%**	**32%**	**31%**	**70%**

After - % reported problems

SYMPTOMS	Tablet	Flat.	Dist A	Vomit	Bowel	Appetit	F.Food
1 J-F-M	7%	30%	30%	11%	30%	33%	22%
2 A-M-J	18%	36%	23%	5%	45%	23%	36%
3 J-A-S	26%	30%	30%	19%	26%	19%	30%
4 O-N-D	25%	56%	38%	38%	31%	31%	56%
5 J-F-M	11%	13%	18%	8%	37%	32%	16%
6 A-M-J	0%	14%	7%	0%	64%	64%	0%
Total	**15%**	**28%**	**24%**	**13%**	**37%**	**31%**	**26%**

*Basic pattern is as before the greatest changes are in vomiting, tolerance of fatty foods, distended abdomen and flatulence. Little improvements is seen in 'bowel' problems.

*Additional questions on the nature of bowel problems were added to the forms.

4. Changes in the NHP

Mean scores before:

NHP		Energy	Pain	Sleep	Emot	SocI	Mobil	PtII
1	J-F-M	27	36.9	24.2	17.0	27.9	8.8	15.6	21.2
2	A-M-J	22	40.4	29.1	16.6	27.8	5.8	11.7	28.6
3	J-A-S	27	42.0	28.2	17.6	27.6	13.4	17.4	25.9
4	O-N-D	16	40.7	23.9	14.6	27.1	5.5	9.0	27.7
5	J-F-M	36	31.0	34.8	20.3	31.9	6.4	18.2	28.6
6	A-M-J	12	42.4	23.4	19.3	34.3	10.9	20.3	28.6
Total		**140**	**37.8**	**28.3**	**17.8**	**29.3**	**8.4**	**15.7**	**26.5**

Mean Scores after:

NHP +3mnths		Energy	Pain	Sleep	Emot	SocI	Mobil	PtII
1	J-F-M	27	22.5	11.3	7.5	13.9	5.7	13.4	8.5
2	A-M-J	22	24.9	16.3	6.0	16.8	2.9	14.0	14.3
3	J-A-S	27	20.1	6.8	10.3	22.6	6.0	12.9	14.3
4	O-N-D	16	29.8	15.4	9.2	25.7	15.5	10.5	16.1
5	J-F-M	35	11.0	4.5	2.3	6.2	3.1	7.6	7.8
6	A-M-J	5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total		**132**	**19.4**	**9.5**	**6.4**	**15.0**	**5.6**	**11.0**	**11.0**

% patients showing improvements in NHP:

NHP		Energy	Pain	Sleep	Emot	SocI	Mobil	PtII
1	J-F-M	44%	63%	59%	67%	19%	30%	48%
2	A-M-J	50%	59%	55%	55%	9%	32%	59%
3	J-A-S	56%	63%	41%	41%	33%	30%	41%
4	O-N-D	31%	44%	44%	38%	6%	13%	38%
5	J-F-M	43%	71%	71%	71%	17%	43%	54%
6	A-M-J	80%	40%	80%	20%	40%	40%	80%
Total		**47%**	**61%**	**57%**	**55%**	**19%**	**32%**	**50%**

5. Changes to 12 months

The general picture of patients 12 months after the operation is similar to that at three months - NHP scores stay much the same from 3 to 12 months

Mean NHP scores:

	Pre-op	**+3mnths**	**+12mnths**
Energy	38.8	24.9	23.5
Pain	26.2	14.0	10.0
Emot Reac	14.9	6.8	6.9
Sleep	25.6	16.3	20.0
Soc. Isol	5.6	4.5	6.3
Mobil	14.4	14.3	15.0
Pt II	21.6	11.0	11.4

* Symptom scores are roughly similar (possibly slightly worse).

Symptoms - No. patients reporting problems:

	Pre-op	**+3mnths**	**+12mnths**
Tablets	11	4	6
Flatulence	24	14	12
Dist Abdomen	22	10	3
Vomit	23	4	3
Bowel	7	15	11
Appet	14	11	11
F.Food	27	12	17

At the moment the results would suggest that the 12 months review does not add much to the picture.

Examples of exception reports ..

The follwing summaries refer to a selection of patients who apppear not to have done well after the operation - either as reflected in their NHP or readmissions. It is interesting to oberserve that in most cases the NHP, the list of symptoms and the facts of readmission tend to agree.

Do these patients hold any useful lessons?

Patient Summary : A Patient id11111111 (4) dd/mm/yy
Indications : Chronic cholecystitis
Co-morbidities :
Previous ops. : Hysterectomy, tonsillectomy
Abdominal pain : >4 times last year
Symptoms since : 6/87 Diagnosed 1/89
Problems : tablets,abdo.dist.,appet.,fatty food
3mth. later : flat.,abdo.dist.,bowels
12 mth. later : tablets,abdo.dist.,fatty foods
Op. complications :
Post-op. complications:

NHP scores	Pre-op	+3mth.	+12mth.
1. Energy	100.00	60.80	63.20
2. Pain	72.96	30.17	35.27
3. Emotional reaction	19.78	10.47	0.00
4. Sleep	34.27	0.00	12.57
5. Social isolation	0.00	0.00	0.00
6. Physical mobility	33.19	31.07	41.86
7. PartII responses	42.87	70.00	42.87
Readmitted?		no	no
GP?		yes	yes
A+E?		no	no
Feel no better?		no	no
Other problems?		no	no

Comments: 3/12 - infection in wound.
12/12 - no reported problems.

Patient Summary : A patient id222222 dd/mm/yy (2)
Indications : Chronic cholecystitis
Co-morbidities : Angio Dec.88 re chest pains
Previous ops. : Tubal tie, eye op.
Abdominal pain : >4 times last year
Symptoms since : 1/87 Diagnosed: 10/88
Problems: tabs,flat.,abdo.dist.,vom.,bowels,appet.,fatty 3mth. 3 mnth later: no change
Op. complications :
Post-op. complications:

NHP scores	pre-op	+3mth.	+12mth.
1. Energy	100.00	63.20	-----
2 . Pain	73.31	100.00	-----
3. Emotional reaction	69.84	55.85	-----
4. Sleep	77.63	77.63	-----
5. Social isolation	64.67	42.14	-----
6. Physical mobility	31.07	31.07	-----
7. PartII responses	42.87	56.00	-----
Readmitted?		yes	
GP?		yes	
A+E?		yes	
Feel no better?		no	
Other problems?		yes	

Comments: Pain - constant(r.side) and coughing up blood. Still getting chest pain.

APPENDIX 5 (c) Notes on Links to Case Mix - DRGs August 1990

The patients on the cholecystectomy outcomes study data base have been linked in to details from the hospital case mix system. The data covered all episodes from January 1989 to May 1990. The following notes cover some observations/issues resulting.

1.In some cases our patients were registered with an experimental number rather than a hospital number so linkage was not possible. In some cases a link was made - but the inpatient episode was the wrong one (typically wrong DRG and too short). This may be due to my software, the download or the absence of the right(subsequent) consultant episode on case mix. I will investigate further.

2.There appeared to be examples where the coding was not complete or not accurate. In my opinion the DRG is a good way of quickly identifying any gross anomalies. The most common problems are either when the coding has not been done, or when a procedure code appears to be missing (probably the DRGs 207,208).

3.The key episodes (ie the stay during which the cholecystectomy was carried out) mainly falls into four DRGs

197 "Total cholecystectomy w/o common duct explore w/o cc"
198 "Total cholecystectomy w/o C.D.E. with cc"

193 "Biliary tract procedure excluding total cholecystectomy with cc"
194 "Biliary tract procedure excluding total cholecystectomy without cc"

These two pairs are distinguished by the procedure codes, J18.3 (Total cholecystectomy only) leading to DRGs 197 and 198, J18.2 (chole + CBD explore) leading to 193 and 194.

The complications/co-morbidities are a fairly mixed and include codes for.. angina, MS, pneumoconiosis, cholangitis etc.

A few patients had other procedures which may lead them to other DRGS - which is as it should be.

4.There are some notable patterns in the way many patients had an earlier admission (in another DRG). Similarly some patients show readmissions though the DRGs tend to be varied. This idea of a bundle of in-patient episodes - even for elective surgery - seems to have implications for contracting etc. If you based your charges on just the chole episode you could lose if you are also doing the diagnostic work-ups on 20% of the patients.

The following figures are taken from 96 cases admitted for cholecystectomy from March 1989 to March 1990

Main Episode

DRG 197	9 cases	Total cholecystect. w/o CDE with cc
DRG 198	44 cases	Total cholecystect. w/o CDE w/o cc
DRG 193	10 cases	Biliary tract proc. exc tot chole +cc
DRG 194	9 cases	Biliary tract proc. exc tot chole w/o cc

No Link: 10 cases
Wrong Link: 5 cases

DRG 207	3 cases	Disorders biliary tract with cc
DRG 208	2 cases	Disorder biliary tract w/o cc
DRG 148	1 case	Maj small & large bowel proc
Blank	1	Not yet coded
DRG 200	1 case	Hepatobil dx proc for non-malig
DRG 192	1 case	Minor pancreas,liver,shunt proc

Typical episodes before cholecystectomy:

DRG 183	9 cases	Mainly endoscopic examination
DRG 207/208	10 cases	Biliary tract dx but no procedure

Others include:

204 Dis. pancreas	416 Septicaemia
143 Chest pain	88 Chron. Obstr. P. Dis
125 Circ. disorders ex AMI	165 Appendicectomy
206 Disease liver	87 Pulm oedema
449 Peptic ulcer	449 Poisening

Typical episodes after cholecystectomy:

394 Other OR proc blood forming organs
111 Major reconstructive vasc. proc
183 Oesaphagitis etc.
100 Other digestive system diagnosis
109 Cardiothoracic procedire

Appendix 5(d) Minutes of CASPE/General Surgery Meeting 20/11/90.

Present:

Martin Bardsley.
Julie Goodfellow.
Alison McCallum.
Joan Watson.
Gill Sanders.
Mr. P. Wright.
Mr. C. Venables.
Mr. N. Jones.
Mr. J. Chamberlain.

1. **Additional Analyses**

MB presented some additional analyses he had done on the results presented in the main report. These included:

a) **Risk Assessment**

The numbers of cases with a pre-operative'risk' assessment was often not completed on the pre-operative form.
Those that have been completed appear to show a strong correlation to the incidence of perioperative complications but not post-op. complications. MB suggested that this was probably because the forms are often completed after the operation. PW & NJ protested that this was not correct and confirmed that they are completed pre-operatively.
CV expected to see a correlation between assessment of risk and incidence of post-op. complications, which as yet has not appeared- this is probably due to the small numbers involved.
MB discussed the apparent relationships between the assessment of risk and in pre and post-operative NHP scores. Those patients assessed as "high risk" by the surgeons have high pre-op NHP scores and show good improvements, whilst the opposite is true for those assessed as "low risk". PW suggested that those patients who were assessed as "medium risk" and who did not show big improvements over this time period, are perhaps a separate category of patients that may have bowel problems as opposed to gall bladder disease.

b) **Bowel Problems**

MB discussed the incidence of bowel problems. It had been noted in previous meetings that reported bowel problems were often worse after the operation than before. The results (which appear to be internally consistent) showed that the incidence of diarrhoea & constipation are roughly equal, at follow-up.

CV pointed out that one might expect a higher incidence of diarrhoea, due to the continuous flow of fluid into the duodenum and hence into the bowel, which occurs after cholecystectomy.

c) **Validations**

MB reported that the validations have been started comparing the NHP for pre-admission, to the NHP completed in hospital. Initial results show very good correlations between the two scores.
JG reported that interviews with 15 patients have now been carried out, in order to validate the questionnaires used to collect data on admission and at follow-up.
There were no further comments on this matter.

General Results

MB discussed the main report with the surgeons. He pointed out that a small number of patients had been missed, mainly due to unexpected, or emergency admissions.
CV and PW discussed the problem of classifying the indications for surgery and reported that they found this difficult. They appeared particularly concerned about the high proportion of patients who have been classified as asymptomatic, when probably they have had at least one episode of biliary colic in the past. It was agreed that JG would investigate this further.
CV suggested that better indication categories would be:

- Acute - cholecystitis
  - biliary colic
  - pancreatitis
- Chronic

He felt that it would be necessary to re-categorise retrospectively at a later stage.

Length of Stay

PW pointed out that 2 days appeared an excessively long pre-op stay and CV stated that he prefers to have his patients in early. However, he feels that the post-op stay is more important and generally, this seems to be decreasing.

Complications

There was a general feeling that the incidence of post-op complications is low. MB was surprised at the apparently high incidence of urine retention, but the surgeons agreed that this is one of the most common post-op complications.
PW pointed out that a lot of the patients have several medical problems and therefore will show a high incidence of problems at follow-up, which may often not be related to the cholecystectomy.
NJ enquired as to whether or not there was any information on retained stones, as the FRH have had a 15 - 20% incidence of this in the past. CV stated that the accepted incidence is about 8% worldwide.

MB replied that he could not recall any such incidence, but that CASPE will recheck the data. The surgeons feel that this is a positive statement on the results.

Results at 12 month follow up

The results at 12 month appear to be similar to those at 3/12 and the improvements seen post-op appear to be holding. The slight deteriorations could perhaps be explained by the occurrence of other illness and the need of some patients to be "intrinsically sick" (PW). MB felt that the indications at present were that, perhaps 12/12 data collection could be stopped. JW stated that problems with non- responders to 12/12 follow up creates a lot of work. GS felt that the collection of 12/12 follow up data should be continued until there are larger numbers. It was agreed to continue the annual follow-up at least until Easter.

Exception report

MB enquired if exception reports on patients who didn't respond were useful in identifying generallessons and areas for improving practice. PW agreed that these patients need to be identified, but they did not as yet contribute any general lessons.

GS asked if reporting on these exceptions would allow the surgeons to learn from them. Although there was agreement that the information is interesting, it was felt that it would not affect practice.

CV stressed the aims of the study as the feasibility of outcomes monitoring, and that using the results to change practice (eg. patient selection criteria) was unlikely. It was pointed out that if there was comparative data available from another unit, which showed better outcomes, then the result could be more useful in changing practice.

Funding

Funding of the CASPE project was discussed. It was decided to carry on with data collection until March 1991, especially as an unusually small number of cholecystectomies have been performed in recent months, due to bed closures.

Other areas were discussed as possible extensions of the outcomes research, should funding become available. Those suggested by the surgeons were:

- colo-rectal surgery, as common.
- day case surgery for hernias, as this is an increasing area.

Julie Goodfellow. 15/1/91.

Appendix 6(a) FREEMAN HOSPITAL DIABETIC CLINIC FORM 1 NEW PATIENT OUTCOMES

PATIENT IDENTIFIERS
HOSPITAL No: .....

Surname: .....

Used name: .....
Sex (M/F) .....

Date of birth: (d/m/y)
..../...../....

Year diagnosed .....

Insulin treated?
1=Yes, 2=No .....

NEW REFERRAL OR INSULIN-STARTER (1/2) .....

Date of visit:.....

Referred by (GP/Cons init).....

New or known diagnosis(1/2).....

FH of diabetes(Y/N).....

Starting insulin(1).....

On insulin now(1).....Year started....

On OHA now (1).....Year started....

OHA in past(1).....Years....to

Weight (kg) .....
Height (m) .....

BMI .....
Urine protein.....

Attended/attending already, or known problem(0/1)

Cardiology .....
Ophthalmology .....
Smoking .....

Drugs-Anti-failure.....
Anti-BP .....
Anti-angina.....
Anti-lipid .....

Nephrology .....
Vascular .....
Symptoms:
Angina .....
Claudication .....
Neuropathic .....
Impotence .....

Lower Limbs	R	L
Poor feet/nails	.....	.....
Ulcer/infection	.....	.....
Amputation	.....	.....
Foot pulses	.....	.....

Eyes	R	L
VA	../..	../..
Cataract	.....	.....
Background	.....	.....
Maculopathy	.....	.....
Advanced	.....	.....

Blood pressure ...../.....

Highest recent glucose ....
HbA1 ....
Creatinine ....

Cholesterol/HDL ...../.... ....
Triglyceride ....
Microalbumin ration ....

Comments/past medical history

FORM 2 DIABETES 3 MONTH FOLLOW UP - NEW PATIENTS

Name: Date:

Weight :.....kg

BP :...../.....

HbA1 :..... Chol:.....

Hypoglaycaemia last month: None[ ]

	AM	PM	EVE	NT
Self-treated	[ ]	[ ]	[ ]	[ ]
Other treated	[ ]	[ ]	[ ]	[ ]

Referrals:

Chirop	[ ]	Cardiol	[ ]
Nephrol	[ ]	Ophthalmol	[ ]
Neurol	[ ]	Other.....	

Seen by.....

FORM 3 NON-INSULIN DEPENDENT DIABETIC PATIENT QUESTIONNAIRE

PATIENT IDENTIFICATION

DATE: / /

PLEASE READ THE FOLLOWING INSTRUCTIONS BEFORE ANSWERING ANY QUESTIONS.

1. On the following pages you will find some questions and statements about diabetes.

2. Each question and statement is followed by a number of choices.

3. You should choose from these choices one or more which you think correctly answers the question or completes the statement.

   FOR EXAMPLE:

   Q Most people normally travel to and from work by .....

   Bus/Train [X]
   Horse [ ]
   Car/Motorcycle/Scooter [X]
   Bicycle [X]
   Aircraft [ ]
   I do not know [ ]

In this example the THREE choices which most people would agree correctly completes the statement are marked thus [X]

4. If you cannot answer a question or complete a statement please make sure to put a cross next to "I do not know".

5. Please attempt to answer all the questions even though some may seem not to apply to you.

6. Make sure you answer the questions on all of the pages.

THANK YOU.

PLEASE TURN OVER

Q1. Which of the following is true about diabetes?

A little sugar in the urine is a good thing	[ ]
It can be controlled with treatment	[ ]
You will have it for the rest of your life	[ ]
There can be too much sugar in the blood	[ ]
Your body does not make enough insulin	[ ]
I do not know	[ ]

Q2. Tablets for the treatment of diabetes.....

May be increased by the doctor if urine tests are positive	[ ]
Help lower the amount of sugar in the blood	[ ]
Need not be taken if a meal has been missed	[ ]
May be stopped by the doctor if urine tests are negative	[ ]
I do not know	[ ]

Q3. Why do diabetics test their urine?

To know if the diabetes is under control	[ ]
To know that the kidneys are working properly	[ ]
To measure the level of insulin in the urine	[ ]
To check for blood in the urine	[ ]
I do not know	[ ]

Q4.If a urine test shows 2% (++++) sugar this most likely means a diabetic has ....

Slightly high blood sugar level	[ ]
A low blood sugar level	[ ]
A high blood sugar level	[ ]
A normal blood sugar level	[ ]
I do not know	[ ]

Q5. The correct way for a diabetic to care for toenails is ....

To cut them every 7 to 8 weeks	[ ]
To cut them straight across	[ ]
To cut them very short	[ ]
To use a magnifying glass if eyesight is poor	[ ]
To clean free edges after cutting	[ ]
I do not know	[ ]

Q6. Minor injuries to the feet are more likely to get infected when the blood sugar levels ....

Occasionally get too low	[ ]
Are low all the time	[ ]
Are high much of the time	[ ]
Occasionally get too high	[ ]
I do not know	[ ]

Q7. Foods containing refined sugar ....

Always cause blood sugar levels to get too low	[ ]
Raise blood sugar levels quicker than starchy foods	[ ]
Will have no effect on blood sugar levels	[ ]
Are slower than starchy food in raising blood sugar levels	[ ]
I do not know	[ ]

Q8. Which of the following contains sugar?

Fruit squash	[ ]
Chocolate mousse	[ ]
Instant pudding	[ ]
Cows' milk	[ ]
Parsnip	[ ]
I do not know	[ ]

Q9. Which of the following is high fat food?

Cottage cheese	[ ]
Skimmed milk	[ ]
Salad cream	[ ]
Cheddar cheese	[ ]
I do not know	[ ]

Q10. High blood sugar levels can be caused by ....

A delayed meal	[ ]
Being less active than usual	[ ]
Drinking alcohol	[ ]
Getting an infection	[ ]
Emotional stress	[ ]
I do not know	[ ]

Q11. Keeping diabetes well controlled over the years can lower the risk of damage to ....

The stomach	[ ]
Nerves in the feet	[ ]
The kidneys	[ ]
The lungs	[ ]
The eyes	[ ]
The hearing	[ ]
I do not know	[ ]

PLEASE TURN OVER

Q12. Common symptoms of low blood sugars are ....

Feeling hungry and sweating	[ ]
Blurred vision	[ ]
Feeling sick and thirsty	[ ]
Slurred speech	[ ]
Passing a lot of urine	[ ]
I do not know	[ ]

Q13. Hardening and narrowing of the arteries is a problem ....

Not seen very often in diabetics	[ ]
Which is only seen in diabetics	[ ]
Which can occur earlier in diabetics	[ ]
No worse for diabetics than non-diabetics	[ ]
I do not know	[ ]

Q14. Some common symptoms of very high blood sugar levels are ....

Feeling faint	[ ]
Blurring of vision	[ ]
Passing more urine than usual	[ ]
A headache and feeling irritable	[ ]
Feeling thirsty	[ ]
Feeling sick and vomiting	[ ]
I do not know	[ ]

Q15. Your urine or blood tests have started to show increased sugar, you ....

Should rest for 4-5 hours	[ ]
Should check your diet is correct	[ ]
Should check for any infections	[ ]
May need to eat less at meal times	[ ]
I do not know	[ ]

Q16. When a diabetic smokes the effect it has is ....

To increase the risk of damage to blood vessels	[ ]
Increased risk of poor blood circulation in legs	[ ]
To increase the risk of heart disease	[ ]
To cause problems with weight control	[ ]
No worse than for a non-diabetic	[ ]
I do not know	[ ]

Appendix 6(b)

CASPE/Freeman Outcome study

Diabetes Outcome Report January 1991- Annual Review Cases

A. Introduction
B. Cross section of clinic popoulation 1987-1990
C. Changes in laboratory indicators 1988-1989 and 1989-1990
D. Changes in clinical indicators 1988-1989 and 1989-1990.
E. Cases not receiving annual review

Appendix I - Definitions of patient types, high,low etc.
Appendix II - Results annual reviews a. Metabolic indicators
b. Clinical indicators

Contact: Martin Bardsley, Donna Swinden,Freeman x26473
Angela Skinner, NGH Diabetic Unit
Martin Bardsley, Jim Coles, CASPE 071-229-8739.

A. Introduction

The following report reviews outcomes/changes in patient health for the Freeman diabetes clinic. The data concerns the 1989 annual review clinic and changes that occurred over the preceding year. In addition data on the first six months of 1990 is presented.

Our aims are to explore:

- the appropriateness of these views on outcome and as descriptions of performance.

- identify whether the results are better/worse than expected.

- explore whether looking at outcomes can improve practice and management and so lead to better quality care.

B. Cross Sectional views of the clinic population

Tables

B1= Cross section 1987. (n=430)
B2= Cross Section 1988. (n=547)
B3= Cross section 1989. (n=637)
B4= Cross section 1990 Jan to July (n=426).

Tables B1 to B4 show for separate years the distribution across a range of key variables for the cases who had an annual review. New cases who have only just been logged onto the database have been excluded. The values represent the percentage of that patient type with particular scores on each variable.

Cases have been split into four patient/treatment types (see Appendix I)

Y=Insulin treated, age at diagnosis < =35
I=Insulin treated, age at diagnosis > 35
T=Patients on tablets
D=Patients on diet only

Briefly the tables suggest:

Sex - remained approximately stable at between 50-60% male, fewer males being in the insulin (dx < 35) group and with slightly more men on diet only.

Age - remained stable with marked differences between treatment groups.
The proportion of elderly cases in the insulin (Dx > 35) appears to be rising.

HbA1 -There are some fluctuations between years -1988 being the 'best'. The results for 1989 and 1990 are disappointing. The increase in the insulin (dx < 35) group from 1989 to 1990 seems particularly alarming.

Cholesterol - Generally stable with an increase in the incidence of non-recording in 1989. By 1990 27% of cases had values over 6.5

BMI - gradually worsening from 32% to 44% greater then 27.5 kg/m2. The changes appear in all patient groups.

BP - Until 1990 the results looked stable. The last six months have seen a rise in patients high on both systolic and diastolic.

Visual acuity - stable at around 10% of cases with some visual impairment

Circulation - stable, about 30% having some problem, 15% having claudication or worse.

Eyes - stable, 25% with abnormal fundi, 5-8% maculopathy, 5-7% proliferative retinopathy.

Angina - Slight increase from 15% to 18% over the years.

Neuropathic symptoms - shows some fluctuations 17-24% but no clear trend

Biothesiometer - erratic (and with patchy recording). There have been some differences in the calibration of the machines and the validity of comparisons over these time scales

must be doubtful.

Smoking - stable at around 15-20% smokers

Albustix - Erratic ranging from 5-13%, this is presumably a reflection of changes in the calibration of the assay.

C. Changes in laboratory indicators 1988-1989 and 1989-1990

Tables C1 - Changes 1988 to 1989 n=483
C2 - Changes 1989 to 1990 n=280

Tables C1 and C2 present the changes in patients between successive annual reviews. The tables are split into five subsections the first summarises all cases, the next four sections give results for the individual treatment types.

Each table shows for a range of indicators the number of valid cases available (and underneath the proportion of the total) followed by the number of cases classified as unsatisfactory(Hi) in the first year and then the second. The numbers who have changed during the year are then shown.

For example table C1 - change 1988 to 1989.
A total of 483 cases came to both annual reviews. We have 443 of these (92%) with two valid HbA1 values. Of these 157 (35%) were 'unsatisfactory'
in 1988, whilst 213 (48%) were unsatisfactory in 1989.
195 were low at the start and stayed low, 122 were high and stayed high,
35 got 'better' moving from high to low, 91 got 'worse' moving from low to high.

Comments:

HbA1 - there is no net fall in patients classified as high over the year -rather patients appear worse. Approximately 80 patients have got 'worse' with only 30 getting 'better' leaving a net change of 50 - spread across all patient types. The proportion classified as high in both years varies by patient type ranging from 10% of diet patients to 44% of Insulin (age at dx<35).
The changes during 1989 appear alarming. Contact with the lab suggest that there has been a shift in average HbA1 levels of about 1 unit % during the year yet there have been no conscious change in the assay procedures. The changes for the first six months of 1990 are not as abnormal though still in the 'wrong direction'.

Cholesterol - There is some general improvement - fewer patients over 6.5 in the second year. Once again the proportion high in both years seems large at 11-20% in 1989 and up to 28% of diet cases in 1990. In 1989 there appeared net improvements in the diet

and insulin (dx>35) groups, these are not seen in 1990.

BMI - as with HbA1 more patients getting 'worse' than 'better' - particularly in the Diet and Insulin (age at dx>35) group where 43% and 57% are over the 27.5 kg/m2 threshold in 1990 with very little change.

This general view of the 'stability' of the annual review population is the same as that from the earlier report. Is this what we should expect for the future? Are there ways in which we can construct better indicators eg concentrating on just the very unsatisfactory patients?

D. Changes in clinical indicators 1988-1989 and 1989-1990.

Tables D1. - Changes 1988-1989 n=483
D2. - Changes 1989-1990 n=280

Tables D1 and D2 deal with the changes in the incidence of clinical problems between successive annual reviews. As changes in this group tend to be slower and more difficult to control less detail is given.

Table D1 charts the incidence of the clinical problems expressed as a percentage of the total cases. The figure in brackets represents the change from the previous year (ie 1988 values minus 1989 values).
Thus in 1989 67% of patients had an ok blood pressure, this was 3% less than 1989.

Comments:

Blood Pressure - In general more patients had problems in the second year than the first (for both 1988 and 1989). These numbers appear rather unstable, which may reflect the relatively smaller numbers in each sample.

Visual Acuity - shows marked changes in three categories - presumably reflecting the deterioration in eyesight one might expect of a population of this age. Nevertheless, for those of whose who don't wear glasses these numbers seem alarming.

Circulation - Once again the difference between the younger and the older patient types is marked. The observed incidence of absent pulses shows some fluctuation which are no doubt due to the problems of consistent measurement. The totals across all case types show much more stability and relatively little change year on year. It appears that with these numbers of cases it is not realistic to try and assess the rate at which these problems

will arise within one year. Longer time periods should be adopted though these run into problems with the size and nature of the sample - relatively few cases will have observations over say 3-4 years.

Retinopathies - The incidence of abnormal fundi (background retinop) appears fairly stable with markedly higher values in the Insulin (dx < 35) group, and lower in the diet group. This pattern is not repeated in the incidence of maculopathy or proliferative retinopathy where the Insulin (dx > 25) and tablet groups score worse. Once again with low numbers it is difficult to be certain of the rate of change in any one year.

Neuropathies - The observation of neuropathic symptoms in approximately one quarter of cases seems quite reliable, with the insulin (dx > 35) groups faring rather worse than the others. There appear to be relatively large increases during 1988 - especially amongst the diet group, given that this condition is potentially reversible (or so I am told) this seems rather alarming.

The biothesiometer readings are again unreliable and show massive swings -probably due to the calibration and the high incidence of 'not recorded'.

Nephropathies - the albustix figures are showing some rather erratic behaviour as observed earlier, with relatively large increases in positive results in 1989 yet relatively large decreases in 1988. I would suggest the creatinine gives a clearer message.

Smoking - Though the proportion of smokers seems fairly stable somewhere between 15-20% of the total, there are a higher proportion in the younger insulin (dx < 35) group.

E.NON-ATTENDERS FOR ANNUAL REVIEW 1989

In 1989 16 patients defaulted from the annual review clinic and were not seen in the regular clinic. A further 50 patients were not sent an annual review appointment and were not seen in the regular clinic - they are being chased at the moment.

Data was collected from 80-85% of the patients who were attending the regular review clinic but did not attend for annual review in 1989 - a total of 123 records. HbAl and BMI were the only metabolic indicators collected as **others were not recorded in a sufficient number of patients** to allow analysis. The data collected was the last recorded in 1989 (or the first in 1990 if there were no records after September 1989), and values 12-15 months previously.

As with the annual review analysis, the patients have been split into four patient/treatment groups:

HbAl

Table El shows the percentages of cases classified as either high or low in 1988 and 1989, with respect to HbAl (cut off value = 8.75%). Values in brackets show comparable annual review percentages.

TABLE El

	Lo Lo	Hi Lo	Lo Hi	Hi Hi
All	32 (44)	10 (8)	23 (21)	35 (28)
Ins(dx<35)	6 (27)	10 (8)	31 (21)	53 (44)
Ins(dx>35)	14 (31)	9 (8)	36 (20)	41 (42)
Tablet	36 (41)	14 (10)	18 (21)	32 (28)
Diet	72 (64)	8 (5)	8 (20)	12 (10)

As with the annual review attenders there is an overall rise (of 13%) in those patients classified as high over the year. However the proportion of patients high over both years is greater (35% v. 28%) and the proportion of patients remaining low over both years is smaller (32% v. 44%) than the patients who received annual reviews - these points are more marked in both insulin treated groups.

BMI

Table E2 - % of cases classified as high or low in 1988 and 1989 with respect to BMI - Cut-off = 27.5 kgm-2. Figures in brackets
show comparable annual review percentages.

TABLE E2

	Lo Lo	Hi Lo	Lo Hi	Hi Hi
TOTAL	56 (55)	4 (4)	6 (7)	34 (34)
Ins(dx<35)	75 (79)	6 (3)	3 (3)	16 (14)
Ins(dx>35)	54 (58)	0 (3)	18 (12)	27 (27)
Tablet	50 (50)	2 (5)	0 (5)	48 (39)
Diet	44 (43)	8 (4)	8 (10)	40 (43)

There is no significant difference between patients who receive an annual review and those who do not, whether you look at the group as a whole or look at different patient/treatment groups.

As with annual review attenders there is a slight overall rise in the proportion of patients classified as high over the year.

EYES, FEET AND BP

Table E3 shows the percentage of patients in whose notes it was documented that their eyes, feet and blood pressure had been checked, either in the routine clinic or during an in-patient stay.

TABLE E3

% checked	Eyes	Feet	BP
TOTAL	31	32	68
Ins(dx<35)	25	13	34
Ins(dx>35)	54	59	82
Tablet	34	34	82
Diet	12	28	76

It is worrying that only 1/3 of the patients have their eyes and feet checked if they do not receive an annual review, and 2/3 of the patients have their blood pressure monitored. On the whole the diet only patients are the least monitored, and the insulin (age at dx > 35) are more closely observed.

SUMMARY Non Annual review Cases

Although patients weight does not seem to be affected if they do not receive an annual review, the data contained in Tables Dl and D3 serve to stress the importance of all diabetics having an annual check. So that the HbAl levels will be lowered slightly and to ensure that their eyes, feet and blood pressure are monitored annually to detect early signs of abnormality.

Appendix I

Categorical Variables - coded when subset created:

BP	Female		Male	
age <=30	> 140	>85	>145	>90
age <=40	> 150	>90	>150	>90
age >40	> 160	>90	>160	>90
otherwise	> 160	>90	>160	>90

Values roughly equate to 95th percentiles for normal populations where over 160/90.

Scored BPSUM 1=systolic, 2=diastolic , 3=both

Visual Acuity (best eye)
0=better than 6/12
1=better than 6/36 and less than 6/12
2=worse than 6/36 or counting fingers or seeing light
9= none of the above (quite a lot)

Circulation
0=ok
1=absent pulses (left or right)
2=claudication
3=ulcers
4=amputation (any)

Retinopathy:
Abnormal fundi =Yes/No Fundi abnormal or background retinopathy
(R or L retinal haemorrhages, hard exudates, venous abnormalities soft exudates)
Proliferative =Yes/No Proliferative
(R or L venous abnormalities, vitreous haemorrhages)
Maculopathy =Yes/No Maculopathy

Angina Yes/No
Smoking Yes/No .. any smoking
Neuropathy Yes/No .. neuropathic symptoms
Claudication Yes/No
Impotence Yes/No or nor relevant

Biothesiometer Yes/No = <25 left or right
Albustix Yes=any positive

Continuous Variables:

1. HbA1(%)
   0 = No value
   1 = Satisfactory >0 and <8.75
   2 = Unsatisfactory >8.75 and <9.5
   3 = Poor >10.0

2.BMI
   0 = No value
   1 = Satisfactory <25
   2 = Unsatisfactory >25 and <27.5
   3 = Poor >27.5

3. Chol
   0 = No value
   1 = Satisfactory >5.2
   2 = Unsatisfactory >5.2 and <6.5
   3 = Poor >6.5

4. CholHDL
   0 = No value
   1 = Poor <0.9
   2 = Unsatisfactory 0.9-1.1
   3 = Good >1.1

5. Triglyc
   0 = No value
   1 = Satisfactory <2.5
   2 = Unsatisfactory 2.5-4
   3 = Poor > 4 Poor

6. Creatinine
   1 = <125
   2 = >125

7. Microalbumen
   1=<3.5
   2=3.5- 10
   3=> 10 Poor

Table B1 Cross Section 1987

Column Percentages(%)

	Ins Dx<35	Ins Dx>35	Tab let	Diet	Total
No	88	56	150	136	430
Male	56.8	55.4	59.3	65.4	60
Female	43.2	44.6	40.7	34.6	40
Age (Not rec)	4.5	5.4	2.7	0.7	3
<40 yrs	71.6	17.9	2.7	5.1	20
40-59 yrs	18.2	35.7	30.0	35.3	30
60-69 yrs	4.5	26.8	34.7	32.4	27
>70 yrs	1.1	14.3	30.0	26.5	21
HbA1 (Not rec)	3.4	8.9	5.3	7.4	6
<= 7.5	25.0	12.5	27.3	41.2	29
7.5 - 8.75	27.3	17.9	26.0	23.5	24
8.75 - 10.0	19.3	30.4	21.3	14.0	20
> 10.0	25.0	30.4	20.0	14.0	20
Cholesterol (N/R)	3.4	3.6	5.3	2.2	4
<=5.2	54.5	32.1	20.7	31.6	33
5.2 - 6.5	26.1	42.9	40.7	41.2	38
>6.5	15.9	21.4	33.3	25.0	26
BMI (Not rec)	1.1	1.8	2.7	0.7	2
<=25	55.7	55.4	43.3	26.5	42
25 - 27.5	26.1	25.0	20.7	27.2	24
>27.5	17.0	17.9	33.3	45.6	32
Chol HDL (N/R)	20.5	16.1	20.0	16.2	18
>1.1	54.5	53.6	28.0	31.6	38
0.9 - 1.1	19.3	17.9	18.7	20.6	19
<=0.9	5.7	12.5	33.3	31.6	24
Creatinine (N/R)	2.3	8.9	8.7	5.1	6
<=125	97.7	75.0	80.7	86.8	85
>125	0.0	16.1	10.7	8.1	8
Triglycerides (N/R	6.8	5.4	7.3	5.1	6
<=2.5	79.5	73.2	40.7	64.0	60
2.5 - 4.0	5.7	12.5	35.3	21.3	22
>4.0	8.0	8.9	16.7	9.6	12

	Dx<35	Dx>35	let	Diet	Total
	88	56	150	136	430
Microalbumen (N/R	98.9	100.0	100.0	99.3	100
<=3.5	0.0	0.0	0.0	0.0	0
3.5 - 10.0	1.1	0.0	0.0	0.7	0
>10.0	0.0	0.0	0.0	0.0	0
Blood Pressure Ok	87.5	87.5	66.7	66.9	74
Systolic high	4.5	3.6	22.0	18.4	15
Diastoloic high	5.7	0.0	2.7	4.4	3
S + D high	2.3	5.4	6.7	8.1	6
Not recorded	0.0	3.6	2.0	2.2	2
Visual acuity Ok	92.0	89.3	84.7	90.4	89
6/12 to 6/36	2.3	5.4	12.7	7.4	8
Worse than 6/36	1.1	0.0	2.0	0.0	1
Other/not rec	4.5	5.4	0.7	2.2	3
Circulation Ok	70.5	57.1	56.0	64.0	62
Pulses	28.4	28.6	22.0	22.8	24
Claudication	0.0	3.6	14.7	8.8	8
Ulcer	0.0	0.0	5.3	2.2	3
Amputation	1.1	10.7	2.0	2.2	3
Eyes - Ab. fundi	28.4	37.5	30.7	14.0	26
Not observe	0.0	0.0	0.0	0.0	0
Maculopathy	3.4	3.6	8.0	3.7	5
Proliferati	10.2	7.1	8.0	2.9	7
Angina	2.3	12.5	18.7	19.9	15
Neuropath. Symptom	9.1	30.4	19.3	20.6	19
Biothesiometer <25	78.4	41.1	27.3	27.9	40
Not recorded	3.4	8.9	4.0	5.9	5
Albustix +ve	21.6	8.9	14.0	9.6	13
Not recorded	0.0	0.0	0.0	0.0	0
Smoking (any)	22.7	28.6	20.7	11.8	19

Ins Ins Tab

Table B2 ..Cross-Section 1988

Column Percentages(%)

	Ins Dx<35	Ins Dx>35	Tab let	Diet	Total
No=	105	80	189	173	547
Male	58.1	62.5	55.6	67.6	61
Female	41.9	37.5	44.4	32.4	39
Age (Not rec)	1.9	1.3	1.1	7.5	3
<40 yrs	69.5	16.3	1.1	2.3	17
40-59 yrs	25.7	37.5	32.3	34.7	33
60-69 yrs	2.9	28.8	34.9	27.7	26
>70 yrs	0.0	16.3	30.7	27.7	22
HbA1 (Not rec)	4.8	2.5	2.6	6.4	4
<= 7.5	21.9	22.5	32.3	60.7	38
7.5 - 8.75	21.9	27.5	24.9	17.9	22
8.75 - 10.0	29.5	26.3	17.5	8.1	18
> 10.0	21.9	21.3	22.8	6.9	17
Cholesterol (N/R)	1.0	0.0	3.2	1.7	2
<=5.2	54.3	28.8	20.1	28.3	31
5.2 - 6.5	29.5	41.3	40.7	37.0	37
>6.5	15.2	30.0	36.0	32.9	30
BMI (Not rec)	0.0	2.5	3.2	3.5	3
<=25	54.3	41.3	31.7	26.6	36
25 - 27.5	26.7	28.8	22.2	26.6	25
>27.5	19.0	27.5	42.9	43.4	36
Chol HDL (Not rec)	9.5	22.5	15.9	12.7	15
>1.1	64.8	46.3	34.4	39.3	44
0.9 - 1.1	20.0	16.3	21.2	20.8	20
<=0.9	5.7	15.0	28.6	27.2	22
Creatinine (N/R)	6.7	7.5	5.3	5.8	6
<=125	90.5	83.8	84.1	86.7	86
>125	2.9	8.8	10.6	7.5	8
Triglycerides (N/R)	1.9	2.5	5.3	1.7	3
<=2.5	86.7	57.5	41.3	61.8	59
2.5 - 4.0	9.5	17.5	34.4	23.1	24
>4.0	1.9	22.5	19.0	13.3	14

Column Percentages(%)

	Dx<35	Dx>35	let	Diet	Total
No=	105	80	189	173	547
Microalbumen (N/R)	51.4	41.3	54.0	53.8	52
<=3.5	0.0	0.0	0.0	0.0	0
3.5 - 10.0	47.6	55.0	42.3	45.7	46
>10.0	1.0	3.8	3.7	0.6	2
Blood Pressure Ok	87.6	81.3	62.4	64.2	71
Systolic high	1.9	7.5	21.7	19.7	15
Diastoloic high	6.7	6.3	6.3	4.6	6
S + D high	2.9	3.8	8.5	9.2	7
Not recorded	1.0	1.3	1.1	2.3	1
Visual acuity (OK)	94.3	88.8	91.5	91.9	92
6/12 to 6/36	2.9	6.3	3.7	3.5	4
Worse than 6/36	1.0	0.0	2.1	0.6	1
Other/not rec	1.0	5.0	2.6	4.0	3
Circulation Ok	91.4	72.5	63.0	71.1	72
Pulses	3.8	7.5	18.0	9.2	11
Claudication	1.9	12.5	13.8	13.9	11
Ulcer	1.0	3.8	1.6	1.2	2
Amputation	1.0	3.8	3.7	4.6	3
Eyes - Ab fundi	30.5	35.0	30.7	17.9	27
Not observe	0.0	0.0	0.0	0.0	0
Maculopathy	3.8	7.5	10.1	4.6	7
Proliferati	8.6	7.5	8.5	2.9	7
Angina	2.9	11.3	20.6	18.5	15
Neuropath. Symptom	9.5	28.8	17.5	15.0	17
Biothesiometer <25	71.4	35.0	24.3	29.5	37
Not recorded	10.5	10.0	10.1	9.8	10
Albustix +ve	17.1	16.3	13.8	9.8	14
Not recorded	4.8	3.8	1.6	2.3	3
Smoking (any)	27.6	21.3	18.0	16.2	20

Table B3. Cross Section 1989

Column Percentages (%)

	Ins Dx<35	Ins Dx>35	Tab let	Diet	Total
No=	122	110	209	196	637
Male	59.8	56.4	57.4	62.2	59
Female	40.2	43.6	42.6	37.8	41
Age (Not rec)	0.0	0.9	1.0	2.0	1
<40 yrs	68.0	8.2	1.9	3.1	16
40-59 yrs	25.4	46.4	31.6	29.6	32
60-69 yrs	4.9	27.3	35.9	35.7	28
>70 yrs	1.6	18.2	30.1	29.6	22
HbA1 (Not rec)	3.3	2.7	4.3	5.6	4
<= 7.5	14.8	18.2	26.8	48.0	30
7.5 - 8.75	20.5	20.0	23.9	24.5	23
8.75 - 10.0	33.6	28.2	17.2	14.3	21
> 10.0	27.9	31.8	28.2	7.7	22
Cholesterol (N/R)	9.0	11.8	6.7	12.2	10
<=5.2	50.0	31.8	25.4	21.9	30
5.2 - 6.5	27.0	40.0	34.9	42.9	37
>6.5	13.9	17.3	33.5	23.0	24
BMI (Not rec)	0.8	4.5	3.8	2.6	3
<=25	45.9	37.3	31.6	20.9	32
25 - 27.5	32.8	24.5	19.1	23.5	24
>27.5	20.5	34.5	45.9	53.1	41
Chol HDL (Not rec)	18.0	24.5	21.1	32.7	25
>1.1	59.0	39.1	30.6	27.6	37
0.9 - 1.1	11.5	14.5	22.0	14.8	16
<=0.9	11.5	22.7	26.8	25.0	23
Creatinine (N/R)	5.7	6.4	5.3	5.6	6
<=125	91.8	79.1	84.2	86.7	86
>125	2.5	15.5	11.0	7.7	9
Triglycerides (N/R)	9.0	11.8	8.1	12.2	10
<=2.5	82.0	60.9	48.8	47.4	57
2.5 - 4.0	6.6	15.5	27.8	28.1	22
>4.0	2.5	12.7	15.8	12.2	12

	Dx<35	Dx>35	let	Diet	Total
	122	110	209	196	637
Microalbumen (Not	99.2	100.0	100.5	99.5	100
<=3.5	0.0	0.0	0.0	0.0	0
3.5 - 10.0	0.8	0.9	0.0	0.5	0
>10.0	0.0	0.0	0.0	0.0	0
Blood Pressure Ok	86.9	79.1	66.5	66.8	73
Systolic high	6.6	10.0	21.1	14.3	14
Diastolic high	3.3	3.6	6.7	6.6	5
S + D high	1.6	7.3	5.3	11.2	7
Not recorded	1.6	0.9	1.0	1.0	1
Visual acuity (OK)	95.1	93.6	91.4	91.8	93
6/12 to 6/36	2.5	5.5	5.7	4.1	5
Worse than 6/36	0.8	0.9	1.9	1.0	1
Other/not rec	1.6	0.9	1.4	3.1	2
Circulation Ok	89.3	69.1	63.6	68.4	71
Pulses	7.4	14.5	17.2	14.3	14
Claudication	1.6	10.0	15.3	14.8	12
Ulcer	0.8	1.8	1.0	1.5	1
Amputation	0.8	5.5	3.3	1.0	3
Eyes - Ab fundi		27.9	28.2	26.3	15.8	24
Not observe	0.0	0.0	0.0	0.0	0
Maculopathy	4.1	7.3	8.6	3.1	6
Proliferati	7.4	6.4	6.2	3.6	6
Angina	3.3	18.2	23.4	21.9	18
Neuropath. Symptom	16.4	28.2	27.3	22.4	24
Biothesiometer <25	46.7	18.2	24.4	24.5	28
Not recorded	35.2	48.2	43.1	44.4	43
Albustix +ve	3.3	3.6	7.2	3.6	5
Not recorded	0.0	0.0	0.0	0.0	0
Smoking (any)	18.9	20.0	14.4	14.8	16

Table B4 1990 Cross section (to July)

Column Percentages(%)

	Dx<35	Ins Dx>35	Ins let	Tab Diet	Total
No=		77	86	143	120	426
Male		49.4	54.7	54.5	65.8	57
Female		50.6	45.3	45.5	34.2	43
Age	(Not rec)	0.0	0.0	0.0	0.0	0
	<40 yrs	66.2	7.0	1.4	1.7	14
	40-59 yrs	28.6	40.7	31.5	28.3	32
	60-69 yrs	2.6	30.2	37.1	35.0	29
	>70 yrs	2.6	22.1	30.1	35.0	25
HbA1	(Not rec)	2.6	1.2	4.2	1.7	3
	<= 7.5	16.9	17.4	23.1	45.0	27
	7.5 - 8.75	13.0	25.6	23.8	27.5	23
	8.75 - 10.0	26.0	31.4	21.0	16.7	23
	> 10.0	41.6	24.4	28.0	9.2	24
Cholesterol	(N/R)	1.3	2.3	2.8	0.8	2
	<=5.2	46.8	39.5	19.6	22.5	29
	5.2 - 6.5	37.7	34.9	44.1	45.8	42
	>6.5	14.3	23.3	33.6	30.8	27
BMI	(Not rec)	0.0	4.7	0.0	1.7	1
	<=25	45.5	37.2	26.6	18.3	30
	25 - 27.5	27.3	19.8	25.9	24.2	24
	>27.5	27.3	38.4	47.6	55.8	44
Chol HDL	(Not rec)	1.3	5.8	4.2	1.7	3
	>1.1	76.6	57.0	43.4	42.5	52
	0.9 - 1.1	13.0	16.3	27.3	20.0	20
	<=0.9	9.1	20.9	25.2	35.8	24
Creatinine	(N/R)	5.2	8.1	10.5	5.8	8
	<=125	92.2	81.4	83.2	83.3	85
	>125	2.6	10.5	6.3	10.8	8
Triglycerides	(N/R)	1.3	2.3	2.8	0.8	2
	<=2.5	84.4	73.3	50.3	60.0	64
	2.5 - 4.0	11.7	10.5	32.2	27.5	23
	>4.0	2.6	14.0	14.7	11.7	12

	Dx<35	Dx>35	let	Diet	Total
	77	86	143	120	426
Microalbumen (N/R)	67.5	72.1	62.2	70.0	67
<=3.5	23.4	15.1	20.3	18.3	19
3.5 - 10.0	6.5	8.1	11.2	9.2	9
>10.0	2.6	4.7	6.3	2.5	4
Blood Pressure Ok	76.6	65.1	60.8	50.8	62
Systolic high	11.7	17.4	14.7	23.3	17
Diastolic high	5.2	7.0	6.3	8.3	7
S + D high	2.6	9.3	18.2	16.7	13
Not recorded	3.9	1.2	0.0	0.8	1
Visual acuity (OK)	92.2	82.6	81.8	85.0	85
6/12 to 6/36	6.5	16.3	12.6	12.5	12
Worse than 6/36	0.0	1.2	5.6	1.7	3
Other/not rec	1.3	0.0	0.0	0.8	0
Circulation Ok	93.5	70.9	64.3	66.7	72
Pulses	3.9	11.6	16.8	15.0	13
Claudication	2.6	15.1	14.0	14.2	12
Ulcer	0.0	2.3	2.8	1.7	2
Amputation	0.0	0.0	2.1	2.5	1
Eyes - Ab. fundi	31.2	25.6	35.0	13.3	26
Not observe	0.0	0.0	0.0	0.0	0
Maculopathy	5.2	7.0	11.9	4.2	8
Proliferati	1.3	4.7	10.5	2.5	5
Angina	5.2	15.1	26.6	19.2	18
Neuropath. Symptom	16.9	32.6	20.3	25.8	24
Biothesiometer <25	74.0	50.0	50.3	57.5	57
Not recorded	19.5	9.3	4.9	10.0	10
Albustix +ve	10.4	10.5	9.8	7.5	9
Not recorded	1.3	1.2	0.0	0.0	0
Smoking (any)	29.9	24.4	13.3	15.8	19

Table C1(a) Changes In Lab indicators 1988-1989

All cases Total=483
NB. Hi=Unsatisfactory (not always high)

	Total	No. unsatisf. 88	89	LoLo	Changes HiLo	88 to 89 LoHi	HiHi
HbA1	443	157	213	195	35	91	122
	92%	35%	48%	44%	8%	21%	28%
Chol HDL	360	138	142	107	35	31	187
	75%	38%	39%	52%	9%	10%	30%
Body Mas	461	175	190	252	19	34	156
	95%	38%	41%	55%	4%	7%	34%
Choleste	441	148	129	256	56	37	92
	91%	34%	29%	58%	13%	8%	21%
Creatini	425	39	50	364	11	22	28
	88%	9%	12%	86%	3%	5%	7%
Triglyce	435	184	164	216	55	35	129
	90%	42%	38%	50%	13%	8%	30%
Microalb	60	1	12	47	1	12	0
	12%	2%	20%	78%	2%	20%	0%

Table C1 (b) Changes in lab indicators 1988-1989

Insulin treated (Age at diagnosis >35 years) Total=69

	Total	No. unsatisf. 88	89	Changes 88 to 89 LoLo	HiLo	LoHi	HiHi
HbA1	65	32	40	20	5	13	27
	94%	49%	62%	31%	8%	20%	42%
Chol HDL	45	15	14	12	2	3	28
	65%	33%	31%	62%	7%	4%	27%
Body Mas	66	20	26	38	2	8	18
	96%	30%	39%	58%	3%	12%	27%
Choleste	62	23	19	35	8	4	15
	90%	37%	31%	56%	13%	6%	24%
Creatini	63	7	8	52	3	4	4
	91%	11%	13%	83%	5%	6%	6%
Triglyce	60	27	20	29	11	4	16
	87%	45%	33%	48%	18%	7%	27%
Microalb	9	0	2	7	0	2	0

Insulin treated (Age at diagnosis <35 years) Total=92

	Total	No. unsatisf. 88	89	Changes 88 to 89 LoLo	HiLo	LoHi	HiHi
HbA1	85	44	55	23	7	18	37
	92%	52%	65%	27%	8%	21%	44%
Chol HDL	71	12	11	5	6	7	53
	77%	17%	15%	75%	10%	8%	7%
Body Mas	91	16	16	72	3	3	13
	99%	18%	18%	79%	3%	3%	14%
Choleste	84	16	13	64	7	4	9
	91%	19%	15%	76%	8%	5%	11%
Creatini	80	3	5	73	2	4	1
	87%	4%	6%	91%	3%	5%	1%
Triglyce	83	8	8	69	6	6	2
	90%	10%	10%	83%	7%	7%	2%
Microalb	12	0	1	11	0	1	0

Treated by tablet Total=171

	Total	88	89	LoLo	HiLo	LoHi	HiHi
HbA1	158	60	77	65	16	33	44
	92%	38%	49%	41%	10%	21%	28%
Chol HDL	131	63	66	52	14	11	54
	77%	48%	50%	41%	8%	11%	40%
Body Mas	159	71	71	80	8	8	63
	93%	45%	45%	50%	5%	5%	40%
Choleste	156	59	60	77	19	20	40
	91%	38%	38%	49%	12%	13%	26%
Creatini	150	16	25	124	1	10	15
	88%	11%	17%	83%	1%	7%	10%
Triglyce	153	89	80	52	21	12	68
	89%	58%	52%	34%	14%	8%	44%
Microalb	20	1	4	15	1	4	0

Treated by diet only Total cases=151

	Total	88	89	LoLo	HiLo	LoHi	HiHi
HbA1	135	21	41	87	7	27	14
	89%	16%	30%	64%	5%	20%	10%
Chol HDL	113	48	51	38	13	10	52
	75%	42%	45%	46%	9%	12%	34%
Body Mas	145	68	77	62	6	15	62
	96%	47%	53%	43%	4%	10%	43%
Choleste	139	50	37	80	22	9	28
	92%	36%	27%	58%	16%	6%	20%
Creatini	132	13	12	115	5	4	8
	87%	10%	9%	87%	4%	3%	6%
Triglyce	139	60	56	66	17	13	43
	92%	43%	40%	47%	12%	9%	31%
Microalb	19	0	5	14	0	5	0

Table C2(a) Changes in lab indicators 1989-Jul 1990

All cases Total=280

	Total	No. high in 89	No. high in J90	Changes 89 to J90 LoLo	HiLo	LoHi	HiHi
HbA1	263	124	140	92	31	47	93
	94%	47%	53%	35%	12%	18%	35%
Chol HDL	193	69	67	52	15	17	109
	69%	36%	35%	56%	9%	8%	27%
Body Mas	273	124	126	136	11	13	113
	98%	45%	46%	50%	4%	5%	41%
Choleste	251	67	73	164	14	20	53
	90%	27%	29%	65%	6%	8%	21%
Creatini	247	25	25	215	7	7	18
	88%	10%	10%	87%	3%	3%	7%
Triglyce	249	80	75	147	27	22	53
	89%	32%	30%	59%	11%	9%	21%
Microalb	0	0	0	0	0	0	0

Table C2(b) Changes in Lab indicators 1989-July 1990

Insulin treated (Age at diagnosis >35 years) Total=59

	Total	No. high in 89	No. high in J90	Changes 89 to J90 LoLo	HiLo	LoHi	HiHi
HbA1	57	35	31	14	12	8	23
	97%	61%	54%	25%	21%	14%	40%
Chol HDL	41	14	14	11	3	3	24
	69%	34%	34%	59%	7%	7%	27%
Body Mas	56	22	22	31	3	3	19
	95%	39%	39%	55%	5%	5%	34%
Choleste	53	10	13	40	0	3	10
	90%	19%	25%	75%	0%	6%	19%
Creatini	53	8	7	45	1	0	7
	90%	15%	13%	85%	2%	0%	13%
Triglyce	53	13	14	36	3	4	10
	90%	25%	26%	68%	6%	8%	19%
Microalb	0	0	0	0	0	0	0

Insulin treated (Age at diagnosis <35 years) Total=50

	Total	No. high in 89	No. high in J90	Changes 89 to J90 LoLo	HiLo	LoHi	HiHi
HbA1	46	34	38	7	1	5	33
	92%	74%	83%	15%	2%	11%	72%
Chol HDL	44	9	9	7	2	2	33
	88%	20%	20%	75%	5%	5%	16%
Body Mas	50	15	14	31	5	4	10
	100%	30%	28%	62%	10%	8%	20%
Choleste	45	6	7	37	1	2	5
	90%	13%	16%	82%	2%	4%	11%
Creatini	45	0	2	43	0	2	0
	90%	0%	4%	96%	0%	4%	0%
Triglyce	45	5	5	37	3	3	2
	90%	11%	11%	82%	7%	7%	4%
Microalb	0	0	0	0	0	0	0

Treated by tablet Total=81

	Total	No. high in 89	No. high in J90	Changes 89 to J90 LoLo	HiLo	LoHi	HiHi
HbA1	74	34	41	25	8	15	26
	91%	46%	55%	34%	11%	20%	35%
Chol HDL	54	20	22	15	7	5	27
	67%	37%	41%	50%	9%	13%	28%
Body Mas	79	35	37	41	1	3	34
	98%	44%	47%	52%	1%	4%	43%
Choleste	73	23	25	41	7	9	16
	90%	32%	34%	56%	10%	12%	22%
Creatini	66	8	7	56	3	2	5
	81%	12%	11%	85%	5%	3%	8%
Triglyce	71	31	33	29	9	11	22
	88%	44%	46%	41%	13%	15%	31%
Microalb	0	0	0	0	0	0	0

Treated by diet only Total cases=90

	Total	No. high in 89	No. high in J90	Changes 89 to J90 LoLo	HiLo	LoHi	HiHi
HbA1	86	21	30	46	10	19	11
	96%	24%	35%	53%	12%	22%	13%
Chol HDL	54	26	22	19	3	7	25
	60%	48%	41%	46%	13%	6%	35%
Body Mas	88	52	53	33	2	3	50
	98%	59%	60%	38%	2%	3%	57%
Choleste	80	28	28	46	6	6	22
	89%	35%	35%	58%	8%	8%	28%
Creatini	83	9	9	71	3	3	6
	92%	11%	11%	86%	4%	4%	7%
Triglyce	80	31	23	45	12	4	19
	89%	39%	29%	56%	15%	5%	24%
Microalb	0	0	0	0	0	0	0

Table D1 Incidence of clinical complications 1988-1989

Percentage 89 [% change 88 to(-) 89]

Variable	Total	Ins dx<35	Ins Dx>35	Tablet	Diet	
Number cases=	483	98	83	177	129	
BP ok	67% [3]	85% [1]	66% [8]	64% [0]	57% [7]	
BP Sys	17% [-1]	8% [-5]	19% [-11]	18% [4]	19% [1]	-
BP Dias	5% [0]	3% [3]	4% [5]	5% [0]	9% [-5]	-
BP Both	10% [-2]	1% [3]	11% [-5]	10% [-3]	14% [-4]	-
BP NotRec	1% [1]	3% [-2]	0% [2]	0% [1]	1% [1]	
V Acuity ok	88% [4]	96% [-1]	87% [2]	84% [7]	86% [8]	
6/12 to 6/36	9% [-5]	3% [-1]	11% [-6]	9% [-3]	11% [-9]	-
Worse 6/36	2% [-1]	0% [1]	1% [-1]	4% [-2]	2% [-2]	-
Not Rec	1% [2]	1% [1]	1% [5]	1% [1]	1% [2]	
Circul Ok	70% [4]	92% [0]	70% [2]	63% [1]	60% [11]	
Pulses	15% [-5]	7% [-3]	11% [-2]	19% [-3]	17% [-9]	-
Claudic	11% [0]	0% [2]	12% [-2]	12% [2]	18% [-2]	-
Ulcer	1% [0]	0% [1]	1% [2]	1% [0]	2% [-2]	-
Amputation	3% [0]	1% [0]	6% [0]	3% [0]	3% [2]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Fundi Ok	74% [-1]	62% [5]	75% [-7]	72% [-2]	84% [0]	
Abnormal	26% [1]	38% [-5]	25% [7]	26% [2]	16% [0]	-
Not rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Maculopathy No	92% [1]	94% [2]	89% [4]	88% [-1]	96% [1]	
Yes	8% [-1]	6% [-2]	11% [-4]	10% [1]	4% [-1]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Prolif Retinop No	94% [-1]	95% [-5]	94% [0]	91% [-1]	96% [2]	
Yes	6% [1]	5% [5]	6% [0]	7% [1]	4% [-2]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Angina No	80% [5]	98% [0]	80% [5]	75% [8]	73% [4]	
Yes	20% [-5]	2% [0]	20% [-5]	23% [-8]	27% [-4]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Neur Symps No	76% [7]	88% [2]	66% [7]	74% [6]	74% [12]	
Yes	24% [-7]	12% [-2]	34% [-8]	23% [-6]	26% [-12]	-
Not Rec	0% [0]	0% [0]	0% [1]	1% [0]	0% [1]	
Biothesiom <25	26% [27]	6% [10]	33% [27]	31% [57]	29% [35]	
>25	45% [-8]	68% [4]	37% [-8]	38% [7]	39% [-12]	-
Not Rec	30% [-19]	26% [-14]	30% [-18]	29% [-12]	33% [-23]	
Albustix -ve	93% [-8]	94% [-12]	95% [-17]	90% [-6]	93% [-1]	
+ve	6% [7]	4% [11]	5% [12]	7% [6]	7% [0]	-
Not Rec	0% [2]	1% [2]	0% [5]	0% [1]	0% [1]	
Smoking No	83% [-2]	77% [-3]	86% [-6]	83% [-3]	84% [2]	
Yes	17% [2]	23% [3]	14% [6]	15% [3]	16% [-2]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	

Table D2 Incidence of clinical complications 1989-July 1990

Percentage 90 [% change 89 to(-) 90]

Variable	Total	Ins dx<35	Ins Dx>35	Tablet	Diet	
Number cases=	280	56	61	89	79	
BP Ok	61% [14]	73% [14]	66% [15]	56% [15]	52% [18]	
BP Sys	18% [-6]	14% [-9]	18% [-10]	13% [1]	24% [-8]	-
BP Dias	6% [-1]	4% [-2]	8% [-3]	3% [2]	8% [-4]	-
BP Both	13% [-7]	4% [0]	7% [-2]	21% [-13]	15% [-6]	-
BP NotRec	2% [0]	5% [-4]	2% [0]	0% [1]	1% [0]	
Vis Ac Ok	83% [8]	91% [2]	82% [11]	73% [11]	84% [6]	
VA 6/12 to 36	14% [-9]	7% [-4]	18% [-11]	16% [-9]	14% [-9]	-
VA Worse 6/36	2% [-1]	0% [2]	0% [0]	6% [-3]	1% [0]	-
VA Not Rec	1% [1]	2% [0]	0% [0]	0% [3]	1% [3]	
Circul Ok	71% [-1]	95% [-9]	67% [3]	62% [1]	65% [0]	
Pulses	14% [0]	4% [5]	11% [2]	18% [-4]	18% [0]	-
Claudic	12% [0]	2% [2]	18% [-7]	12% [2]	14% [1]	-
Ulcer	1% [1]	0% [2]	3% [0]	0% [1]	1% [0]	-
Amput	1% [0]	0% [0]	0% [2]	2% [0]	3% [-1]	-
Not rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Fundi Ok	74% [4]	62% [2]	79% [-5]	65% [1]	85% [0]	
Fundi Ab	26% [-4]	37% [-2]	21% [5]	29% [-1]	15% [0]	-
Not rec.	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	-
Maculop No	93% [4]	96% [4]	92% [5]	84% [6]	96% [0]	
Yes	7% [-4]	4% [-4]	8% [-5]	10% [-6]	4% [0]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	-
Prolif Ret No	96% [-2]	98% [-4]	97% [-2]	89% [-2]	97% [0]	
Yes	4% [2]	2% [4]	3% [2]	6% [2]	3% [0]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Angina No	80% [2]	96% [-4]	82% [3]	69% [3]	73% [3]	
Angina Yes	20% [-2]	4% [4]	18% [-3]	26% [-3]	27% [-3]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	
Neur Symp No	75% [0]	84% [-5]	67% [2]	74% [-2]	70% [6]	
Neur Symp Yes	25% [-1]	16% [4]	33% [-2]	20% [2]	30% [-6]	-
Not Rec	0% [0]	0% [2]	0% [0]	0% [0]	0% [0]	
Biothes<25	32% [11]	5% [27]	43% [2]	37% [16]	35% [18]	
Biothes>25	57% [-37]	77% [-45]	48% [-41]	51% [-27]	53% [-34]	-
Not Rec	11% [26]	18% [18]	10% [39]	7% [39]	11% [16]	
Albustix -ve	89% [8]	86% [12]	90% [3]	83% [9]	92% [6]	
Albutsix +ve	10% [-6]	11% [-9]	10% [-3]	10% [-7]	8% [-6]	-
Not Rec	0% [0]	2% [-2]	0% [0]	0% [0]	0% [0]	
Smoking No	81% [-1]	66% [2]	75% [-2]	82% [-1]	90% [-3]	
Smoking Yes	19% [1]	34% [-2]	25% [2]	12% [1]	10% [3]	-
Not Rec	0% [0]	0% [0]	0% [0]	0% [0]	0% [0]	

Appendix 6(c) Minutes of CASPE Diabetes Outcomes Meeting
Friday 11th January, 1991 at 12.00 a.m.

Present:-

Martin Bardsley
Donna Swinden
Judy Wilbourne
Alison McCallum
Ann Wilson
Felicity Tunbridge
Philip Home

Apologies from:- Angela Skinner
Sally Marshall
Gill Sanders

Agenda

1. New Staff and functions.
2. Annual review data - HbA1 assays
   - Multi-hospital meeting?
3. G.P's study.
4. Progress on Insulin starters.
5. New patients - latest results.
6. Health status survey results.
7. More Information to Patients? Progress.
8. Any other business.
9. Progress - writing papers.

Minutes

1. Introduced to Ann Wilson - the new 'H' grade Audit Nurse employed by the Diabetic Unit. PH explained that National Audit is designed to assess whether it is feasible to collect information on performance indicators in different diabetes centres.

   It is partly funded by the Department of Health. Ann will not be directly involved with CASPE. Angela's replacement is Niamh Hanlon, who will be starting later this month. She will be more involved with CASPE.

2a HbA1 assays - reference the shift of 0.7%. PH has been in touch with biochemists. The higher mean level of 0.5 in 1989 also affecte the RVI. The RVI data is more reliable due to pregnant patient monitoring. PH has a graph that demonstrates that the HbA1 levels at the RVI and the Freeman moved in unison during 1989. 20% of patients had been reclassified as a result of the change in levels. This is a

"massive problem" and the biochemists concluded that the assay must be ditched.

Internal quality control is month by month, and therefore the bias is absorbed. It is a general problem as there is only one supplier of reagents. AM suggested writing this up as a paper and PH agreed.

2b Multi - hospital meeting. Martin has produced a series of tables on the Annual Review data. He suggests circulating annual review data only and omitting 3/12 follow up data for the moment.

PH agrees and would like to see 1990 data for purpose of Writing papers FT & PH suggest a separate report for new patients.

Those patients that do not come to the Annual Review Clinic have been identified and coded by DS. MB will get DM report completed. DS to ring Shelagh on Friday 18th Jan, to remind PH to organise a multi- hospital meeting.

3. FT - things are going well with GP study. Tables are automated. Needs to update for the 3rd practise. Waiting for 1990 data from all three practices. Computer program - need to introduce ability to cope with GP's that do not do certain measures. GP's have found the study helpful. PH - asked GP's for short summaries on the usefulness of the study - about one side of A4 - with an emphasis on the changes they have made. MB - wants to put FT's work into final report.
FT - information collected by Diabetes Facilitator is useful as it can help correct BML measurements used by one practise that wasn't measuring patient heights. FT looking for fourth practise. Thinks there are problems. MB thinks there is mileage for a wider basis but not sure if it is possible here. PH suggests that a Central Database is needed and study could benefit from BDA involvement.

PH is to look at form design.
Action on GP's to be agreed by next meeting.

4. FT - negative news in that is is impossible to collect data retrospectively. The good news prospective collection works well, particularly if it is done when JW has the notes. JW identifies insulin starters (from 1/01/91) and 200 forms are completed:- patient personal data and present diabetic status, plus Bradley's Health Questionnaire. JW said some patients have requested greater clarification when talking about 'treatment'. AM had problems with questionnaire but very difficult to change it, PH said not to clarify the question, let the patient use their own interpretation. JW was having problems with the questionnaire herself as regards defining medical problems as opposed to diagnosis and suggested contacting FT/PH when completing forms. FT & PH both agreed. PH suggested that FT catches patients in clinic. Also patients need to be caught later on after initial recruitment. Need to decide how to do this and when.

MB was not sure whether 3/12 follow up will be useful as CASPE are thinking of dropping 3/12 SIP collection. He suggested a Treatment Satisfaction Questionnaire

might be useful at 3/12 follow up - the same one as the initial questionnaire. Also a 12/12 follow up, plus relevant metabolic data. Complications cam be looked at during annual review. Metabolic data to be collected pre-insulin, one year and following year. AMc suggests health status rather than treatment satisfaction. Little experience of treatment satisfaction in diabetes studies. MB suggested trying health status questionnaire in addition to treatment satisfaction questionnaire.

5. New patient latest results are out of sync. with Annual Review data. No great change in SIP's to 3/12 or to 12/12. Education questionnaire wasn't scored correctly but is now. Need to look at scores in relation to other measurements. Significant differences seen between educated and non-educated patients but no large quantitative differences.

   MB suggested dropping 3/12 follow up. It was agreed by the rest meeting. FT asked in the clinical data was ok? DS said there were no problems. JW asked is it necessary to do pre-education questionnaires with new NIDDMs, MB said yes. JW the 1st group of insulin starters have nearly completed education course. Questionnaires will be sent to their homes.

6. Sue Astelle's work - there were no significant differences between clinic completion and post. Order doesn't matter. Reminders vs. immediate responders - no real differences. Non-responders probably are different (MB). Most people score 0 in the different dimensions. The Bradley has a more even distribution as would be expected at it is specially designed for diabetes.

   AMc was not sure what well-being measures. Usefulness may be limited as scores do not always reflect how patient feels at the time, as opposed to generally throughout one's life. Treatment groups have large differences on all scores except emotional reaction, social isolation and treatment satisfaction.

   MB discussed briefly report on SA's study. He concluded that Bradley, is not useful at all in relation to clinical indicators.

   AMc has 3 interviews left to do, and will finish coding.

7. Still in progress. Trying to identify holes in information - inspired by BDA meeting. Need to decide who to send questionnaires to.

8. There was no other business.

9. PH needs to look at MB's reports/data and will write some papers.

Appendix 7(a) Data collection form ANGIOPLASTY (PTCA)

QUESTIONNAIRE NO 1

Pre-operative Assessment

Date: [/ /]
Ward: [ ]

PATIENT IDENTIFICATION:
Name.....
Address.....
.....
Hospital No:.....
Age.....
Telephone No:.....

G.P.Name.....
Address.....
.....
Telephone No:.....

Severity of angina I II III IV [ ]

Significant current/past disease

	Years	Months
Time since beginning of anginal pain	[ ]	[ ]
	Yes	No
Previous myocardial infarction	[ ]	[ ]
Previous CABG	[ ]	[ ]
Previous PTCA	[ ]	[ ]
Chronic respiratory disease	[ ]	[ ]
History of congestive heart failure	[ ]	[ ]
History of hypertension	[ ]	[ ]
Current (or previous) smoker	[ ]	[ ]
History of hyperlipidaemia	[ ]	[ ]

Current Medications

	Specify	Dosage	Frequency/day	For what condition
1				
2				
3				
4				

Summary: Single (S) Double (B) Double + (B+) Triple (T) [ ]

Indications from coronary angiography

Artery involved specify	Percentage (%) 50 - 70	71 - 90	Stenosis >90
Left main	[ ]	[ ]	[ ]
Left anterior descending	[ ]	[ ]	[ ]
Right coronary artery	[ ]	[ ]	[ ]
Circumflex	[ ]	[ ]	[ ]
Vein graft	[ ]	[ ]	[ ]
Other specify.....	[ ]	[ ]	[ ]

Vessel involvement (summary)	Single	Two	Triple
	[ ]	[ ]	[ ]
Ventricular Function	Poor	Moderate	Good
	[ ]	[ ]	[ ]

Summary Indications:
Stable Angina [ ]
Unstable Angina (Emergency) [ ]
Myocardial Infarction [ ]

Other (specify) .....

QUESTIONNAIRE 2

HERE ARE SOME QUESTIONS ABOUT YOUR ABILITY TO EXERCISE

1 A) HOW FAR DO YOU WALK EACH DAY?

	Please tick one box only
drive everywhere?	[ ]
never leave the house?	[ ]
walk less than 200 yards?	[ ]
walk less than quarter of a mile?	[ ]
walk less than half a mile?	[ ]
walk more than half a mile?	[ ]

If more than half a mile, what is the furthest you walk each day?

[ ] miles

B) WHAT IS IT THAT LIMITS YOUR EXERCISE?

Angina	[ ]
Shortness of breath	[ ]
Fatigue	[ ]
Other	[ ]

Please specify.....

C) ARE YOU MORE OR LESS ACTIVE THANWHEN YOU FIRST SUFFERED WITH ANGINA?

More active	[ ]
Less active	[ ]

D) DO YOU STILL SUFFER WITH ANGINAL CHEST PAIN?

If "NO" leave the next page. Go on to the final 3 pages. NO [ ]

If "YES" complete the next page and the final 3 pages. YES [ ]

HERE ARE SOME GENERAL QUESTIONS ABOUT CHEST PAIN/ANGINA
GIVE THE ANSWER THAT BEST DESCRIBES HOW YOU ARE AT PRESENT

DO YOU EXPERIENCE CHEST PAIN WITH ANY OF THESE ACTIVITIES?

- performing unusually strenuous activities eg. hill walking, washing your car, sports like squash/badminton etc., running.

OR
- performing heavy work eg. labouring, rushing for a bus with shopping, washing windows/spring cleaning, digging the garden.

Yes [ ]
No [ ]
Dont know [ ]

WOULD YOU EXPERIENCE CHEST PAIN WHEN YOU.....?

- climb stairs quickly (more than one flight)

OR - walk up-hill with out slowing down

OR - walk more than 400 yards on the flat at a normal pace (keeping up with your friends).

Yes [ ]
No [ ]
Dont know [ ]

WOULD YOU EXPERIENCE CHEST PAIN IF YOU WERE TO PERFORM THESE ACTIVITIES IN:

- in cold or windy conditions
OR - when angry, upset or emotionally stressed

OR - after meals

Yes [ ]
No [ ]
Dont know [ ]

WOULD YOU EXPERIENCE CHEST PAIN IF YOU WERE TO....?

- walk less than 400 yards on the flat at a normal pace (keeping up with your friends)
OR - climb one flight of stairs or less

OR - walk up-hill at a slowed pace (such that your friends have to wait for you).

Yes [ ]
No [ ]
Dont know [ ]

DO YOU EXPERIENCE CHEST PAIN WITH ALL PHYSICAL ACTIVITY?

eg. washing, dressing, any walking (ie. less than 20 yards)

Yes [ ]
No [ ]
Dont know [ ]

THANK YOU FOR YOUR HELP

ANGIOPLASTY (PTCA) QUESTIONNAIRE NO 3

FREEMAN CASPE OUTCOME STUDY

Procedure Details

Patient Identification:

Date of procedure: [/ /]

Name.....

Procedure Indication:
Stable angina Elective [ ]
Unstable angina Emergency [ ]
Mayocardial Infarction [ ]
Previous CABG [ ]
Previous PTCA [ ]

Address.....
.....
.....
Hospital No.....

Operating Cardiologist :.....

Complexity of PTCA Risk of problems: Low [ ] Medium [ ] High [ ]

Successfully dilated?	Yes	No
1	[ ]	[ ]
2	[ ]	[ ]
3	[ ]	[ ]

Number of lesions attempted 1 [ ] 2 [ ] 3 [ ] 4 or more [ ]

Number of catheters used [ ]

Were there any problems during the procedure? Yes [ ] No [ ]

Please specify
.....

Post PTCA hospital outcome

(including complications)

		Yes	No
1	Success and routine discharge	[ ]	[ ]
2	Angina requiring repeat PTCA	[ ]	[ ]
3	Haemorrhage requiring transfusion	[ ]	[ ]
4	CABG - emergency	[ ]	[ ]
5	CABG - elective	[ ]	[ ]
6	Myocardial infarction	[ ]	[ ]
7	Death	[ ]	[ ]

Other complications (specify) .....

QUESTIONNAIRE 4

FREEMAN HOSPITAL CASPE OUTCOMES STUDY

3 - 12 MONTH ANGIOPLASTY (PTCA) QUESTIONNAIRE

PATIENTS NAME ..... DATE [/ /]
Address.....
.....
Hospital No.....

Drug Treatment

Could you indicate your current drug treatment and dose.

	DRUG	DOSAGE	FREQUENCY
1.			per day
2.			per day
3.			per day
4.			per day
5.			per day

HOSPITAL RE-ADMISSION

	YES	NO
Have you been seen at a hospital Accident & Emergency (A & E) Department since your operation?	[ ]	[ ]

If yes, why?.....

Have you been re-admitted to hospital since your operation?	[ ]	[ ]

If yes, please name the hospital.....

and date.....

If yes, what was the reason for your admission?

Chest pain	[ ]
Heart attack	[ ]
Repeat angioplasty	[ ]
Coronary artery surgery (CABG)	[ ]

Other specify.....

Appendix 7 (b) Angioplasty report - December 1990

INTRODUCTION

This is the second report of the angioplasty outcomes study. It includes data collected over an 18 month period, beginning in January 1989. The total number of patients included in the report is 189, who have all reached at least the 3 month follow-up stage.

The outcome indicators are, as reported in the first report:

- Angiographic success - vessels dilated.
- Rate of CABG.
- Survival.
- No readmissions.
- Decreased angina score at follow-up.
- Increased walking distance at follow-up.
- Improvement in general health status (Nottingham Health Profile).
- Decreased medication at follow-up.

PRESENTING CHARACTERISTICS

1/ Age & sex distribution

The sample consists of male and female patients distributed in the following proportions:
male = 76.5% female = 23.5%
The mean age is 53 years, with a range of 28 - 74 years.

2/ Type of angina

The patients were classified as suffering from either Stable or Unstable Angina Pectoris.
- stable = 65%
- unstable = 35%

3/ Vessels attempted & stenosis severity

Table 1.1 - Nos. of vessels stenosed and severity of stenoses (no.patients=189).

STENOSIS SEVERITY	LAD	LMAIN	RCA	CIRCX	DIAG	OBTUSE
A = 50-70%	14	1	11	9	1	2
B = 71-90%	40	0	12	13	2	0
C > 90%	61	0	33	21	6	1

* There was one vein graft reported to be stenosed >90%.

The most commonly stenosed vessels are the LAD artery, followed by the RCA and then the Circumflex artery.

Is this as expected?

4/ Class of angina

Angina has been classified according to the Canadian Angina Classification system on a scale of 1 to 4 (1= least severe and 4 = most severe).
The percentages of patients in these classes are:

Class 1 - 3%
Class 2 - 26%
Class 3 - 13%
Class 4 - 33%

The remaining 26% were not able to be classified as they did not successfully complete the pain questionnaire.

5/ Medication

Medication was assessed using a simple scale of 1 to 6 (1= single therapy and 5[& 6] = triple therapy). The percentage of patients on triple therapy for angina was 44%.

RESULTS

1/ IMMEDIATE OUTCOMES

Table 2.1

OUTCOME CATEGORY	STABLE/ DIL.	UNST/ DIL.	%	STABLE/ FAIL	UNSTA /FAIL	%
DIED IN HOSP.	0	1	0.6	1	0	3.8
CABG IN HOSP.	0	1	0.6	4	5	35
MI IN HOSP.	0	1	0.6	0	0	0.0
REDO IN HOSP.	0	1	0.6	0	0	0.0

The total number of patients = 189 (163=successful; 26=unsuccessful).

-There was an 88% initial success rate, or dilation rate (89% for stable patients and 86% for unstable patients).

-When the number of stenosed vessels is observed with the number of vessels that are dilated and failed, there are a higher proportion of failures among the double and triple vessel groups - is this as expected?

Table 2.2

OUTCOME	NO.VESS.=1	NO.VESS.=2	NO.VESS.=3	TOTAL NOS.
DILATED	92%	82%	71%	134
FAILED	8%	18%	29%	19

When the vessels attempted are observed with the rate of success and failure, there do not appear to be significant differences. See table 2.3:

Table 2.3 - No. vessels dilated/failed.

VESSEL	DILATED NOS.	FAILED NOS.	% FAILED
LAD	88	12	12%
RCA	38	10	21%
DIAG	7	2	22%
LMAIN	1	0	0%
OBTUSE	4	1	20%
CIRCX	32	5	14%

Age & Sex

There is a difference in the mean ages of patients whose procedure was either successful or unsuccessful:

- dilated = 50.3 years
- failed = 61.7 years.

There do not appear to be any real differences between the proportion of females in the samples:

- dilated = 25% female.
- failed = 21% female.

2/ OUTCOMES TO 3 MONTHS

Table 3.1 - Outcomes from discharge to 3 months(no.=189)

OUTCOME CATEGORY	STAB/ DIL.	UNST/ DIL.	% ($)	STABL/FAIL	UNSTA/ FAIL	% (#)
OK TO 3m	90	43	81.6	6	2	30.8
DIED TO 3m	0	1	0.6	0	0	0.0
CABG TO 3m	0	2	1.2	3	2	19.2
PTCA TO 3m	5	2	4.3	0	0	0.0
READMIT 3m	9	4	8.0	0	0	0.0

$ patients successful = 163; # -patients unsuccessful = 26.

a/ Readmissions

Total readmissions in the successful group(including those readmitted with major cardiac events: CABG, MI, PTCA) = 19/163 = 11.6%.

Other reasons for admission were reported as:

- gastroscopy.
- kidney transplant.

- gastric reflux.
- chest pain / rpt.angiograms.

b/ PTCA to 3m

The 7 patients who underwent a repeat PTCA to 3 months had a repeat procedure to the original vessel - which was successful in all cases on both occasions.

c/Stenosis severity & vessels

A simple stenosis score was given to the vessels. This was calculated by:

- 1 point for <70%
- 2 points for 70-90%
- 3 points for >90%.....

This was then summed across all"stenosed"vessels.

Tab3.2

PATIENT OUTCOME	DILATED AV.STEN.	FAILED AV.STEN.
STAB./CABG/PTCA	2.2	3.6
STAB. -OK TO 3m	2.6	3.8
UNSTAB. -DIED	5.5	/
UNSTAB. -CABG/PTCA	3.1	2.8
UNSTAB. -OK TO 3m	3.0	3.5
STAB. -DIED	/	5.0

Those who died have the highest stenosis scores. Also, those who are successfully dilated, had lower scores than those who failed, although numbers are small.

d/ Type of angina

The general picture is positive. Many more patients show an improvement in angina score from admission to 3 m.-see table 3.3.

Table 3.3 - %IV at presentation and at 3/12 follow-up

PATIENT TYPE	BETTER (%)	SAME (%)	WORSE (%)	%IV BEFORE	%IV AFTER
STABLE/DIL.	52.6	42.1	5.3	33.3%	14.0%
UNSTABLE/DIL.	46.7	33.3	20.0	44.4%	20.0%
CABG	62.5	25.0	12.5	45.5%	13.0%
READMITTED	50.0	33.3	16.7	44.4%	17.0%
FAILED PTCA	25.0	50.0	25.0	40.0%	25.0%

There appears to be a marked difference between the incidence of class IV angina in the stable and unstable groups (33.3% compared to 44.4%), and both groups show a reduction to 3 months. The pain dimension of the Nottingham Health Profile supports the reduction in the percentage of patients on Triple therapy before and after PTCA (for both stable and unstable patients):

	mean NHP pain score before	at 3m
Stable	26.33	11.52
Unstable	25.56	15.26

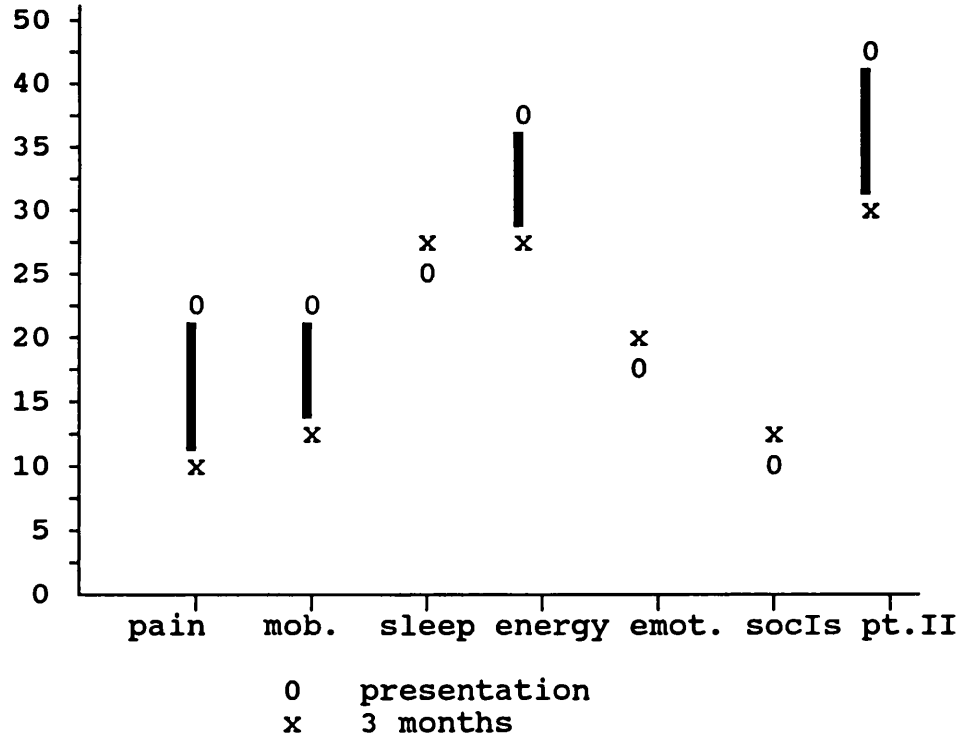
e/ Medication
Table 3.5

PATIENT TYPE	BETTER (%)	SAME (%)	WORSE (%)	% Tx BEFORE	% Tx AT 3m
STABLE/DIL.	36.3	42.5	21.3	42.9	47.5
UNSTABLE/DIL.	16.7	45.8	37.5	26.7	52.5
CABG	58.3	25.0	16.7	41.2	33.3
READMITTED	45.5	36.4	18.2	45.5	63.6
FAILED PTCA	20.0	80.0	0	62.5	100.0

The percentages of patients on Triple Therapy, in general, have increased from presentation to 3 month follow-up - especially for the unstable, the readmitted and the failed groups (this has been necessary to produce an improvement in angina scores - see table 3.4). It is interesting to note that the stable dilated group (whose PTCA was successful) do not improve the potency of medication afterwards.

f/ General well-being

Figure 3.1 - Mean NHP scores before & at 3m after PTCA



The Nottingham Health Profile has been used as a measure of general well-being. Each item is scored from 0 to 100, with 0 being the most healthy score and 100 being the least healthy score.

As can be seen from Figure 3.1, there is a general improvement in well-being (reduction in NHP scores) over the follow-up period(this is for all patients -successes and failures). Sleep and social isolation show slight rises in NHP score, but neither appear to be significant.

g/ Activity levels

The percentage of patients that report an increase in activity level to 3 months is 47%. 38% of patients report a decrease in activity level from presentation to 3 months. This is confirmed by the mobility dimension of the NHP:

At admission - mean score = 21
At 3m - mean score = 13.89

3/ OUTCOMES TO 12 MONTHS (Patients who have completed 12m).

Table 4.1 - Outcomes from presentation to 12/12 -no.=99

OUTCOME CATEGORY	STAB/ DIL.	UNST/ DIL.	% ($)	STAB/ FAIL	UNST/ FAIL	% (#)
NON-RESP.	1	1	2.3	0	0	0.0
DIED-HOSP.	0	0	1.1	1	0	8.3
CABG-HOSP.	0	0	0.0	4	1	41.7
DIED-12/12	0	1	1.1	0	0	0.0
OK - 12/12	30	11	47.0	0	1	8.3
CABG-12/12	3	3	8.0	1	2	25.0
MI-12/12	1	0	1.1	0	1	8.3
REDO-12/12	9	5	17.3	0	0	0.0
READ-12/12	10	10	22.9	1	0	8.3

$ -patients successful = 87; # - patients unsuccessful = 12.

a/ Readmissions to 12 months

In total, 48.4% of the successful patients who have reached 12 month follow-up have been readmitted to hospital during the year. Of these, 54% have had major cardiac events (26.4% of total that have reached 12 month follow-up) i.e. CABG, re-do PTCA, or MI. The other reasons reported for admission include:

- Chemotherapy catheter insertion.
- D & C.
- Chest infection.
- Virus infection.
- Several chest pains and repeat angiograms.

Altogether, 55% of the patients who are stable at presentation and who reach 12 month follow-up do not

report any cardiac events compared to 34% for the unstable group.

There are no new cases of death to 12 month follow-up, with all deaths occurring during the first 3 months after PTCA.

b/ PTCA to 12 months

Of those patients who underwent a second angioplasty to 12 months, the majority had a redo to the original vessel dilated (which was successful at original attempt). All redo angioplasties were successful.

c/ Angina Class

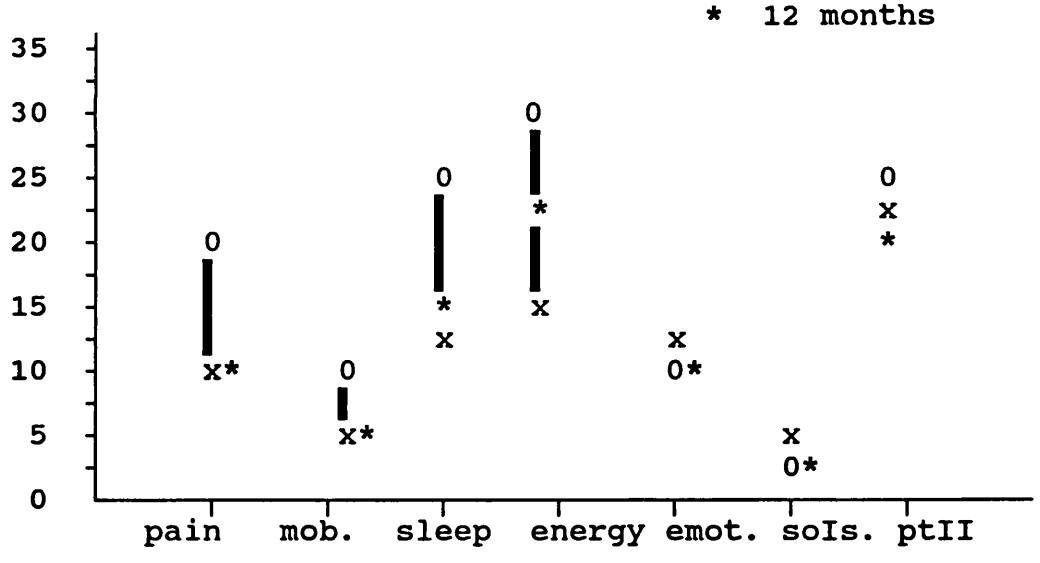
The percentage of patients who belong to class IV in the Canadian Angina Classification at 12 months remains fairly stable from 3 months and therefore is lower than the percentage of class IV patients at presentation. The mean NHP pain dimension scores show a decrease to 3m and in those reaching 12m follow-up, this improvement is maintained (see Figure 4.1).

d/ Medication

The percentage of patients who are taking Triple Therapy medication at 12 months decreases from 3 months, especially in the stable dilated group (47.5% at 3/12 to 24% at 12/12). Those who are readmitted to 12/12 also show a decrease in the potency of medication taken from 3/12, but the degree of change is not so large (63.6% to 50%). It therefore appears that the medication prescribed is not reduced significantly until after 3/12 from the PTCA and that those patients who are readmitted continue to take more potent medication.

e/ General Well-being

Figure 4.1 - Mean NHP scores at 0m, 3m & 12m for patients.



The mean NHP scores for those who have reached 12 month follow-up show a general improvement from presentation to 12 months, although energy, sleep and mobility scores have crept up from 3 months.

When the scores are broken down further into outcome categories to 12 months, the results can be summarised as follows:

CABG: - good improvements in pain, sleep and partII to 12 months, after initial worsening at 3 months.

MI: - only 2 patients. Much worse to 12 months.

Re-PTCA - big improvements in all categories to 12 months.

OK: - generally better than presentation, although some worsening from 3 months in most categories.

f/ **Activity levels**

45% of patients who reach 12 month follow-up report an improvement in activity level, while 33% report being less active than at admission.
The mean NHP scores of these patients for the mobility dimension confirm this:

At admission - mean score = 10.33
At 12 months - mean score = 8.45.

DISCUSSION

Two articles have recently been published which look at 12 month outcomes following PTCA. The first study reports on 100 consecutive patients to undergo PTCA at The Brompton (Glazier et al,1990). The other study followed up 406 stable and 202 unstable patients for a period of 36 months after PTCA, at Mainz, W.Germany (Rupprecht et al,1990).
The proportions of stable and unstable patients in the studies was as follows:

- CASPE/Freeman stab=65%, unstab=35%
- Glazier stab=84%, unstab=16%
- Rupprecht stab=67%, unstab=33%

All studies included a high proportion of patients with multivessel disease.

The incidence of death did not vary greatly across the studies. In our outcome study, there was a 1.2% death rate to 3 months (2 out of 3 deaths in hospital) and no further deaths to 12 months. Glazier reported 1 death in hospital and no more deaths at 12 months (1%), while Ruppecht

reported 0.5% death rate in hospital and a further 1.5% to 12 months (with a further 1.8% cardiac deaths to 36 months).

The incidence of CABG during the hospital period varied between the studies. Rupprecht reported an incidence of 3.5%, Glazier reported 4% and Freeman reported 0.6%(for those whose PTCA was successful.

The Freeman study was the only one to follow-up at 3 months -therefore no comparisons can be made at 3m.

To 12 months, the rate of CABG was the same at Freeman and The Brompton (8%) and Rupprecht reported a slightly lower rate of 6.9%.

Non-fatal MI was reported as less than 1% at Freeman to 3 months and only 1.1% of those reaching 12 month follow-up. This compares to 4% incidence in hospital and a further 1% at 12 months by Glazier and 3.3% and 1.2% respectively over the same periods by Rupprecht.

The re-angioplasty rate at 12 months varied between 11.5% (Rupprecht) to 15% (Glazier) to 17.3% (Freeman).

The Freeman/CASPE outcomes research is the only study of the three to attempt to measure patient well-being.

Are these results as one might expect?
Do the results compare to the other studies as expected?
Do they highlight any areas which have not been covered adequately?

References

Glazier J.J. et al - "Clinical outcome following balloon angioplasty in 100 consecutive patients with multivessel coronary artery disease"
J.Royal Coll. Phys. of London. Vol24#4.p292-294,Oct.1990.

Rupprecht H.J. et al - "Short and long-term outcome after PTCA in patients with stable and unstable angina" European Heart Journal Vol.11.p964-973,1990.

Appendix 7(c) Notes the meeting between CASPE/Cardiologists

Held on Thursday 28th February 1991

Present:

Mr. M. Bardsley
Mrs. J. Goodfellow
Dr. D. Reid
Mr. Wrightson
Dr. M. Griffiths
Mr. Hedley-Brown
Dr. Furniss
Dr. G. Sanders
Dr. P. Mounsey
Dr. G. Parry
Dr. M. Farrer
Dr. Bexton
Prof. Campbell

Introduction:

Dr. Reid described the history of PTCA procedures at the FRH. Mr. Bardsley then described the basic outcome indicators for successful angioplasty and presented the results of the project to 3 month follow-up and 12month follow-up.

Discussion:

It was asked if the readmissions included those admitted with major cardiac events eg. for CABG or MI. Dr. Reid said that he had checked those admitted to 3month follow-up and that the majority of chest pains reported, were infact non-cardiac in origin. He explained that by checking out these patient - reported events, he is attempting to "harden up" the outcomes.

The lack of a decrease in the potency of medication was commented on and the group agreed that it probably reflects the policy of FRH, to not reduce medication.

There was a general concern that the NHP may not be reliable over time and therefore the group were reassured that the test-retest validations showed good correlations.

Another concern was that because the weighting of scores is non-linear, those
who score higher have a greater potential for improvement than those who score low. However, Mr. Bardsley explained that although there is a slight regression to the mean, this cannot fully explain the large improvements seen in the majority of the patients.

Dr. Reid was asked his own feelings about the NHP - he replied that although it shows those patients who are not feeling better, it doesn't differentiate between those who are readmitted with real cardiac and non-cardiac pain. Therefore, he felt that hard outcomes are better. Mr. Bardsley replied that these "hard outcomes" describe the health service process and not the patient. The cardiologists felt that the NHP falls down as an outcome measure on those patients who have multiple problems and are not expected to do well, as they appear not to show an improvement, although the stenosis may be reduced. Dr. Sanders responded by saying that this may, in fact, be it's strong point - the patients perception of their illness being important. As an example she cited the knee replacement patients who do not progress as expected, after what is apparently a "good result". The orthopaedic surgeons are now looking at common characteristics in those who do not progress, which may possibly lead to a change in their patient selection criteria.

Dr. Reid was asked if he can pick out patients from those referred for PTCA who are likely not to do well. He felt that he could do so with some patients and not with others.

Although the time for discussion was limited, the response to the presentation of results was fairly positive, with several clinicians expressing a keen interst - especially in the use of the NHP.

Appendix 8(a) Data collection proforma Orthopaedics

(Patient Label Here Please)

Height_____ins/cm Weight_____lbs/kg

INDICATION FOR SURGERY (Tick only one):

Osteoarthritis.....	[ ]	Trauma (fracture).....	[ ]
Rheumatoid arthritis.....	[ ]	Ankylosing spondilitis.....	[ ]
Avascular necrosis.....	[ ]	Post-infection arthritis...	[ ]
Congenital dislocation...	[ ]	Psoriatic arthritis.....	[ ]
Traumatic arthritis.....	[ ]	Other.....	[ ]

PROPOSED SURGERY:

	New	OR	Revision	ONLY IF REVISION No. of Oper	Reason Loose	Infect	Tech	Other	Previous Operation This Team	Currant Implant
Knee L	[ ]		[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	_____
Knee R	[ ]		[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	_____
Hip L	[ ]		[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	_____
Hip R	[ ]		[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	_____

PREVIOUS SURGERY: Write the year of the most recent operation in any appropriate slot.

	Hip L	Hip R	Knee L	Knee R	Ankle L	Ankle R	Foot L	Foot R
Join Replace								
Arthrod								
Other								

SIGNIFICANT CURRENT/CHRONIC DISEASE: (Tick where appropriate)

Arthritis Upper Limb: Right [ ] Left [ ]
Arthritis Lower Limb: Right [ ] Left [ ]

Spine: Cervical [ ] Thoracic [ ] Lumbar [ ]

Write in the codes for up to 5 current/chronic diseases from the list below.

[ ] [ ] [ ] [ ] [ ]

1	Peripheral Vascular Disease	11	Peptic Ulcer
2	Myocardial Infarction	12	Hiatus Hernia
3	Hypertension	13	Lower GI
4	Angina	14	Diabetes
5	Stroke	15	Renal Failure
6	TIA	16	Recurrent UTI
7	Neurological Disorder	17	Prostatism
8	Asthma		
9	COAD		
10	Malignancy	99	Other.....

FREEMAN /CASPE JOINT REPLACEMENT STUDY (2 of 2)

(Patient Label Here Please)

Post-operative form (attached to pink case-mix form)

SIGNIFICANT COMPLICATIONS DURING THE OPERATION

MAIN OPERATION:

	YES	NO	
Surgical	[ ]	[ ]	If yes, specify .....
Anaesthetic	[ ]	[ ]	If yes, specify .....

SECONDARY OPERATION 1:

	YES	NO	
Surgical	[ ]	[ ]	If yes, specify .....
Anaesthetic	[ ]	[ ]	If yes, specify .....

SECONDARY OPERATION 2:

	YES	NO	
Surgical	[ ]	[ ]	If yes, specify .....
Anaesthetic	[ ]	[ ]	If yes, specify .....

SECONDARY OPERATION 3:

	YES	NO	
Surgical	[ ]	[ ]	If yes, specify .....
Anaesthetic	[ ]	[ ]	If yes, specify .....

SIGNIFICANT POST-OPERATIVE COMPLICATIONS:

Write in the codes for up to 5 post-op.complications from the list below.

[ ] [ ] [ ] [ ] [ ]

1 Wound infection
2 Post-op bleeding
3 D.V.T. / Pulmonary embolus
4 Respiratory infection/complication
5 Cardiovascular complication (MI)
6 CNS complication (stroke)
7 Urinary complication (retention/infection)
8 Septicaemia
9 Renal failure

99 Other (specify).....

Significant = a condition that : prolongs the expected length of stay
: requires extra clinical resources (nursing, diagnostic tests, extra theatre time)

Completed by..... Date .....

Date of assessment/out-patient appointment .....

(Patient Label Here Please)

ARTHROPLASTY: Condylar [ ] Kinemax Cemented [ ]
Stabiliser [ ] Kinemax Uncemented [ ]

SIDE: Left [ ] Right [ ]

	Pre-Op.	3 month Post-Op.	12 month Post-Op.
WALKING PAIN			
No pain at any time	[ ]	[ ]	[ ]
No pain on walking	[ ]	[ ]	[ ]
Mild pain on walking	[ ]	[ ]	[ ]
Moderate pain on walking	[ ]	[ ]	[ ]
Severe pain on walking	[ ]	[ ]	[ ]
REST PAIN			
No pain at rest	[ ]	[ ]	[ ]
Mild pain at rest	[ ]	[ ]	[ ]
Moderate pain at rest	[ ]	[ ]	[ ]
Severe pain at rest	[ ]	[ ]	[ ]
CLIMB STAIRS			
Without support	[ ]	[ ]	[ ]
With support	[ ]	[ ]	[ ]
Cannot climb stairs	[ ]	[ ]	[ ]
TRANSFERS			
Without support	[ ]	[ ]	[ ]
With support	[ ]	[ ]	[ ]
Cannot transfer	[ ]	[ ]	[ ]
FLEXION DEFORMITY			
No deformity	[ ]	[ ]	[ ]
< 5 degrees	[ ]	[ ]	[ ]
5 - 10 degrees	[ ]	[ ]	[ ]
> 10 degrees	[ ]	[ ]	[ ]
INSTABILITY			
None	[ ]	[ ]	[ ]
Mild (0-5 degrees)	[ ]	[ ]	[ ]
Moderate (5-15 degrees)	[ ]	[ ]	[ ]
Severe (> 15 degrees)	[ ]	[ ]	[ ]
EXTENSION LAG	[ ]	[ ]	[ ]
5 degrees	[ ]	[ ]	[ ]
10 degrees	[ ]	[ ]	[ ]
15 degrees	[ ]	[ ]	[ ]

	Pre-Op	3month Post-Op	12month Post-Op
ASSESSMENT DATE:	.....	.....	.....
WALK/STAND FUNCTION			
Walk and stand unlimited	[ ]	[ ]	[ ]
Walk 400 yards,stand 1/2hr	[ ]	[ ]	[ ]
Walk<400 yards,stand<1/2hr	[ ]	[ ]	[ ]
Walk < 100 yards	[ ]	[ ]	[ ]
Cannot walk	[ ]	[ ]	[ ]
MUSCLE STRENGTH			
Excellent-cannot break quadriceps	[ ]	[ ]	[ ]
Fair-moves through arc of motion	[ ]	[ ]	[ ]
Poor-cannot move through arc	[ ]	[ ]	[ ]
RANGE OF MOTION:DEGREES OF ARC			
	.....	.....	.....
VALGUS DEGREES OVER NORMAL	.....	.....	.....
VARUS DEGREES OVER NORMAL	.....	.....	.....
ASSISTED BY: One stick	[ ]	[ ]	[ ]
One crutch	[ ]	[ ]	[ ]
Two crutches	[ ]	[ ]	[ ]
Unable to use crutches	[ ]	[ ]	[ ]
PATELLO-FEMORAL CREPITUS			
(one only) Audible	[ ]	[ ]	[ ]
Palpable	[ ]	[ ]	[ ]
Absent	[ ]	[ ]	[ ]
SITE OF PAIN Global	[ ]	[ ]	[ ]
(can be Medial	[ ]	[ ]	[ ]
multiple) Lateral	[ ]	[ ]	[ ]
Posterior	[ ]	[ ]	[ ]
Anterior	[ ]	[ ]	[ ]

FREEMAN/CASPE JOINT REPLACEMENT STUDY KNEE X - RAY EVALUTIONS

(Patient Label Here Please)

PRE-OPERATIVE X - RAY EVALUATION: DATE ..... SIDE: Left [ ]

Right[ ]

Tibio-femoral angle:.....

MARK Y IN ANY SLOT WHERE CONDITION IS PRESENT

	Sclerosis	Osteophyte	Collapse	Cyst
Medial femoral				
Medial tibial				
Lateral femoral				
Lateral tibial				
Patello-femoral				

****

POST-OPERATIVE X - RAY EVALUATION: 3 - month Post-OP 12- month Post-Op

DATE: ..../..../.... ..../..../....

Standing tibio-femoral alignment: ....

Tibial prosthesis roll back: ....

Tibial radiolucent line:	1mm	2mm	>2mm	1mm	2mm	>2mm
Zone 1	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 2	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 3	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 4	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 5	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 6	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 7	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

Femoral radiolucent line:	1mm	2mm	>2mm	1mm	2mm	>2mm
Zone 1	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 2	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 3	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 4	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Zone 5	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

Anterior Flange: (one only)	Proud	[ ]	[ ]
	Flush	[ ]	[ ]
	Cutback	[ ]	[ ]

Patello-femoral Joint position (one only)	Neutral	[ ]	[ ]
	Neutrolateral	[ ]	[ ]
	Lateral	[ ]	[ ]
	Subluxed	[ ]	[ ]
	Dislocated	[ ]	[ ]
Patello-femoral Joint condition	Normal	[ ]	[ ]
	Fragmented	[ ]	[ ]

Appendix 8(b) Orthopaedics Report August 1990

Basic data

Presenting characteristics. Table 1.

	OA	RA	Others(Unk)
No. cases	85	67	36
% Male	48%	13%	19%
% revisions:	5%	12%	-
% arthritis			
2 sites:	15%	34%	-
3 sites:	5%	49%	-
Age<60	16%	38%	50%
60-69	36%	38%	28%
>69	48%	23%	21%
Co-morbidities:			
None:	44%	45%	97%
One :	29%	15%	3%
>one:	26%	15%	-
Post-op : complications	12%	7%	-

Most common co-morbidities:

11 27 cases Peptic ulcer
3 26 cases Hypertension
12 11 cases Hiatus hernia
4 8 cases Angina
2 6 cases Previous MI
14 6 cases Diabetes

Common post-Op complications

3 cases Post-op bleeding
2 cases CV comp/MI
1 case Periph. ulcer
Oesaphagitis
Cellulititis
Haematoma
Urine retention
DVT/PE

Slow knee flexion
Spine-fell out of bed
Dyspepsia known DV
Surg MUA slow mobil

Knee score

*There is a significant improvement in Knee scores to three months after the operation. OA and RA cases responding in a similar way.

*Men have slightly higher scores than women - both before and after the operation.

*The younger age group (under 60) do slightly less well.

*The presence of a single co-morbidity does not appear to make much difference to the scores, however the group of patients with more than one co-morbidity fare less well.

*Post-operative complications do not appear to be affaecting the scores at three months

		Pre-score	Post-score
All	122	40.3 (1.5)	72.9 (2.6)
OA	83	43.2	71.6
RA	64	40.9	71.4
Oth.	22	30.7	78.6
Men		47	74
Women		37	67
Age<60		38.7	63.2
60-70		35.4	78.4
>70		45.5	67.0
no co-morbid.		40.0	73.2
one		42.0	83.7
two		40.9	59.8
Post-op complications:			
Yes		43.7	73.0
No		40.7	70.7

NHP Scores:

*As with the knee score there are signifcant improvements on the NHP - on all dimensions - including the 'social ones'. As you might expect pain shows the best improvement followed by mobility and energy.

*There are no enormous difference in the relative improvements of OA and RA patients

*As before there are some age effects - the 60-70 band generally doing better than to inder 60's or over 70's. The reasons are still mysterious.

*There are some interesting and quite marked sex differences. Men tend to score lower pre-operatively and post-operatively.

	NHP scores	
	Before	+3 mnths
	---	---
Energy	43.0	28.0
Pain	63.4	31.6
Emotion R	23.3	12.7
Sleep	50.6	39.9
Soc Isol	13.9	8.0
Mobility	46.7	31.6
Part II	41.7	30.2

All sig <p.01

Changes in NHP score from pre-op to three months later. (larger positive values=greater improvement)

	NHP scores		Age groups		
	OA	RA	<60	60-70	>70
	---	---	---	---	---
Energy	9.8	20.1	17.7	13.5	18.4
Pain	35.8	30.7	27.6	47.3	29.4
Emotion R	8.4	14.0	13.2	12.1	7.3
Sleep	12.5	11.6	5.8	17.0	6.3
Soc Isol	9.6	3.4	-3.6	7.8	8.5
Mobility	16.3	11.0	6.1	22.3	7.3
Part II	17.7	3.6	13.5	16.1	9.0

Pre op NHP Scores	SEX		
		F	M
Energy	40.0	51.9	34.6
Pain	61.6	67.0	57.3
Emot Reacs	21.2	30.7	22.3
Sleep	49.7	49.6	44.2
Soc Isol	12.4	16.4	13.7
Mobility	48.0	53.7	38.8
Part II	38.3	44.0	45.5

Post op NHP Scores	SEX		
		F	M
Energy	26.6	33.0	16.6
Pain	36.0	31.6	23.4
Emot r	13.0	14.2	7.3
Sleep	40.9	40.8	21.6
Soc Isol	7.2	10.1	2.8
Mobil	30.1	38.0	24.9
Pt II	30.1	32.7	23.6

Postop preop	SEX		
		F	M
Energy	13.07	19.07	10.15
Pain	25.27	36.79	34.05
Emot R	10.01	12.27	7.99
Sleep	11.38	5.31	21.43
Soc Isol	7.70	3.24	7.77
Mobil	18.63	13.09	12.23
Pt II	9.03	9.87	20.32

	SCORE1	SCORE2	DSCORE
	Mean	Mean	Mean
SEX			
	42	79	35.84
F	37	67	31.93
M	47	74	28.27

Additional questions.

Does Pre-op Xray evaluation reflect lower Knee scores? No.

The preop Xray evaluation was scored simply by adding up the number of reported abnormalities by site. With 4 posssible abnormalities (sclerosis,osteophyte,collapse,cyst) and 5 sites gives a maximum of 20. This scoring may be nonsense. There appears to be no clear relationship between the extent of damage (recorded in this way) and the presenting knee score.

Summaries of SCORE1 By levels of PREOPX

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			40.6982	16.8805	169
No evaluation (assume ok)			32.9615	21.5471	26
No. abnormaliites	0		43.3030	17.1177	66
	1		41.0000	0.0	1
	2		36.0000	18.3848	2
	3		39.0000	12.2882	3
	4		44.3750	9.5350	16
	5		41.5714	15.9881	7
	6		43.0000	14.4827	9
	7		41.0000	9.7234	12
	8		37.1111	22.6740	9
	9		45.0000	10.1653	4
	10		55.2500	13.3010	4
	11		29.6667	15.8850	3
	12		29.0000	16.4924	5

(No visible relationship to outcome)

Does post-op Xray evaluation (translucent zones) reflect lower knee scores? No.

The post-op x-ray was scored for the number of translucent zones - dies this make senses - and the idea tested that higher scores would result in lower post-operative knee scores and worse outcomes (improvements on knee score).
There is no observable relationship (at this stage).

Summaries of Improvement in Knee score.
By levels of POSTOPX

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			32.7131	33.2259	122
Not available assume ok			3.2381	37.6522	21
0	Zones		40.6800	29.5818	75
1	Zone		27.1250	23.1235	8
2			43.5714	36.4274	7
3			17.0000	11.6905	4
4			28.0000	0.0	1
5			45.0000	27.7849	3
6			26.0000	1.4142	2
7			67.0000	0.0	1

Appendix 8(c) Notes of Outcome meeting- Jan 1991 Orthopaedics

FREEMAN/CASPE ORTHOPAEDIC MEETING

Present: Steve Smith
Ian Pinder
Martin Bardsley
Donna Swinden
Gill Sanders

1.Data Collection.

MB informed SS and IP of the current CASPE position as regarding funding. Both were keen to continue data collection in the meantime. DS informed IP that some annual review forms had not been completed, and IP agreed to complete them retrospectively from the notes/x rays whilst DS would post out NHPs to the respective patients. Apart from that, data collection was going well with 300 patients in the study, 77 of which had reached the Annual Review stage.

2.Mr. Len. Fenwick's comments at the recent Steering Group Meeting were passed on to IP and SS by GS. These comments regarded presenting some hard data to colleagues IP and SS would like to see some recent results before going any further. SS mentioned that Frank Sibley had written a paper on the outcomes of 30 knee replacement patients. Mr. Sibley will be writing to MB in the near future. SS will get a copy of his paper for CASPE. SS and IP suggest writing a paper when 100 annual reviews have been completed.

IP wants to talk about outcomes of knee replacements at the surgical sub-group of the next British Society of Rheumatology Meeting.

3. MB mentioned that validations of the knee score forms and the NHP forms had been carried out using notes and interviews as comparisons. MB did not have the results with him but there seemed to be a reasonable match.

4.Results - The August Report.
MB talked through the report. The fact that younger patients do slightly less well on the knee score was picked out by SS and IP. They both showed concern over this group and would like to investigate it further. It was suggested that MB looks at different age groups within and between RH and OA patients. MB said he would identify those younger patients that were not responding as well and we would look at them in further detail. SS suggested that just one or two patients could drag the whole group down. He mentioned litigations, psychiatric illness, hypochondriasis and unrealistic expectation as possible causes of non-response. IP and SS were interested to compare the relative improvements of RA and OA patients at annual review. MB pointed out that knee

scores correspond with NHP
scores. IP and SS would like this demonstrated with greater clarity.

There was further speculation over the differences in knee scores between men and women, with little conclusion.

MB had several questions about the X ray data collected. IP confessed that it was problematic to interpret. The pre-op X ray indicators are not discriminative enough. IP wondered whether large tib-femoral angles equate with low knee scores. MB to have a look.

SS mentioned that post-op X ray indicators such as the radiolucent lines were only significant when two observations were compared. Therefore needs to be a longitudinal study. MB then drew attention to the case histories of those patients who showed poor improvement on NHP, knee score or had suffered post-operative complications. SS and IP found that the main value of this exercise was that it showed that Deep Vein Thromboses are not a big problem in knee replacement surgery.

5. MB talked through his NHP graphs. SS and IP would be interested to see how the knee replacement results compare at annual review. It was agreed to produce a new report when 100 annual reviews have been completed and for a discussion to take place in light of that.

Appendix 8(d) - RHEUMATOLOGY PROJECT

RHEUMATOLOGY FOLLOW-UP QUESTIONNAIRE

The questions on this page refer to the state of your health following your admission to the Freeman Hospital approximately one year ago.

(For questions 1 & 2 please tick appropriate box)

1. Do you feel your stay in hospital improved the state of your health in general?

Yes	[ ]
No	[ ]
I don't know	[ ]

2. Do you feel your stay in hospital improved your arthritis?

Yes	[ ]
No	[ ]
I don't know	[ ]

3. For the following questions please circle the number on the scale which most accurately reflects your present state:
(0 = very much better than one year ago, 6 = very much worse than one year ago).

a) **How well controlled do you feel your arthritis is now?**

Much better than one year ago 0 1 2 3 4 5 6 Much worse than one year ago

b) **How often have you felt the level of pain you experience has been unacceptably high?**

Much less often than one year ago 0 1 2 3 4 5 6 Much more often than one year ago

c) **How well do you feel you understand your arthritis now?**

Much more than one year ago 0 1 2 3 4 5 6 Much less than one year ago.

PLEASE TURN OVER

4. What do you see as your main problems due to your arthritis? (Please state in order of severity).

1)

2)

3)

5. Do you have any health problems, other than those due to your arthritis? (Please state in order of severity).

1)

2)

3)

6) Which of the following social services do you receive? (Tick any that apply).

Home help	[ ]	District nurse	[ ]
Meals on Wheels	[ ]	Day centre	[ ]
Bath attendant	[ ]		

Other (please specify) .....

Thank you for your help.