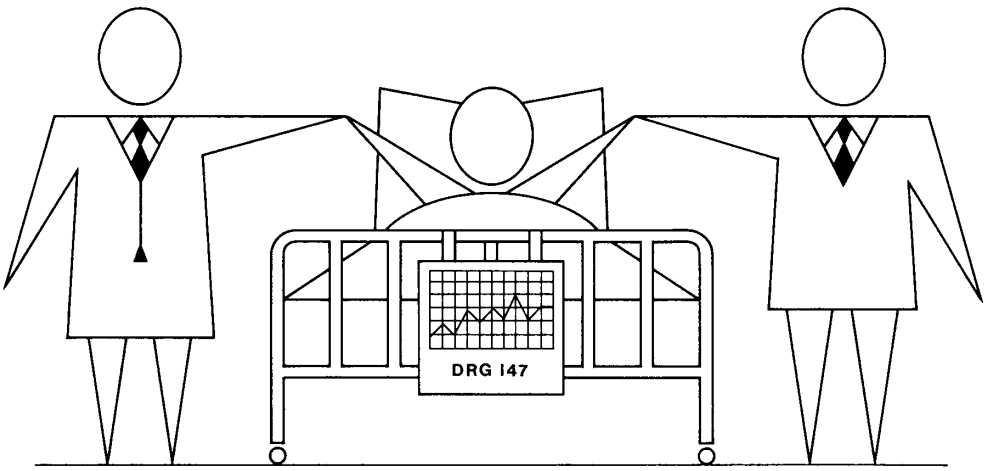
DRGs and health care

The management of case mix



Edited by
Martin Bardsley, James Coles and Linda Jenkins

King Edward's Hospital Fund for London

Major shifts in the role of health service managers have led them no longer merely to respond to central government or professional pressures. They are now expected to take the lead in determining the shape of the health service by meeting local needs and making efficient use of resources.

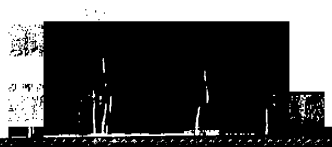
The technological revolution in information retrieval has offered managers increasingly sophisticated systems to support their day to day decisions. Simultaneously the development of diagnosis related groups (DRGs) in the late 70s has provided an effective means of handling case mix.

DRGs classify patients into types that are similar both clinically and in the resources they consume. Because of these unique features DRGs have been chosen by many health services, including the NHS in its resource management approach, as a robust and workable classification of hospital inpatients.

This book is about practical issues of case mix management. It reviews the current applications of DRGs at hospital level and examines their use in policy matters at all levels of health services. It is for everyone interested in and responsible for the way hospital resources are used, such as general managers, hospital doctors, nurses, health service researchers and specialists in finance, planning and information.

The editors, who have also contributed chapters, have been actively involved in research into planning services with clinical budget holders in the NHS, and have acquired a considerable depth of knowledge and experience in the use of DRGs in the UK and elsewhere. The introduction is by Professor Robert Fetter, one of the developers of the DRG classification.

ISBN 0 19 724640 0



DRGs
and health care
The management of case mix

King Edward's Hospital Fund for London

Patron: Her Majesty The Queen

President: HRH The Prince of Wales

Treasurer: Robin Dent
Chairman of the Management Committee: The Hon Hugh Astor JP
Secretary: Robert J Maxwell JP PhD

King Edward's Hospital Fund for London is an independent foundation, established in 1897 and incorporated by Act of Parliament 1907, and is a registered charity. It seeks to encourage good practice and innovation in the management of health care by research, experiment and education, and by direct grants.

Appeals for these purposes continue to increase.

The Treasurer would welcome any new sources of money, in the form of donations, deeds of covenant or legacies, which would enable the Fund to augment its activities.

Requests for the annual report, which includes a financial statement, lists of all grants and other information, should be addressed to the Secretary, King Edward's Hospital Fund for London, 14 Palace Court, London W2 4HT.

DRGs
and health care

The management of case mix

EDITED BY
Martin Bardsley James Coles Linda Jenkins

King Edward's Hospital Fund for London

Typeset by Tradespools Ltd
Printed in England by Hollen Street Press
Distributed for the King's Fund by
Oxford University Press

ISBN 0 19 724640 0

King's Fund Publishing Office
2 St Andrew's Place
London NW1 4LB

CONTENTS

Tables

Figures

ACKNOWLEDGEMENTS

The editors gratefully acknowledge the willing cooperation of the authors and their observance of seemingly impossible deadlines. They also wish to thank the King's Fund for publishing the book, the Research Management Division, DHSS, who funded the UK DRG research team, and Charlotte Rooney for typing numerous scripts.

Thanks also to Robert B Fetter, John Thompson, and Dr James Woolliscroft for their helpful comments on earlier versions of chapter 2; and Annette Schmidt for manuscript preparation.

Finally, thanks to Roger Beech, John Brazier, Marian Craig, Walter Holland, Francesco Taroni and officers in the DHSS for ideas, comment and criticism; to Marian Craig, Tony Swan, Kate Richmond and to the DHSS and the Special Trustees of St Thomas' Hospital for financial support for the work described in Chapter 8.

THE CONTRIBUTORS

Martin Bardsley Research analyst, CASPE Research

Gwyn Bevan Senior Lecturer in Health Economics, Department of Community Medicine, UMDS, St Thomas' Hospital

James Coles Associate Director, CASPE Research, and Fellow in Case Mix Accounting, King's Fund College

Robert B Fetter Professor of Health Care Management, Yale School of Organisation and Management

Linda Jenkins DRG Project Leader, CASPE Research

Laurence F McMahon Assistant Professor, Department of Internal Medicine, University of Michigan Medical Center

Antoinette B Newman Financial planner, Paddington and North Kensington HA

Jean-Marie Rodrigues Professor of Public Health, University of St Etienne

Hugh Sanderson Specialist in community medicine, Wessex RHA

Iden Wickings Director of CASPE Research and Deputy Secretary, King Edward's Hospital Fund for London

PREFACE

Why read this book?

The issue of case mix has for many years been recognised as important in the management of health services, yet it has also proved a source of irritation and tension between clinicians and administrators. Doctors have often felt constrained by their inability to command additional resources when more complex cases are admitted, while administrators have been frustrated in attempts to monitor the efficient use of resources without estimates of the case mix effect.

There have been major changes in recent years in both the technology of data processing and the role that the resulting information plays in management. It is still true to say though that management information in health services is generally limited – in helping either to understand the organisation or to monitor performance. The measurement of case mix on a routine basis is an inevitable development, not only as information but as a means of improving management. A simple classification of case types which can identify where available resources are deployed is important if the oft-quoted aims of efficiency and effectiveness in health care are to be achieved. Diagnosis related groups (DRGs) have been chosen by many health services as the most robust and workable classification of acute hospital inpatients currently available. DRGs, developed at Yale University, are the subject of many books, articles and papers in the United States where they have become the basis for hospital reimbursement within the federally funded Medicare programme. However, apart from these publications on the US experience an authoritative source of information on the application of DRGs in different health care settings has been sadly lacking. This book is the first outside the US to encompass the practical problems of DRGs, review their current applications and examine their potential in health services policy issues at a national level.

The book is aimed at all who have an interest in the way resources are used in the hospital setting, particularly those with some managerial responsibility. Among them will be general managers at regional, district and unit level, consultants, treasurers and heads of support departments. Health service planners and information specialists, as well as economists and research workers, are also likely to find something to interest them. While the prime focus is the British National Health Service it is clear that much will be of relevance to people in other countries throughout western Europe and in Australia who are developing their own use of DRGs. Chapter 4 gives a summary of their progress to date.

Given pressures on the time of senior managers as well as on clinicians, it is unlikely that many readers will wish to read the book from cover to cover, but will prefer to select individual chapters of particular interest to them. To this end, each chapter has been written to be read independently. Those with only an elementary knowledge of the subject would be advised to read Professor

Fetter's introduction and chapters 1 and 2 first in order to gain an insight into the concept of case mix and the construction of DRGs.

The book begins by putting DRGs in perspective, starting with an explanation of the need for case mix information that led to the development of DRGs. This is given in the introduction by Professor Fetter of the Yale School of Organisation and Management. DRG development was first prompted by the need to monitor the utilisation of services, but in times of high inflation in the health care sector in the US it was quickly seen as a tool for cost containment. Under Medicare's payment by DRG, a large proportion of US hospital stays are reimbursed at a price fixed by the case type. It is argued that such a fundamental change has led to better information, encouraging a new style of management which can address issues of efficiency and effectiveness at the level of patient type. These ideas are pursued in later chapters, in particular those in the section on DRGs in management.

Before choosing DRGs as a measure of case mix some consideration should be given to the requirements of a classification scheme. Chapter 1 examines the concept of case mix and the levels at which the hospital output might need to be described. A number of classification systems are available and their suitability and appropriateness depends on their intended application. An evaluation of a particular scheme cannot be made without reference to the way it is to be used. In particular, this chapter highlights the value of case mix measures in hospital resource management making groupings such as DRGs, that are defined as iso-resource groups, attractive.

The development of DRG definitions is described in chapter 2 by someone who was closely involved with the research group which refined the original scheme. The current classification is the result of several redesigns and refinements during a developmental period of 10–12 years. By involving doctors in the analysis of a great number of patient records the twin goals of clinical appropriateness and homogeneity of resource consumption within case types was achieved. The result was a classification based on a manageable number of mutually exclusive and exhaustive inpatient groups. As the scheme has been improved and refined in the past, so in future it is expected that DRGs will continue to evolve to cope with developments in clinical practice. The full DRG titles and diagrams of their coverage across areas of disease are given in Appendices I and II.

Much has been written about the controversial use of DRGs in funding hospitals for elderly patients insured by the US government's Medicare scheme. Chapter 3 highlights some of the problems of the scheme that would need to be addressed by other health care providers if they wished to adopt a similar strategy. It also suggests that the more comprehensive coverage of the New Jersey experiment in fixed price reimbursement was in some aspects better than the nationally adopted Medicare scheme.

In the last three to four years many other countries have looked to DRGs to help control rising costs. The section on DRG developments covers this expansion; for example, the problems being addressed in Europe, Canada and Australia are reviewed in chapter 4. Apart from the obstacles most countries face – a lack of a national standard patient abstract and compatible classification of diseases and operations – this chapter also covers the wide range of potential applications. DRGs are being seen as either tools at

national level for estimating global budgets, at regional level for allocating resources, or at hospital level for budgeting and performance review. Issues of performance include both quality assurance and utilisation review.

Focusing on the UK experience, the role of DRGs in the NHS is seen as complementary to the need for improved information about the resource consequences of alternative patterns of care. Chapter 5 outlines the practicalities of categorising patients by case type using the patient abstract available from a national sample of half a million cases. Some results are given, both of the ease of assigning cases to DRGs and of a comparison of teaching and non-teaching hospital workloads in DRG terms. It was found that in the UK, as in other countries, DRGs were not difficult to assign; the classification scheme was broadly acceptable for inpatients in the acute setting and offered a considerable advance on previous methods of accounting for case mix.

Since most health services are struggling with cash constraints, it is important to assess the expected resource consequences of a given mix of cases. Various methods of costing case types have been proposed, from detailed patient costing to cost allocation models. Chapter 6 considers the advantages and the disadvantages of different approaches and addresses the particular problems of allocating costs to DRGs in a system that does not record resource use at the level of individual patients.

The third section, DRGs in management, is in four chapters describing applications and opportunities for managing case mix. Chapter 7 sets the scene, describing the important role for case mix in raising the level of management debate and achieving a more output-oriented approach to the planning and provision of hospital services. The increased specificity of DRGs can give a better indication than simple specialty averages of those benefits that hospital management should be seeking to provide. On a more practical note the management of the inherent variability within case type is examined and increased use of variance analysis techniques is recommended to treasurers and managers alike.

Chapter 8 spells out how it is possible to use the most frequently seen case types as reliable indicators of performance. So long as sufficient cases fall in a group, the shape of the distribution of length of stay, or other variables, can be determined and appropriate confidence limits attached for use in performance review. Calculations of confidence limits for the top 50 DRGs which use 50 per cent of a hospital's resources suggest these would be suitable for selective comparisons of hospital performance.

Chapter 9 uses a case study to examine the value of describing clinical work by case mix groupings and the feasibility of costing patient types within a specialty. Working in a single specialty hospital the study used existing information systems, both manual and computerised, or developed new ones that would allow costing at the level of patient type. While it still requires considerable time and effort to achieve this, continually improving information systems make costing DRGs a feasible option.

Chapter 10 extends these applications by examining the value of case mix measures in the general management of health services, for example the use of case mix information in the context of clinical budgeting. It is suggested that they provide the basis for a language of health care management. This

language can be used as the framework for devolving budgetary responsibility to those committing the resources, with the aim of making overall improvements in health service outputs.

The final chapter attempts to bring together the separate contributions and proposes that the development of case mix measures will assist general managers but will also require them to address issues which have previously been found intractable. It points out that DRGs are not merely a system for reimbursing hospitals; their international adoption in health services with a range of funding mechanisms bears this out. The chapter debates the need for managers to develop a deeper understanding of clinical matters and also looks at the need to consider sectors of health care apart from acute inpatient work. General managers will find DRGs useful as descriptors of case mix. With costs attached, they can discuss the implications of changes in the level of provision or style of treatment with doctors, treasurers and other professionals. Methods of describing long term care patients and outpatients will soon be available for use in considering the broader issues of balancing appropriate care and efficient use of resources.

At present, DRGs offer the most complete and manageable classification of inpatient case mix available to health service managers. Using readily available information from the patient abstract, the groupings are easily determined and, on analysis, provide insight into many of the issues that managers should be addressing. Moreover, they offer a 'portable' vocabulary within and across national boundaries. The message of this book is that while improved systems will undoubtedly develop in due course, DRGs are available now. Managers should be willing to examine the benefits they offer.

INTRODUCTION

Robert B Fetter

In the last decade, national health care expenditures in most developed countries in the world have increased faster than the rate of inflation and in the United States now represent 10.5 per cent of the gross national product. The largest single component is for hospital care, which has accounted for an increasing share of total expenditures and has been a major contributor to the relative growth of the health care portion of the GNP[1]. Information on factors influencing hospital costs has therefore been critical to the management of health care institutions.

Hospital output and cost functions have been theoretically and empirically investigated in a variety of research settings[2-8]. The major limitation of all these studies is the method used to account for the multiproduct nature of the hospital. While there is little agreement on the definition of these products, there is a consensus that a hospital produces an extensive variety of them and that differences in product-line play an important role in understanding cost variations among institutions and among patients within an institution.

Diagnosis related groups (DRGs) are a system for describing the types of patients discharged from acute care hospitals. The current version of the groups contains 467 classes of patients, each defined in terms of one or more of the following variables: principal diagnosis, surgical procedures, additional diagnoses (comorbidities and complications), age, sex, and discharge disposition. The groups were designed to be clinically coherent in the sense that they are expected to evoke a set of clinical responses which result in a similar pattern of resource use[9]. Hence, the profile of services ordered by a physician is expected to be fairly similar for all patients treated in a given DRG.

Since the US federal government began paying a fixed price per DRG for providing services to Medicare patients in 1983, physicians and hospital managers are well aware of the implementation of DRGs as a payment mechanism. The original development of the groups, however, had nothing to do with prospective payment. In fact, the initial development of a patient classification scheme at Yale University began in the 1960s and was largely motivated by the needs of two utilisation review programs that were attempting to identify unusual cases with exceptionally long lengths of stay.

It appeared at the time that industrial control methods, commonly used by manufacturing firms, could be applied provided that the products of the hospital were identified. However, although product definition is often straightforward for a manufacturing firm, it is not so apparent for hospitals. One of the earliest attempts at defining the product of a hospital was that of Codman[10], who defined the products of the Massachusetts General Hospital in 1912 in terms of patients treated, students receiving medical instruction, nurses graduated, medical and surgical papers published, and important ideas

demonstrated. While the concept of the hospital product may not have changed substantially since Codman's time, there was clearly no consensus in the late 1960s on a useful operative definition.

Defining the concept of the hospital product

Chase and Aquilano[11] define a product as 'the output from a productive system offered for sale (in the case of a business) or otherwise made available (in the case of a governmental or philanthropic organisation) to some consumer'. In this context, the outputs of a hospital are the specific goods and services it provides to patients. The specific set provided to each patient is a 'product' of the hospital.

The development of DRGs initially began, then, as an attempt to define operationally the products of a hospital in terms of groups of patients receiving similar sets of outputs or services (such as laboratory tests, x-rays, nursing care). An approach was developed during this early research that was to be used in all future versions. Briefly, it divided all principal diagnosis codes into major diagnostic categories. These major categories were then partitioned into subgroups based on the values of variables associated with length of stay – the only utilisation measure available at that time. The development of the scheme is described more fully in Chapter 2.

Under the current version of DRGs, all principal diagnosis codes are condensed into twenty three major diagnostic categories (MDCs). The category to which a particular diagnosis is assigned is a function of the organ system it predominantly affects or the specialty which would typically provide care. Each hospital discharge is assigned to one and only one MDC based on its principal diagnosis code. In most MDCs, medical hospitalisations are then partitioned into clinically coherent groups of principal diagnoses while surgical hospitalisations are partitioned into groups of operating room procedures, referred to as procedure categories. There is an established hierarchy to these procedure categories based on intensity of resource use. Hence, a discharge with multiple operating room procedures is assigned to the most intensive category containing one of these procedures. Finally, both medical and surgical discharges may be further partitioned on the basis of age, the existence of substantial comorbidities and complications and discharge status. Some procedure categories are also partitioned on the basis of principal diagnosis (such as presence of malignancy). All these partitions were made using variables that were highly associated with resource use.

Implications for hospital management

The implementation of the DRG based payment system in the US represented the most significant change in Medicare policy since the programme began in 1965 and may have also created a profound change in the management style of acute care hospitals. Specifically, DRG based payment has encouraged administrators to view the utilisation and cost of hospital services along product lines.

In the context of these product lines, various aspects of production and operations management commonly employed by manufacturing firms –

product selection and design, quality control, and cost accounting – can be applied to hospitals for the purpose of increasing efficiency and quality of care. For example, the set of products which constitutes the business of each hospital can be used as the basis for a flexible budgeting and cost control system[12]. Each product is identified in terms of the treatment plan and set of services expected to be delivered to the patient. A patient hospitalised for acute appendicitis without peritonitis and without comorbidity problems might be expected to consume 12 meals, four days of hotel services, 16 hours of nursing care, 50 minutes of surgery, and so on. Each element would be costed so as to produce in each service-providing department or cost center the expected costs of this treatment. A budget would be an explosion of the hospital's forecasted mix of cases in terms of the components of each case type (product) and their cost.

As actual patient load became known, variance analysis would reveal the extent to which costs incurred in each cost center were above or below expected values. Causes would then be assigned based on the type of variance. In the hospital setting these sources of variance are changes in the following:

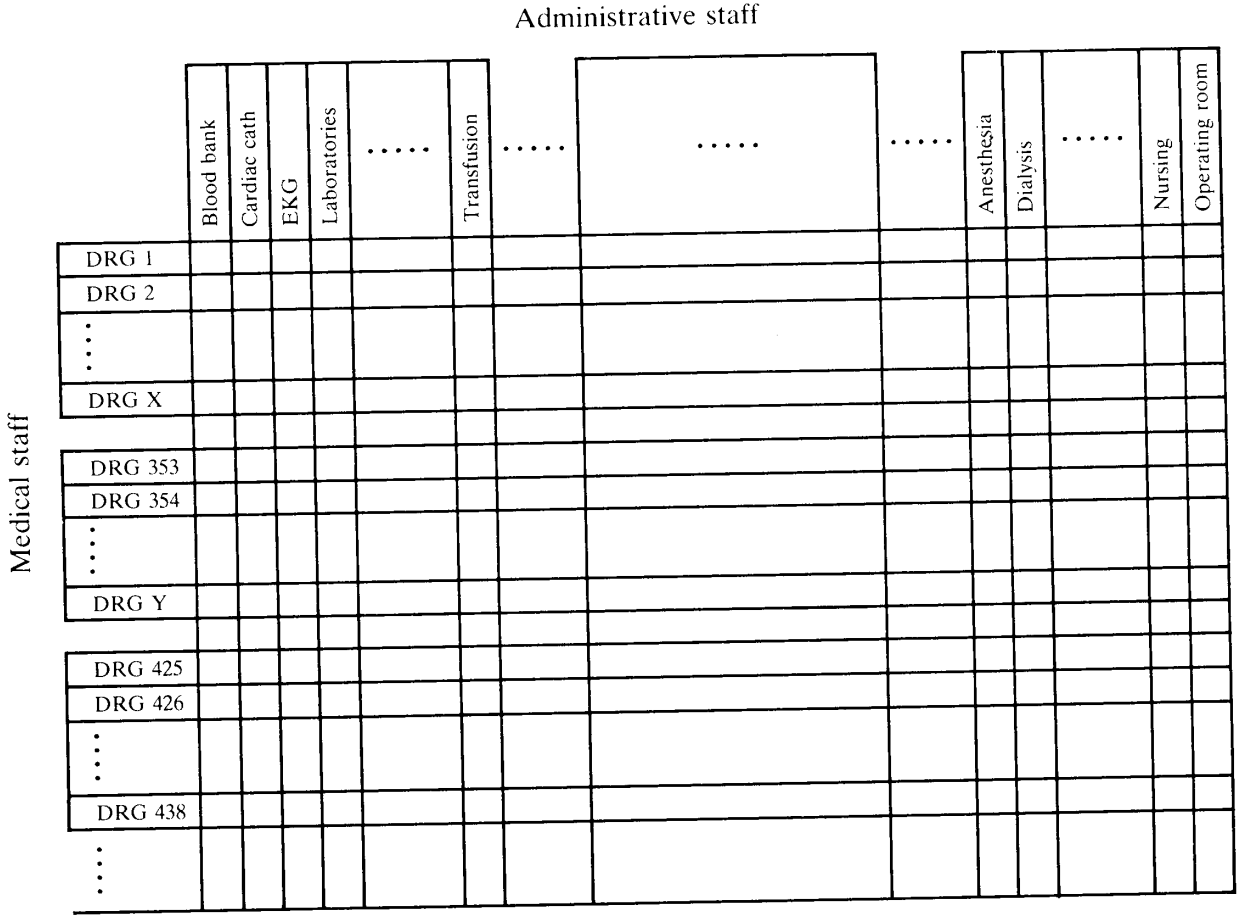
Input prices (personnel, materials)
Volume (number of patients treated)
Case mix (types of patients treated)
Efficiency (usage of input factors)
Treatment pattern (variations in physician prescription of services)

Thus, the analysis addresses simultaneously the administrative concerns of department managers and the clinical concerns of the providers. This will allow for a constructive dialogue between management and clinician. On a more global level, differences in practice and their cost can be compared across hospitals, allowing for the first time an accurate assessment of the reasons behind the widely divergent costs apparent in this sector. Fetter[13] has demonstrated the value of this approach in hospital service departments and has constructed mechanisms for its implementation.

Implementation of the product line approach requires a new organisational structure to the hospital's medical and administrative staffs. Under the traditional hierarchical structure of a hospital with administrators, assistant administrators and so on, the implication is that by managing the various departments (pharmacy, housekeeping, laboratory, radiology, and so forth) one is managing the institution. This structure does not recognise the fact that the ultimate product of the hospital is the complete set of services provided to each patient on the orders of physicians, not clean linen, nutritional meals, and appropriate medications.

The matrix structure, as described by Neuhauser[14] for the hospital setting, captures the concept of product line management in operational terms for the hospital's internal organisation (see Figure 1). The individual departments are responsible for providing the necessary support services required in the treatment of patients. The department heads oversee the conversion of inputs (labour, materials, equipment) to outputs (laboratory tests, x-rays, hours of nursing care). The physicians, on the other hand, are the product managers. They are responsible for assembling a package of outputs which are

Figure 1 Product line management within the hospital as a matrix organisation



ultimately provided to patients. The formal involvement of physicians in hospital administration and the identification of their responsibilities in resource management using a generalised form of this model is also being introduced in the health care systems of other countries. For example, a similar approach is fundamental to the introduction of management budgeting in the UK health service.

Once the matrix has been developed, performance of the medical and administrative staffs can be monitored in the delivery of patient care for the defined product lines. Physicians are responsible for determining the mix of the hospital's resources necessary to diagnose and treat each type of patient. Hence, they must be able to support any significant variances in the use of resources against some defined standards for the same group of patients. Those in charge of the laboratory, kitchen, blood bank, and so forth, are responsible for the production, including quality control, of their respective department's services. As such they are accountable for the efficiency with which specific services are provided.

Monitoring and accountability by product lines has become critical in the US with a system of DRG based hospital payment. However even under other systems it is clearly important for the financial viability of a hospital to have accurate information pertaining to the costs of treating different types of cases.

Conclusion

Concern over the rising costs of medical care has resulted in increased pressure on hospitals to control patient care costs through the adoption of more efficient management techniques, such as those commonly employed by manufacturing firms. The successful transfer of these methods to the hospital sector requires a structure for examining utilisation of services and for establishing standards and criteria for identifying areas which offer opportunities for improvement. The first step in providing this structure is to define what the hospital is producing.

As currently constructed, the DRGs provide hospital administrators and physicians with a powerful mechanism to understand and control hospital costs. Specifically, total costs within an institution can be broken down not only by type of patient (product) but also by service type (output). Control systems can therefore be designed to monitor the patient care and treatment process in a manner far more precisely and comprehensively than has been previously possible.

The potential role of the DRGs in controlling hospital costs is clearly apparent in the areas of case mix accounting and strategic planning. Changes in product or diagnostic mix can be identified and planned for under this system. The cost and revenue implications of actual or proposed changes in diagnostic mix can be estimated. The implications of evolving or alternative patterns of medical practice can be identified. It is felt that these implications can be stated in terms which both physicians and administrators can understand. The way physicians allocate resources, and the consequences of that allocation process on the administrator's concern with the finances of the hospital, can now be explained in terms more meaningful to the physician

than nursing costs per day or raw food costs per meal. Whether this will result in a change in behaviour of either is another matter, but the main assertion is that it can now be seen whether or not the increased information and subsequent education changes behaviour.

Our existing departmental costing mechanisms do not permit the physician to make the connection between the units of service department resources he uses and the way he is treating certain patients, and, up until now, this inability has too often been termed irresponsibility by those managing the hospital. The product oriented approach of DRGs allows both administrators and clinicians to ascertain the cost and quality implications of the various treatment plans within an institution.

References

1 National Center for Health Statistics: Health, United States. Hyatsville Md, US Department of Health, Education and Welfare, 1980. DHEW Pub. No. (PHS) 81-1232.
2 Lave J R and Lave L B. Estimated cost functions for Pennsylvania hospitals. Inquiry, 7, 1970: 3–14.
3 Lave J R. A review of the methods used to study hospital costs. Inquiry, 3, 1966: 57–81.
4 Lave J R and Lave L B. The extent of role differentiation among hospitals. Health Services Research, 6, 1971: 15–38.
5 Berry R E. Cost and efficiency in the production of hospital services. Milbank Memorial Fund Quarterly, 52, 1974: 291–313.
6 Berry R E. Product heterogeneity and hospital cost analysis. Inquiry, 7, 1970: 67–75.
7 Feldstein M S. Hospital cost variation and case-mix differences. Medical Care, 3, 1965: 95–103.
8 Ruchlin H S and Leveson I. Measuring hospital productivity. Health Services Research, 9, 1974: 308–323.
9 Fetter R B. The new ICD-9-CM diagnosis-related groups classification scheme. HCFA Pub. No. 03167. Health Care Financing Administration. Baltimore Md, US Department of Health and Human Services, 1983.
10 Codman E A. The product of a hospital. Surgery, Gynaecology and Obstetrics, 18, 1914: 491–496.
11 Chase R B and Aquilano N J. Production and operations management: a life-cycle approach. Homewood Il, Richard D Irwin Inc., 1977.
12 Thompson J D, Averill R F and Fetter R B. Planning, budgeting, and controlling – one look at the future: case-mix cost accounting. Health Services Research 14, 1979: 111–125.
13 Fetter A L. Cost measurement and control in a hospital service department: a new approach. Unpublished Master's essay, Department of Epidemiology and Public Health, Yale University, 1980.
14 Neuhauser D. The hospital as a matrix organization. Hospital Administration, 17, 1972: 8–25.

DRGs in perspective

1 CONCEPTS OF CASE MIX

Martin Bardsley

The process of distinguishing specific groups of patients, as embodied in the definition and measurement of case mix, is not especially new. In one sense the classification of patients into groups on the basis of diagnosis, aetiology, pathology and so on are central to both the theory and practice of modern medicine. What is relatively new is the practical application of these ideas of case mix into routine health service management.

Such measures are an example of the increasing sophistication of the tools available to aid decision making on health issues. In particular they address the economic aspects of health care which are currently on the ascendent as determinants in the policy making and management process. This chapter is intended to describe the rationale behind measures of hospital case mix, with particular reference to economic considerations. The aim being to place DRGs as a measure of case mix in a wider perspective and to consider the suitability and appropriateness of some selected alternative or additional systems classification.

Defining hospital outputs

The development of case mix measures is just one part of the increasing acknowledgement of the relationships between the types of care that can be provided and the money available. The *raison d'être* of health economics is that society will not necessarily value health to the exclusion of everything else. At a social level, health services do not have a blank cheque to cover any costs incurred, and total expenditure on health care is limited by a complex interaction of political, social and personal values. From within a health service the emphasis is therefore on allocating what monies are available in the way that will do most good. An acknowledgement of these financial constraints in health service funding leads inevitably to attempts at measuring the relative efficiency and effectiveness with which money is being spent. In its widest sense, 'efficiency' can be seen as summarising the relative relationship between some measure of output and the associated costs. Whether this relationship is expressed as benefits per pound or pounds per benefit, the idea is basically the same. The problem in the health service, or more specifically the hospital, is to be able to specify precisely what the relevant outputs are.

> There appears to be no agreement either on a conceptual or merely definitional level, among those who have most intensively studied the economics of hospitals, on what the appropriate measures of output is or should be.[1]

Hospital outputs have been variously defined – from the provision of services to patients[2], to improvements in health[3]. This diversity of approaches is a

reflection of some important theoretical and practical problems in specifying the nature of hospital output. The economic evaluation of efficiency requires, for practical purposes, a rather undimensional view of what benefits have been accrued and at what costs. Trying to find a measure that will be consistently and unambiguously associated with a net benefit from all perspectives, and at all levels, has proved impossible.

This diversity often leads to the acknowledgement of the multi-dimensionality of outputs[4] and the idea that any one measure may be capturing only part of the overall benefits. In addition to the strains and accommodations at the theoretical level, a number of compromises are made in practice. For example, it is very often the extent and the nature of the information available which leads to different specifications and formulations of a relevant measure.

In the past, attempts to develop standardised reference groups for economic analysis (that is, an appropriate specification of output) have split into two camps. On one side the primary concern has been with the quantification of patient health status, and health indicators, in an attempt to develop aggregate measures, at the societal level, of the benefits of health care technology; for example, the Quality Adjusted Life Year (QALY)[5]. The other route to 'output' measurement for evaluating 'efficiency' has been more pragmatic, and typically has compromised by using 'intermediate outputs' as proxies for real patient benefits in an attempt to study cost variations and, by implication, efficiency variations that exist in practice. Measurement of case mix is part of this line of research. As the definition and measurement of case mix improve it may well prove possible to unite these two lines of research.

Final or intermediate outputs?

An expression of hospital output in terms of the health benefits accruing to individual patients is as yet some way ahead of the management information tools available. Present mechanisms for reviewing hospital performance against cost rely heavily on the use of intermediate measures of output.

Hospital care is often represented as a bundle of services given to the patient on the recommendation of the doctor. The representation of hospital efficiency or productivity is then based on the costs associated with the production of a given unit of service (for example, an x-ray test or and inpatient day), which can form the basis of comparative indicators of performance. It is clear that such indicators only partially represent the activities of the hospital as a whole, the ultimate objectives of the institution, or the benefits accrued by the patient. These measures have been variously described as intermediate outputs of the hospital (the terminology adopted by Professor Fetter in the Introduction), intermediate inputs to medical care, or throughputs.

Though the ability to produce comparative costs of patient days or costs per discharge are a clear improvement on simply comparing aggregate hospital costs, it is by no means the final word in measuring 'efficiency'. Simple comparative measures of hospital costs per case will not be realistic if they ignore differences in the types of case treated. In order to overcome this, two strategies have been advocated[6] and can be seen, albeit in the early stages, in

the NHS. The first approach is to restrict the types of institution which can be compared, effectively classifying hospitals into groups. The second approach is to classify more directly according to the types of patient treated, using an analysis of case mix.

Service mix

One of the important determinants of overall hospital costs will be the range and volume of services offered, and much of the literature on hospital costs has been concerned with the search for economies of scale. It is fairly clear that in some areas a greater volume of services will correspond to lower marginal costs and a more efficient use of initially capital-intensive facilities. Therefore a comparison of costs which ignores the possibilities of these economies would be inequitable. In addition, the range of services available can be an important indicator of the types of treatment offered and the quality of care. For example if certain sophisticated and expensive diagnostic procedures are only available at certain specialist hospitals, is it reasonable to compare diagnostic costs in these departments with other hospitals lacking these facilities?

These are just two of the examples of the way in which a classification of hospitals can be justified. There are many others, and the selection of relevant criteria for hospital classification is itself a difficult process[7]. A variety of measures have been used which try to standardise for known differences in the range of services offered or the frequency of certain types of care[8].

In the UK, an acknowledgement of these factors is shown in the classification of hospital types and the grouping of hospitals on regional cost returns, a recognition that comparisons between specialised high technology teaching hospitals and small rural geriatric facilities may be something less than wholly fair to either.

Case mix

The second approach to overcoming the obvious effects of differences in service provision between hospitals has been to standardise for the type of patient – the case mix. Here the unit of comparison for use in the efficiency equation is a specific type of case. As such, this may well preclude the necessity to distinguish some aspects of the service mix.

Although the specification of a particular type of patient, as embodied in case mix, is still a long way from our ultimate health service output, it is an important step beyond the measurement of input efficiency in patient-days or costs per test. Implicit in the definition of patient types as the basis for comparison, is the ability to compare not only the costliness of producing individual service components of care, but the quantities and combinations of these inputs in the treatment process.

Within the NHS an acknowledgement of case mix in hospital cost accounting has been relatively limited. It is over twenty years since Feldstein's[9] original studies developed simple measures of case mix and demonstrated the important cost implications. More recently, the develop-

ment of specialty costing recognises the differences between the expected costs for successful treatments in different areas of clinical medicine[10]. Furthermore, the development of management/clinical budgeting will require more precise information than the speciality, allowing a distinction of case mix at the level of individual consultants[11].

However, the analysis of cost data at the specialty level, although a requirement for the future, is still not routine practice. Beyond the specialty, it is clear that there can be differences within the same specialty in the types of cases treated with important cost implications.

The classification process

The definition and measurement of case mix revolves around the process of classification which condenses the infinite variety of hospital patients into appropriate groups. One of the advantages of this process is that it makes certain forms of analysis practicable by reducing the number of groups that have to be considered. This does not deny the uniqueness of the individual, but rather facilitates complementary forms of an analysis at an aggregate level which would not otherwise have been possible.

When faced with the activities of a hospital, it is fairly obvious that individual patients are different in many ways – age, height, sex, shoe size, style of pyjamas, for example. Every patient is a unique blend of physical, mental and social characteristics. The aim of the case mix classification is to pick out characteristics of patients that are deemed relevant to a particular purpose. Patients are then split into groups on the basis of the selecting characteristics. Any one group will contain patients with similar chosen attributes, yet different from the members of other groups. The groups can then be used on either a quantitative or qualitative basis to understand and predict changes in the universe of hospital patients.

One of the important characteristics of this process is that the criteria for classifying patients will be related to the ultimate application of the system. A variety of different patient characteristics is available for classification, ranging from iso-symptom groups based on a similarity of patient symptoms on admission to iso-value groups with similar social valuations of the care provided[12]. A number of different classifications can be derived, each suited to a particular purpose. To complicate matters further, a classification may develop which is based on one or more similarities in patient types.

An appropriate classification scheme can be used in a number of ways. At one level it can provide a qualitative description of hospital case mix. In another sense the standardisation provided by the classification can facilitate comparisons between different institutions, for example the frequency of cases or the average cost to treat a given group. More importantly, a vector of case mix proportions can be combined into a simple scalar measure by applying relative weights to each group. Thus the expected costs of treating a given case mix can be represented as the sum of the number of cases in each group multiplied by the expected treatment costs of each group. It then becomes possible to condense the complexity of case types, in terms of expected costs, into a single descriptive statistic.

Evaluating case mix classifications

Many different schemes of patient classifications have been proposed which identify different aspects of the patient and/or treatment process, and are applicable to a variety of purposes. There are no definitive criteria which say whether one classification is a better description of case mix than another. There is however a battery of yardsticks with which to consider the individual merits of each scheme. Hornbrook[12] identifies a variety of criteria for evaluating the performance of a case mix measure. These are summarised in Table 1.

Table 1 Hornbrook's criteria for evaluating case mix classifications

1 Reliability	Consistency, not susceptible to random errors.
2 Validity	a) Content – representative and comprehensive b) Predictive – ability to predict some hypothesised outcome c) Construct – ability to explain differences in a way that is theoretically coherent
3 Sensitivity	Discriminates between hospitals
4 Cost-effectiveness	Least cost method of measurement without significantly compromising performance
5 Flexibility	Can be used for a variety of purposes
6 Acceptability	Measure is accepted by all users

For any one classification, or any one purpose, there tend to be trade-offs between these different aspects of performance. For example, the more sensitive a measure in its ability to discriminate between hospitals, the less reliable it will tend to become. It is also the case that different criteria will receive different weightings, and by different protagonists, in terms of their contribution to the overall performance of a measure. Though some comparative measures of alternative classifications do exist[13] these are not sufficiently comprehensive to favour unequivocally one measure against another. If, instead of asking what are the characteristics of a good measure of case mix, we consider why some measures are successful (in the biological sense that they survive and multiply) then a different picture emerges.

In particular it would seem that the two most important elements in case mix classifications are:

1 The extent to which the practical basis of the classification corresponds to the sort of measure that is theoretically needed – that is, validity in its widest sense.

2 The practicability of the scheme and, in particular, the extra costs of acquiring the information itself.

Precisely how these two elements are judged, the trade-offs between the two and the range of options considered, will be dependent on the perceived need

to change the way of looking at case mix. In practical terms this means that the evaluation of a case mix measure is not a technical exercise which can be conducted in isolation from the surrounding issues. The advisability of any one measure will be dependent on a choice between competing options. In the first case this will be the option of changing to a new definition of case mix, and then a consideration of relevant alternatives.

One of the most persuasive arguments in favour of a measure of case mix is that it can provide descriptions where none was previously available. Whatever the merits or de-merits of any one scheme, it will still be 'better than nothing'. If such systems of case mix can be implemented at little cost then there appears to be little to lose, and possibly a great deal to gain. It is interesting to observe that even the critics of DRG based reimbursement acknowledge that there is a place for measures of case mix and that the current scheme is not necessarily less equitable than fee for service.

Diagnostic classifications

Patient classifications based on diagnosis are perhaps the most well established ways of differentiating between patient types. Diagnosis is seen to sit at the centre of the medical decision making process, a synthesis of the patients' symptoms/problems and a determinant of expected treatment. Standardised classifications of diagnoses have been available for some time through the offices of the World Health Organization ICD schemes[14]. The classification is broadly statistical in attempting to group together conditions which are basically similar under one heading in the four digit diagnostic code. Though typically associated with epidemiological and clinical uses, the system in various guises has also been used to provide management information.

Despite their widespread popularity, diagnostic classifications are not immune from criticism. They are said to contain examples of many diverse conditions under one label, while in some areas have two codes for the same disease. The ICD system has been criticised for failing to make the distinction between health problems (symptoms, physical abnormalities and pathological manifestations) and diseases.

> Therefore a diagnosis should have information documenting four elements: the cause of the problem, the location of the problem, the manifestations of the problem and the severity of the problem. Unfortunately many of the diagnostic labels traditionally used by the medical profession and many of the diagnostic rubrics in the ICD coding systems, do not give these types of data[15].

However adequate ICD classifications are in categorising diseases there are a number of other considerations which must be borne in mind when these groups are used to study economic efficiency. For example, there is no exact and inevitable relationship between the condition of the patient and the diagnosis that results or the code that is entered on the discharge abstract. Even if this variability is ignored, as almost inevitably it must be, it can be debated as to whether we are interested in the patient's condition rather than the medical rationalisation of it – the health problem as opposed to the disease. Though these possible objections to diagnostic coding appear to be

rather abstract and unassailable it is important to bear them in mind when such codes form the basis for most computer-based patient clinical data and measures of case mix.

There are also a number of other areas in which a diagnostic classification may not fully match up to the hypothesised measure desired. The most pressing must be the lack of any indication of patient outcomes or the quality of care. It is necessary when using diagnostic classifications to assume that these factors are constant between patients and hospitals, or have no cost implications. There are also a number of areas of hospital activity where diagnosis is irrelevant, for example the provision of preventive services or organ donors.

Despite these theoretical objections, diagnostic classifications have been widely and usefully used to determine relative efficiencies. Their applicability as routine measures of case mix has been superseded in recent years by others for more practical reasons. In particular, diagnostic codes have not proved good at predicting resource variations due to the mix of patient types, and the number of groups (several thousand ignoring age and sex distinctions) makes them too unwieldy for many management information applications. In order to derive comprehensive descriptions of case types it is often necessary to collapse the classification to the 3-digit level, and even then the list of diagnoses needed to describe one area of clinical workload can be uncomfortably long. In a quantitative sense, diagnostic groupings have been used to account for case mix; for example, the UK performance indicators standardise relative length of stay using a combination of 3-digit ICD code, age and sex divisions. Though computers make such analysis possible for large numbers of individual cells, the small number of cases in each cell make comparisons difficult and statistically unstable (see Chapter 8).

Iso-resource groups

The recent advance of case mix into hospital management owes much to the use of DRGs as an iso-resource group. The terminology in these cases may be a little confusing. Iso-resource groups are not defined on the basis of expected resource use (that would mean a system classifying patients as £100 per case or £200 per case) but on the ability to discriminate between costs of treatment. Thus one group could be described more accurately as being iso-age within diagnosis, and coincidentally iso-resource. With these types of classification, the rationale is to identify variables which explain variations in cost. The variables embodied in the classification system can then be used to assess the cost implications of differences in case mix for individual hospitals.

A number of different iso-resource classifications have been proposed. One approach developed by the Commission on Professional and Hospital Activities (CPHA)[16] uses diagnostic codings in a system of 398 groups. Codes were grouped in a manner that was medically similar and contained patients with similar lengths of stay, as judged by panels of physicians. The groups were then broken down into five age-categories, the presence and absence of co-morbidities, and whether or not an operation was performed. The result was a classification containing 7,960 potential cells which for practical purposes is rather a lot.

As an alternative, the developers of the DRG classification placed an important emphasis on the practicability of the resulting groups, both in explaining resource differences due to case mix and in their potential use for routine management information[17]. Thus, DRG definitions were derived by using a statistical algorithm to maximise reductions in resource variation. This was constrained by the required elements of 'medical meaningfulness', a parsimony in the number of groups formed, and the requirement for routinely available computer abstract data (see Chapter 2). The result was a classification on a number of variables selected through their ability to predict variations in resource use. Perhaps the greatest asset of DRGs is their ability to use data immediately available on computer.

When considering the financial management of hospitals, an iso-resource classification which is technically efficient at teasing out cost differences due to case mix is most appropriate. In order to achieve satisfactory explanations of resource variance and medical meaningfulness, the DRG classification scheme includes some treatment-related variables, in particular the type of surgery performed. The resulting descriptions are therefore slightly further away than diagnostic codes from direct representations of the patient, it being assumed, for example, that the surgery performed was necessary.

When reimbursement takes a relatively passive role in hospital management, such a scheme is clearly attractive as a way of determining acceptable incurred costs (a judgment encoded into the DRG reimbursement rate). For other purposes, it is important to be aware of the discrepancies that may occur between the description of the hospital product, as embodied in the DRG, and the wider objectives of either the institution or the health care system. Thus the use of DRGs requires that aspects of the quality of patient care and outcomes are constant when making comparisons. This does not negate the value of DRGs; it is rather a second order problem that may follow an appreciation of case mix, and the classification itself may well provide the means to study the sort of cost/quality trade-offs that are currently implicit in the system.

A more precise specification of the relationship between, say, a DRG treated and the wider social benefits may also be important when determining normative costs for DRGs. An average cost for a DRG would unfairly penalise those hospitals which, through no fault of their own, had high input prices. Thus Medicare reimbursement in the US adjusts for local area wage differences. However it has been argued that other factors can affect the relative costliness of a hospital; for example the need to have specialised facilities on stand-by. It may be that an additional classification of hospitals is also required[18]. The problem here is to determine precisely what treatments should be done, where, and who should pay.

The severity debate

The application of DRGs in Medicare reimbursement has evoked some criticisms. One of the most consistent has been that DRGs fail to account for more severely ill patients. Thus a number of different schemes have been proposed as alternatives to, or improvements on, current DRG based reimbursement. Some are described below.

The debate on severity can be seen as representing just one of the areas in which the hospital product, as defined, say, by the DRG, does not match up to a desired goal in terms of true hospital outputs. With the system of reimbursement as it is, criticisms have therefore tended to concentrate on within-DRG differences in the state of the patient on admission rather than problems due to outcomes or quality. In particular, it is felt that inner city and teaching hospitals will be unfairly penalised by DRG reimbursement since they admit cases which are more severe than average and require correspondingly more resources to treat effectively.

A severity measure must be defined in such a way that it is independent of any single institution or clinician. This means that the definition must be a clear and unambiguous description of the patient on admission. To be relevant to financing arrangements, the severity level must be associated with higher costs of treatment and must be unequally distributed between hospitals. The need to deal with variations in severity would arise when one hospital received more than its fair share of 'severely ill' patients who were associated with higher treatment costs in order to achieve successful outcomes. If these conditions can be considered to be met satisfactorily, then the advisability of any routine measure of severity will revolve around the practicability of the scheme and, more specifically, the cost involved in collecting any additional data that may be required.

The traditional idea of severity is drawn from a specifically clinical view of a patient which may not be associated with differences in resource use and may ignore non-clinical aspects of the patient which necessitate higher costs. For example, a patient who receives routine surgery on the arm but who has no legs, is clinically identical to other patients yet, in practice, will require more nursing support.

If a strictly clinical view of necessary adjustments to DRGs is abandoned, then the delineation between the condition of the patient on admission and the appreciation of the quality of care during the stay can become increasingly blurred. To specify what patient variables present on admission are 'acceptable reasons' for increased resource use may require some very specific judgments about the sort of care that must be provided, as a minimal rule, and to what type of patient. It can be questioned therefore whether we wish to measure severity in the clinical sense at all, but rather are concerned with a more complex phenomenon.

The Patient Severity Index (PSI)

The PSI measure developed originally by Horn in Johns Hopkins has undergone a number of changes which make successive versions of the scheme more practicable as routine information systems[19].

The original scheme rates patients on seven variables: stages of principal diagnosis, co-morbidities, complications, dependancy, residual response to therapy, rate of response to therapy, and performance of non-operating room procedures. Each variable was rated on a scale of 1–4 corresponding to four problems; either none, mild, severe, or catastrophic. From this matrix a single score on a scale of 1–4 is obtained by a process of 'implicit integration'. Critics of the system were unhappy with this process, which failed to make the

rule for aggregating a score clear. However, Horn has demonstrated a degree of consistency between raters after training[20]. In general, the PSI overcomes the criticism of subjectivity through its ability to pick out, fairly consistently, the obviously unusual and probably more severely ill patients. The scheme is extremely good at identifying differences in resource use, as denoted by the very large reductions in variance obtained within a DRG. Part of this success may be due to the way the score can identify cases which are considered outliers in the DRG scheme.

Nevertheless it is clear that the PSI still performs well on trimmed data. The main barriers to the more widespread use of the scheme are possible theoretical problems in its definition and measurement of severity, and the costs of collecting the relevant information. There must be some questions about the ability of raters based in one hospital to evaluate the extent to which a patient's response to therapy is unduly poor compared with the national norm. Moreover, a poor response to therapy may be a reflection of either iatrogenic disease or deficiencies in the technical aspects of the quality of care. The development of PSI as a proposed sixth digit extension to ICD-9-CM, has involved more explicit criteria on how patients should be classified. These developments make the scheme more attractive in that the relevant criteria governing the scoring system are made explicit, and the encoding of severity at the same time as the diagnosis on the discharge abstract may reduce the costs.

Disease staging

The development and application of disease staging as a possible indication of severity has been championed by Gonnella and Systemetrics Incorporated[21]. The basic idea behind disease staging was originally developed in oncology, where the development of different types of cancer can be seen to progress along a distinct path which consists of four clearly defined stages.

Stage I: Conditions with no complications or problems with minimal severity

Stage II: Problems limited to an organ or system; significantly increased risk of complications

Stage III: Multiple sight involvements; generalised systemic involvement; poor prognosis

Stage IV: Death.

This approach has been used to identify successive stages for over 400 different types of diseases corresponding to a wide range of acute inpatient admissions. The criteria for each stage are specific to individual diseases and are defined by an identification of aetiology, a relevant organ, pathophysiology and the severity of a condition. The identification of each stage was originally based on specific clinical criteria developed by a panel of medical experts. Within any one major stage a number of different substages were identified which

> . . . should place the patient at a significantly higher risk of morbidity and/or mortality than the previous substage and should be clinically differentiable from other substages[22].

Having obtained a consensus among clinicians, these criteria could then be coded according to relevant diagnostic coding schemes. Though some loss of specificity occurred it was possible to translate, for the most part, the staging criteria into individual diagnostic codes which could be used to identify the stage of a disease from the information available on routine computer discharge abstracts. The stage of a disease is therefore not simply determined by the presence or absence of a specific diagnosis, but by all the diagnoses available and possible interactions that can occur between different conditions.

The results of computer-based staging have been compared to those derived manually in reabstracting studies. From a sample of 2,500 medical records, 77 per cent showed the computer and manual stages to be in agreement. Mismatches were due to either a lack of specificity in the ICD-9-CM coding systems, or a failure to record complications or laboratory findings on the discharge abstract, even though the information was available on the medical records. The resulting groups have been found to explain some variation in resource use[23] and give some prediction of patient outcomes, more specifically the risk of death.

However it is generally accepted that staging is not as good a predictor of resource use as DRGs or Horn's PSI[13], and the structure of the classification is more complex than either. The distinction of diseases required for staging does not coincide with those found in the DRGs, making it an uneasy partner as an additional amendment to the present DRG reimbursement systems.

As a classification, staging, like DRGs, must suffer from the disadvantages of using diagnostic coding schemes; however, it is particularly sensitive to errors of omission or commission in diagnoses. Nonetheless, as Gonnella himself points out:

> ... while the shortcomings of the coding systems and discharge abstract process are well known, the reality is that most current health services research, reimbursement approaches and other areas dependent on case mix measures are performed using automated discharge abstract data bases.[21]

It is also clear that a stage of disease determined on discharge may not necessarily identify the condition of the patient on admission. If poor quality care has meant that the disease has spread unnecessarily, a different stage will be assigned and presumably a higher level of reimbursement recommended.

Patient Management Categories (PMCs)

An alternative approach to developing an iso-resource classification was used by Young. In this case patient categories were defined by a panel of appropriately qualified clinicians. Each category was based on a consideration of both the form and extent of a disease as well as a recognition of the reason the patient was admitted to the hospital. In addition to developing the basic classification, the clinicians were asked to specify the components of treatment and diagnosis that a typical patient would expect to receive. Thus the variety of resources for each category could be represented by a path

through these individual components – a patient management path (PMP). The relative costliness of each component of care could then be calculated

> . . . to provide a basis for a relative value scale based on actual hospital costs (as opposed to charges or charge adjusted costs) of services required (as opposed to services rendered).[23]

As with staging, the demands for a classification which can be driven by routinely available abstract data meant that group definitions based on diagnostic and procedure codes were later developed which lost information on the reason for admission. The advantage of defining groups on an *a priori* clinical view is said to be their potential to be more specific about diagnoses indicating greater severity or economically relevant comorbidities. The operational definition of the categories therefore uses all the diagnostic information that is available. Young stresses the clinicians' view of severity within specific disease areas rather than as a generic concept that can be applied across all patient types. The success of any scheme based on combinations of diagnostic codes will depend on the extent to which the limited information provided is sufficient to make often very subtle distinctions between patients.

MEDISGRPS

The Medical Illness Severity Grouping System (MEDISGRPS) was developed by MediQual Systems and has been used to date in a limited number of hospitals[24]. The system identifies five severity groups determined according to specific 'key clinical findings' (KCFs). These may be the results of laboratory, radiological, pathological or physical examinations of the patient and recorded in medical notes. Each KCF is scored, according to explicit criteria, on a scale of zero for a 'normal' finding, to four when the observation is unusual and indicates a more severely ill patient. A patient is then scored on the basis of the most extreme KCF with some modifications if a group of KCFs with similar values occurs.

It is intended that the score be applied to patients within the first four days of admission to hospital and recalculated after ten days if necessary. By using the severity score within a classification of reasons for admission to the hospital, it is possible to judge not only the relative levels of resource use during the stay but also the quality of care that results.

Once again the clinical specificity required for this type of scheme means that data collection is potentially an expensive business. The other main questions must be over the reproducibility of the judgments initially built into the scoring system when the scheme is used in a wider range of hospitals.

APACHE

The Acute Physiology and Chronic Health Evaluation (APACHE) system was originally developed[25] in order to study patients admitted to intensive care units. The proponents of the scheme have recently advocated its use as a severity adjustment across the whole range of inpatient care[26]. The APACHE

score was developed through the screening of a selection of clinical variables for their ability to predict resource use and patient outcomes (death). The first APACHE scheme identified 33 clinically relevant patient variables recorded on admission to ICU which could be converted into a single score by a weighting system. The selection of variables, and their weighting and scoring, was based on the consensus view of a panel of clinicians. This scheme was later pared down to only twelve physiological variables, listed in Table 2. It includes age and an assessment of chronic health status in the APACHE II system. By scoring each variable on a predetermined scale according to the status of the patient on admission to the hospital, a single score representing a measure of severity can be derived.

Table 2 APACHE II severity of disease classification

Temperature Mean arterial pressure Heart rate Respiratory rate Oxygenation Arterial pH Serum sodium Serum potassium Serum creatinine Haematocrit White blood count Glasgow coma score	Acute physiology score Score up to +/− 4 points on each variable
Age	Score 0 to 6 points
Organ insufficiency or immuno-compromised state prior to admission	Chronic health evaluation Score 0 to 5 points

One of the obvious problems with this system is the level of detail recorded on each patient. Though computerisation of the monitoring of vital signs is increasing, there are not many hospitals which can claim to record this information on computer on a routine basis. When the system is advocated, not just for ICU patients but for all admissions, it must be questioned whether the information is collected at all. It remains to be seen how such a system would cope with incomplete data from some institutions and whether a workable scheme could be developed.

In general it is to be hoped that the variables chosen could be recorded fairly unambiguously in different institutions. The systems ability to evaluate the cases on admission is important to differentiate poor quality care from the more severely ill patient. A possible reservation about this approach is its dependence on detailed quantifiable variables. Though these variables lend themselves to reproducibility between hospitals, it may be that their selection as the basis for the score, over aspects of the patient which provide valid descriptions yet are less easily quantifiable, may limit the potential of the scheme.

Conclusions

A variety of case mix measures can be developed for a variety of purposes. The success of any one can be judged by considering the extent to which it will provide a measure that is theoretically relevant as well as being practical. DRGs as a measure of case mix have been presented as a measure of hospital 'throughput' which is superior to earlier indicators when studying the economic consequences of hospital care. The descriptions that DRGs provide do however fall some way short of an ultimate classification of hospital or health service output. The complexity of institutional objectives and the diversity of personal and social values cannot be easily condensed into simple uni-dimensional scales.

References

1 Berki S E. Hospital economics. Lexington, Mass, Lexington Books, 1972.
2 Mann J K and Yett D E. The analysis of hospital cost: a review article. Journal of Business, 41, 1968: 191–202.
3 Dowling W L. Hospital production: a linear programing approach. Lexington Mass, Lexington Books, 1976.
4 Luke R D. Dimensions in hospital case mix management. Inquiry, 16, 1979: 39–49.
5 Williams A. Economics of coronary artery bypass grafting. British Medical Journal, 291, August 1985: 326–329.
6 Tatchell M. Measuring hospital output: a review of service mix and case mix approaches. Social Science and Medicine, 17, 13, 1983: 871–883.
7 Klastorin T D and Watts C A. The determination of alternative hospital classifications. Health Services Research, 16, 1981: 205–220.
8 Horn SD and Schumacher D N. An analysis of case mix complexity using information theory and diagnostic related grouping. Medical Care, 17, 1979: 382–389.
9 Feldstein M S. Economic analysis for health service efficiency. Amsterdam, North-Holland, 1967.
10 Magee C. The potential for specialty costing in the NHS. Public Finance and Accountancy, March 1981.
11 Wickings I, Coles J, Flux R and Howard L. Review of clinical budgeting and costing experiments. British Medical Journal, 286, 1983: 575–577.
12 Hornbrook M C. Hospital case mix: its definition, measurement and use. Part I. The conceptual framework. Medical Care Review, 39: 1982: 1–43. Part II. Review of alternative measures. Medical Care Review, 39, 1982: 73–123.
13 Ament R P, Dreachslin J L, Kobrinski E J and Wood W R. Three case-type classifications: suitability for use in reimbursing hospitals. Medical Care, 20, 1982: 460–467.
14 World Health Organization. International classification of diseases: manual of the international statistical classification of diseases, injuries and causes of death. Geneva, WHO, 1977.
15 National Center for Health Services Research. Disease staging: a clinically based approach to measurement of disease severity. Volume 1. Executive summary. Hyatsville Md, US Department of Health and Human Services. 1983.
16 Commission on Professional and Hospital Activities (CPHA). Length of stay in short-term general hospitals, 1963–1964. New York, McGraw-Hill, 1966.
17 Fetter R B, Shin Y, Freeman J L, Averill R F and Thompson J D. Case mix definition by diagnosis-related groups. Medical Care, 18, 2, 1980 (supplement): 1–53.

18 Berki S E. The design of case-based hospital payment systems. Medical Care, 21, 1983: 1–13.
19 Horn S. Measuring severity of illness: comparisons across institutions. American Journal Public Health, 73, 1983: 25–31.
20 Horn S D, Chachich B and Clopton C. Measuring severity of illness: a reliability study. Medical Care, 21, 1983: 705–714.
21 Gonnella J S, Hornbrook M C and Louis D Z. Staging of disease: a case mix measurement. Journal of the American Medical Association, 251, 5, 1984: 637–644.
22 Garg M L, Louis D Z, Gliebe W A, Spirka C S, Skipper J K and Parekh R R. Evaluating inpatient costs: the staging mechanism. Medical Care, 16, 1978: 191–201.
23 Young W W. Incorporating severity of illness and co-morbidity in case mix measurement. Pittsburgh, Blue Cross of Pennsylvania Health Care Research Department, 1984.
24 Brewster A C, Karlin B G, Hyde L A, Jacobs C M, Bradbury R C and Young M C. MEDISGRPS: a clinically based approach to classifying hospital patients at admission. Inquiry, 22, 1985: 377–387.
25 Knaus W A, Zimmerman J E, Wagner D P, Draper E A and Lawrence D E. APACHE – acute physiology and chronic health evaluation: a physiologically based classification system. Critical Care Medicine, 9, 8, 1981: 591–597.
26 Wagner D P and Draper E A. Acute physiology and chronic health evaluation (APACHE II) and Medicare reimbursement. Health Care Financing Review, annual supplement November 1984: 91–105.

2 THE DEVELOPMENT OF DIAGNOSIS RELATED GROUPS

Dr Laurence F McMahon

Introduction

The development of diagnosis related groups must be viewed in the context of the United States hospital industry. While a formal review of the US hospital system is beyond the scope of the current discussion, a brief overview is necessary to appreciate the genesis of DRGs.

The hospital industry in the United States is relatively modern.[1] At the turn of the century there were only a handful of hospitals in the major American cities. These adopted their organisational structure from the British system and served the poor.[2] The affluent and middle class were treated at home. These early hospitals were financed by voluntary donations, not public funds or patient charges[3]. Around this time, the development of cities and advances in medical practice made the hospital more attractive to the American worker. The number of hospitals grew. A key feature of this growth was the hospital's identification with a geographic region of the city and/or with a particular subset of the city's population (that is, Catholic, Jewish and the like)[4]. Each institution had its own particular constituency and looked to the needs of that constituency when developing care programmes. This fragmented nature of the hospital industry is a key feature of the system today.

With the growth of the hospital system, the earlier methods used to support patient care, that is charity and some municipal support, were found to be inadequate. In the early 1900s patients began to be charged for hospital care. The system of charity plus increased direct patient charges was sufficient to support middle class and poor patients in voluntary hospitals from the turn of the century until the late 1920s. Hospitals then began to face rising losses, largely from the inability of patients to pay for their care. A new system was introduced, that of third party payment.

In 1929, Baylor University Hospital in Dallas, Texas, agreed to provide school teachers with up to 21 days of hospital care per year for six dollars per person[5]. This early system of third party payment developed into the non-governmental hospital insurance plans. Most employed persons' hospital care is paid through these plans, usually on a pre-negotiated percentage of the hospital's charges.

It was not until 1965 that the federal government moved into health care insurance and introduced Medicare, a federal programme to pay for the health care of those aged 65 and over, and Medicaid, a joint federal and state programme to pay for the health care of the poor. By this time US hospitals had grown into a multi-billion dollar a year industry. Payment for hospital care came from several sources and each major payer charged a different rate for the same service. Commercial insurance companies and self-paying patients were charged substantially more for the same service. Additionally,

if hospitals had costs that could not be recovered (bad debts), these were also shifted to the commercial insurance and self-paying patients. In a word, the hospital system that evolved in the US was fragmented, with separate institutions and a host of separate third party payers.

Development of diagnosis related groups (DRGs)

Against this backdrop, research in the 1970s began to focus on the nature of the hospital industry. Because of the independent hospitals, payers and physicians, a key research goal was to identify similarities and explain differences between hospitals. The focus of this analysis was on the resources expended, which had grown rapidly after the adoption of the federal insurance programmes[6] and with the inflationary sixties and seventies.

It soon became clear that resource use was closely linked with a hospital's case mix. Lee and Wallace in a 1972 paper[7] noted:

> The importance of case mix – the flows of different types of cases through a hospital – for determining hospital production costs has been widely recognised. Case mix is a more meaningful measure of hospital output than aggregate days; and more fruitful analysis of hospital production costs should result from taking differences in case mix into account.

A number of studies in the early and mid 1970s examined the effect of case mix in explaining hospital resource use [8, 9, 10]. A principal problem with the emerging case mix analysis was the lack of an agreed standard for a 'case', and the inability of early case mix measures to account adequately for the clinical differences in the cases.

An interdisciplinary research group at Yale University led by Robert Fetter from the Department of Administrative Sciences (later the School of Organisation and Management), and John Thompson from the Department of Epidemiology and Public Health of the School of Medicine, began in the late 1960s to look at hospital management, planning, utilisation review, and the like. Although each research topic was unique, they all raised the same fundamental issue. To study hospital management, planning and utilisation, one needed a focus – patient care. The business of hospitals was patient care, so any study involving hospitals needed a patient care orientation. It was also clear that care differed as a function of patient attributes, such as age and sex, and different states of disease. Therefore, if one was to address the above research questions, an *explicit* definition of the different types of patient care was needed.

As a first attempt to segregate patients into unique groups, existing hospital-based patient classification systems were evaluated. One potential classification system was to segregate patients solely on the basis of their principal diagnosis coded in the International Classification of Disease (ICD) system[11]. This approach was felt to be too simplistic. Important patient attributes such as age were ignored, as was the interaction of the principal diagnosis with other diagnoses, such as diabetes and pneumonia, or with surgery, for example diabetes and amputation. In addition, the classification of patients into groups based on their principal diagnosis created so many groups that the system would have been unworkable.

Another system evaluated had been developed by the Professional Activity Study (PAS)[12]. This classification was based upon the principal diagnosis, the presence or absence of additional diagnoses, the presence or absence of any procedure, and five age categories (0–19, 20–34, 35–49, 50–64, 65+years). The PAS system presented many of the same problems as the ICD. It failed to distinguish among secondary diagnoses, causing patients with diabetes and pneumonia and diabetes and hypertension to fall into the same group. The problem also occurred with surgical procedures, and a patient with diabetes and a toe amputation and one with diabetes and an abdominal aortic aneurysm repair found themselves in the same group. The use of five arbitrary age divisions resulted in 7,000 groups, most of questionable utility. PAS was an improvement over the simple use of the principal diagnosis as a method of classification, but it was clinically inadequate and, because of the large number of groups, administratively cumbersome.

It became clear that to define unique types of hospital-based patient care a new classification system would be required to meet four principal objectives:

1 It must be interpretable medically, with subclasses of patients from homogeneous diagnostic categories. That is, when the patient classes are described to physicians, they should be able to identify a particular patient management process for them.
2 Individual classes should be defined on variables that are commonly available on hospital abstracts and are relevant to output utilisation, pertaining to either the condition of the patient or the treatment process.
3 There must be a manageable number of classes, preferably in the hundreds instead of thousands, that are mutually exclusive and exhaustive. That is, they must cover the entire range of possible disease conditions in the acute-care setting, without overlap.
4 The classes should contain patients with similar expected measures of output utilisation.[13]

Using these guidelines, it was expected that the patients within a given patient care class or group would use similar hospital resources (iso-resource) and their aggregation in a group would make sense medically to physicians (medically meaningful). The requirement of using only available abstracted data was necessary if the grouping system was to be useful in a wide variety of institutions, for management, or for agencies to assist in health planning. Finally, the attempt to limit the number of groups was felt to be necessary to ensure a manageable system. Thus, the goal was to develop a manageable number of medically meaningful iso-resource groups that could be used for hospital management, planning, utilisation review and the like.

Creating the DRGs

In order to define a grouping of patients that was medically meaningful, the medical characteristics of the patients in the group had to be available for physicians to evaluate. To develop iso-resource groups, the resources utilised during a hospitalisation needed to be aggregated. These two requirements dictated a unique interaction between statistical analysis, for resource partitioning, and medical review to ensure that the medical characteristics of

the patients were similar. This two step process was necessary to ensure that resource partitioning and clinical partitioning did not overwhelm each other. For example, patients who have a hernia repair may use the same amount of a hospital's resources as those who have a cataract removed, yet grouping their records together does not make sense clinically. Similarly, a 50 year old patient with an ulcer and a haematocrit reading of 40 per cent is not very different clinically from a 40 year old patient with a haematocrit of 30 per cent, yet the resources utilised are likely to be quite different.

Assuring this balance between the iso-resource goal and the clinical coherence goal proved to be most challenging. It was initially decided to use length of stay (LOS) as the hospital resource measure, given its presence in all hospital data bases and the fact that it had the same meaning in all hospitals. Early attempts to partition hospital resource data into iso-resource groups used a statistical approach, the Automatic Interaction Detector (AID) developed by Sonquist and Morgan[14].

Because of the necessity for rapid evaluation of the resources consumed by group members *and* their clinical characteristics, it became clear that an interactive statistical system capable of rapidly displaying clinical information would be required. A new statistical system called AUTOGRP (Autogroup) was developed for this task[15]. Its key features include:

1 A partitioning algorithm similar to the AID which suggests groupings of an independent variable (for example, age) based on its ability to partition the dependent variable (in this case LOS).
2 A rapid display capability of both the statistical and clinical parameters of the proposed groups.

Using this system physicians could evaluate, for example, which of a selection of independent variables (age, secondary diagnosis, sex and the like) best segregates statistically non-surgical patients who have diabetes. After reviewing the statistical results of the groupings for each of the independent variables, the clinical characteristics of the proposed groups could then be reviewed. In this way, both the statistical and clinical characteristics of possible groups, based upon the various independent variables, could be evaluated efficiently.

Early DRGs

Following the development of an interactive statistical system capable of integrating statistical and clinical analysis, the early development of patient grouping began. In the early 1970s a key concern was, as today, the rising cost of hospital care and its control. At this time it was felt that costs could be lowered through the review of both unnecessary hospital days and services. This process of utilisation review received a major emphasis nationally under the Professional Standards Review Organisation (PSRO) programme [16]. The advent of the PSRO programme provided an early stimulus to work on DRGs.

With the development of AUTOGRP, work on the definition of groups began using data from individual hospitals. This hospital-specific phase of group development was spurred on by each hospital's desire to obtain a more

accurate sampling framework with which to conduct utilisation review. It was felt that creating groups of similar patients would improve the hospital's ability to highlight cases whose resource use deviated from the expected, and were more likely therefore to have problems with the appropriate utilisation of hospital resources[17].

The development of DRGs moved ahead rapidly in the 1970s. The concept of defining groups of similar patients for purposes of utilisation review and management raised the prospect of the same groups being used as a template for payment[10]. A major project was undertaken to develop a new set of patient groups specifically for the purpose of third party hospital payment. It was undertaken by the Yale research group in conjunction with the federal government's Social Security Administration (the agency which then oversaw the federal health care programmes) and the State of New Jersey's Department of Health[18].

The process of group definition required the close cooperation of health service researchers and physicians who adopted a twofold goal:

> to form groups of patients which displayed the least variance in resource use (LOS) while,
> creating groups that were medically coherent.

This joint analysis resulted in the creation of 383 groups which became known as diagnosis related groups because the first partition into major diagnostic categories (MDCs) was on the basis of the principal diagnosis. Details of this 383 grouping process were published in a supplement to *Medical Care*[13].

The 383 group version of DRGs was used to construct a hospital payment system for the State of New Jersey under a cooperative agreement between the state and the federal government. This system was the first large scale attempt in the US at prospectively paying for hospital care on the basis of the hospital's case mix[19].

Development of the new DRGs

In the late 1970s it became clear that there were significant structural problems with the 383 set of DRGs. At the same time, there was growing interest in expanding the New Jersey DRG based hospital payment experiment and the new International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) was released.[20] If the New Jersey experiment was to expand, a more representative database was needed upon which to construct DRGs utilising the new ICD-9-CM dataset. With experience gained from the construction of the first set of DRGs and taking account of the problems identified by critics of the system, the construction of a new set began in 1979 with the support of the Health Care Financing Administration (HCFA), the governmental agency which assumed authority for all federally supported health programmes[21].

A principal goal of the revision was to improve the clinical coherence of the groups by using a nationally representative database. A national stratified sample of 1.4 million records chosen from 325 hospitals selected for their quality coding was assembled for the DRG revision. Because of concern about the clinical coherence of the resulting groups, it was decided, in

consultation with physicians, to partition the data into organ system based groups. Using this approach, the ICD-9-CM system was reclassified into organ systems based on sequences of codes; in this way major diagnostic categories (MDCs) were defined.

Rather than include both individual diseases and organ systems, as had happened in the 383 set of DRGs, the MDCs were defined on organ systems to which physicians' practice largely conformed. Thus, the 383 set MDC called diabetes, was now included in MDC 10: endocrine, nutritional, and metabolic disease and disorders. By assigning most patients initially into organ system based categories, the number of MDCs dropped from 83 to 23. The next partition of the data, again clinically based, was to segregate patients into those who had surgery and those who did not. In response to criticisms of the 383 set, surgery was specifically limited to procedures typically performed in an operating room. Before any statistical analysis of the data had been carried out, 44 groups were formed (two MDCs do not contain surgical groups, see Appendix 1)[21].

After dividing the cases first into the MDCs and then into either surgical or non surgical groups, the AUTOGRP-aided partitioning of the data began in conjunction with panels representing every medical and surgical specialty. To provide a measure of uniformity across the MDCs it was decided to segregate all surgical cases according to the *type* of surgery performed, and all medical cases on their principal diagnosis. Additionally, all surgical procedures on each discharge abstract were reordered into a hierarchy of resource consumption, making their actual order on the record abstract irrelevant. This initial partitioning was supported by the clinical panels as being clinically logical.

Additional variables were then evaluated to assess their statistical and clinical influence on resource use (LOS). As might be expected the specific clinical variables found to influence resource use varied from MDC to MDC. Depending on the MDC, diagnostic groups (corresponding to medical hospitalisations) and procedure categories (corresponding to surgical hospitalisations) may be further partitioned on the basis of age, the existence of specific comorbidities and complications, and, in a few cases, discharge status (that is, death). Some procedure categories are also partitioned on the basis of principal diagnosis.

A compound variable, age > 69 and/or CC, is used extensively throughout the system. This is a dichotomous (2 level) variable which takes on the value 'yes' if age > 69 and/or there are substantial comorbidities or complications; otherwise, 'no'. Analysis of actual data found that this age/CC partition affected patient care resource utilisation to a significant degree. This empiric data therefore supplanted the more traditional > 64 year old partition that served as the basis of former reimbursement policies. Substantial complications and comorbidities are defined as those *specific* additional conditions which, in the judgment of the clinicians constructing the system, would increase the length of stay for 75 per cent of the patients by at least one day. The definition of a *specific* set of diagnoses that elevate a patient into a CC group was motivated by the criticism of the earlier 383 set of DRGs in which *any* additional diagnosis could increase a patient's group assignment. Thus, while essential hypertension would elevate a patient into a comorbidity or complication group in the 383 set of DRGs, it does not in the new set.

The most significant changes in the new DRG system result from:

1 The redefinition of the major diagnostic categories in terms of organ systems.
2 The restriction of surgical categories only to operating room procedures related to the principal diagnoses in their respective MDCs.
3 The definition of significant comorbidities and complications that are specific and based (as is the entire DRG system) on physician review.
4 The reordering of surgical procedures into a hierarchy based on their resource consumption prior to AUTOGRP partitioning.

This new version resulted in a total of 467 groups.

Example of a DRG partition; MDC 12: disease and disorders of the male reproductive system

Reviewing an example of the partitioning of a major diagnostic category will highlight the process undertaken in the definition of the new set of DRGs. The development of DRGs in this example will follow the general form illustrated in Figure 2. The first step was to define the major diagnostic category. In this example the MDC consists of all principal diagnoses related to disease and disorders of the male reproductive system[21]. The patients' records are then divided into two groups, those having a surgical procedure normally performed in an operating room (surgical patients), and those who did not (medical patients). In the third step, surgical procedures were reordered into a resource-based hierarchy making the order in which they actually appear on the abstract irrelevant. In the subsequent analysis, the most resource intensive surgical procedure on the abstract determines into which DRG a surgical patient is placed.

The results of this three-step partitioning are listed graphically in order of resource intensiveness. In this example, surgical procedures range in intensity from major pelvic procedures to circumcision (see Figure 3). The final surgical category, other OR procedures, refers to a group of procedures performed on patients who have diagnoses related to disease and disorders of the male reproductive system which, on an individual basis, occur infrequently. A residual category for all surgical procedures not conforming to the principal diagnosis are assigned to group 468; for example a urinary retention principal diagnosis with a toe amputation procedure. Records assigned to group 468 require individual analysis.

In a similar manner, non surgical patients are stratified in order of decreasing resource use based upon their principal diagnosis. This ranges from malignancy at one extreme to sterilisation at the other. As in the surgical example, a residual group of diagnoses labelled 'other diagnosis', is listed as the final DRG (352) in this MDC. Note that sterilisation of males is listed in the non surgical category. This is consistent with the stipulation that a surgical procedure must customarily be performed in an operating room. In the US, male sterilisation is often performed in the physician's office.

The fourth and final step attempts to identify additional independent variables that define iso-resource groups and are clinically sensible. As might

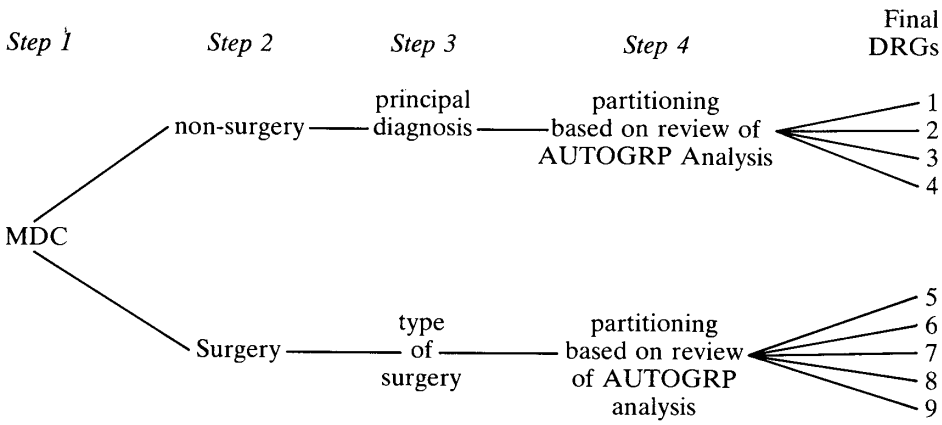
Figure 2 Steps in the definition of the new ICD-9-CM diagnosis related groups

Step 1: Partition into Major Diagnostic Categories (MDC) based on principal diagnosis. MDCs based on organ system specific groups of diagnoses to which clinical practice largely conforms.

Step 2: Separation of each MDC into surgical groups (defined as patients having a procedure customarily performed in an operating room), and non-surgical groups.

Step 3: Surgical groups – surgical procedures first arranged into a hierarchy based upon the most resource intensive procedure performed during the hospitalisation then partitioned on the basis of the type of surgical procedure.
Non-surgical groups – partitioned on the basis of the principal diagnosis.

Step 4: Within each surgical group and non-surgical group, additional partitioning is based on review of statistical analysis using AUTOGRP, and physician review of the clinical characteristics of proposed further partitioning.



be expected, the relevant variables will vary among groups of procedures or diagnoses, even in the same MDC. In the surgical grouping of our example, major pelvic procedures are divided according to whether patients also have a significant comorbidity or complication. In the non surgical category, patients whose principal diagnosis was either malignancy or benign prostatic hypertrophy were partitioned further, depending upon whether they had a significant comorbidity or complication, or were 70 or older.

Response to criticisms of the 383 set of DRGs

The two most important criticisms of earlier versions of DRGs concerned:

1 the ability to boost a patient into a higher paying DRG with minor changes in coding, so called DRG creep, and

2 the ability of DRGs adequately to account for differences in severity of illness [22, 23].

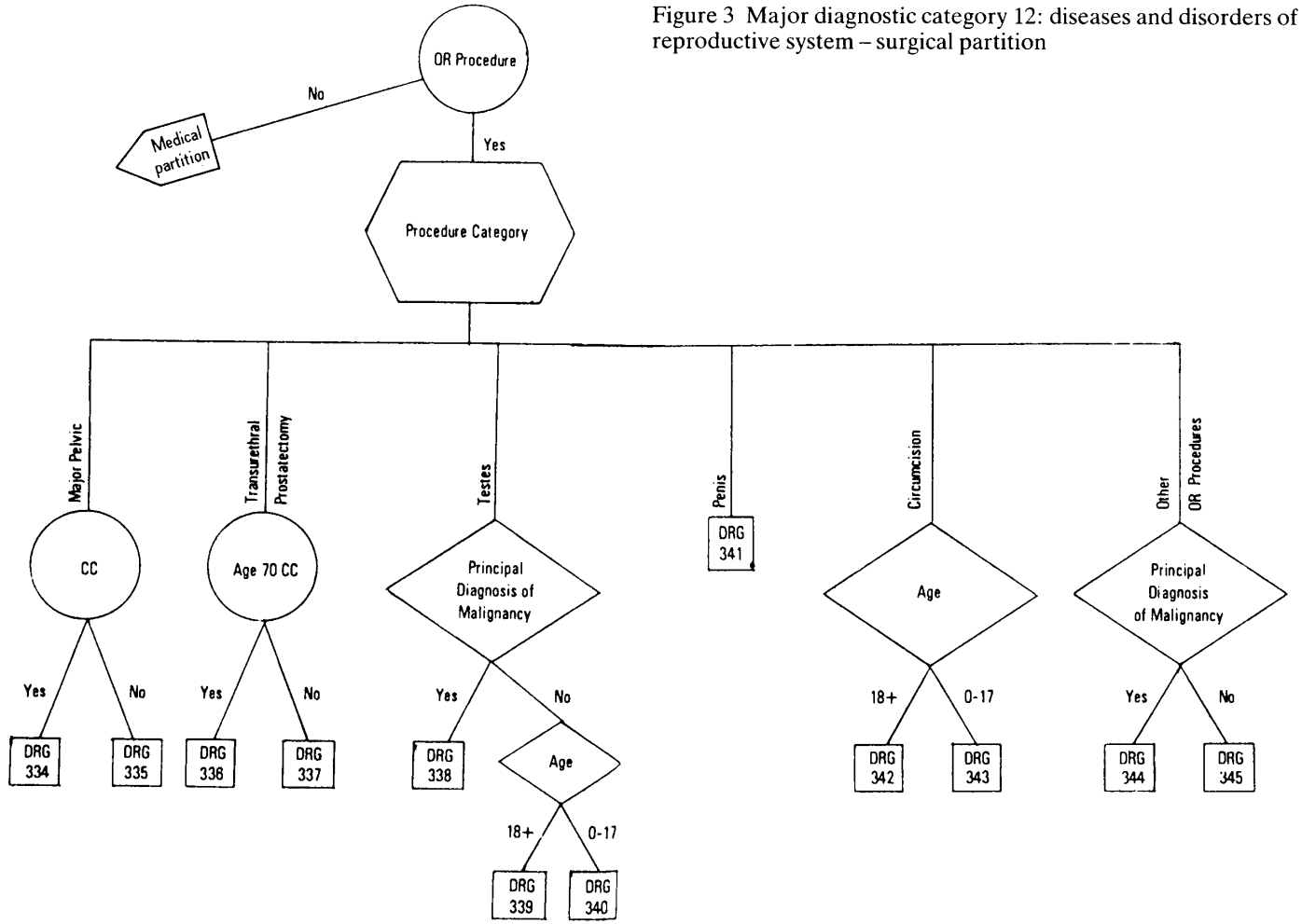


Figure 3 Major diagnostic category 12: diseases and disorders of the male reproductive system – surgical partition

The first concern, DRG creep, was raised by Simborg. He identified the problem as upgrading DRG assignment by adding clinically unimportant information to the discharge abstract, or rearranging the diagnostic information on the discharge abstract. The significance of this problem in the 383 set of DRGs was related largely to that system's underlying structure which allowed *any* second listed diagnosis or procedure to elevate the patient into a new group. (The additional problem of *arbitrarily* changing the order of coding solely to optimise DRG assignment, could not be directly dealt with in the DRG development phase. It had to be addressed in the context of the rules of the payment system). Upgrading DRG assignment by the addition of clinically unimportant diagnoses and/or procedures was, however, subject to correction at the level of the DRG definitions. The issue of intra-DRG severity of illness measurement is a complex problem. At the time of the 383 set, there was concern that the DRG's partitioning structure did not identify the *specific* types of secondary diagnoses and procedures, and would mask important differences in severity of illness.

The new 467 set attempted to address each of these issues by developing explicit criteria for defining significant comorbidities and surgical procedures. The concern about fraudulently mislabelling a principal diagnosis solely to optimise payment raised by Simborg in his discussion of DRG creep, was tackled through the administrative rules of the payment system. (The principal diagnosis in the new DRG-based prospective payment system is defined by law as that diagnosis which on discharge and after analysis of the data from the hospitalisation, was the principal reason for the hospital admission).

The question of severity of illness was addressed via:

1 The definition of surgical and medical hierarchies within each MDC.

2 The specific definitions of surgical procedures, comorbidities and complications.

3 The segregation of statistically unique patients as outliers.

The extent to which residual intra-DRG severity of illness variation remains is subject to continued debate as outlined in a recent paper[24].

Adoption of DRGs for hospital payment

The retrospective third party reimbursement system outlined earlier proved very costly. In particular the federal government's contribution to hospital care increased from 3.1 billion dollars in 1967, to 36.3 billion dollars in 1982 and grew at 15–22 per cent per year in the early 1980s.[25] Many efforts to moderate the rate of rise in hospital costs – such as utilisation review, certificate of need, and second opinion programmes – had failed.

The adoption of the DRG-based Medicare Prospective Payment System was motivated by provisions in the earlier Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA) PL 97–248. This law was designed to cap the amount of Medicare money available for hospital-based care. Under section 101(c) of TEFRA, the secretary of Health and Human Services (HHS) was required to develop a proposal for the prospective payment of hospital care for Medicare beneficiaries.

In December 1982, the then secretary of HHS, Richard Schweiker, outlined a prospective payment system based upon DRG patient definitions which was meant to ensure that Medicare would become a prudent purchaser of hospital care for its beneficiaries[26]. Secretary Schweiker reviewed the familiar litany of problems with the retrospective cost-based payment systems, which could pay one hospital $1,500 for a patient with a heart attack and another $9,000 for an apparently similar patient. He noted that:

> Since patients have different diagnoses, require different treatments, are of different ages, and differ in other ways, it is important to develop a payment system that explicitly adjusts for these differences. Prospective payment systems which do not recognise differences in case mix will severely harm the tertiary care hospitals which treat more complex illnesses, as well as rural hospitals, which have a volatile case mix. The lack of a case mix adjuster would also make the severely ill patient a financial liability to all hospitals and encourage some hospitals to admit only less severely ill patients.[27]

The ability of a case classification system to differentiate clinically distinct patient types is therefore crucial to the equity of case-based prospective payment.

In his report to Congress, secretary Schweiker reviewed the existing case classification systems available for either severity of illness adjustments or as alternatives for the DRGs to serve as the basis for prospective hospital payment. Three systems in addition to DRGs were reviewed in this report; disease staging, the APACHE system, and the severity of illness index. While these systems were very different from DRGs and each other, they were each developed to address the issues of measuring hospital resource use or severity of illness that came to the fore in the 1970s.

The disease staging system is based upon physician-defined stages (from one to four with a variable number of substages) for each of 406 diseases [28, 29, 30]. The patients within a stage are assumed to have a similar level of severity. The acute physiology and chronic health evaluation (APACHE) score was developed by Knaus et al. to assess the severity of illness of intensive care unit cases. [31, 32] Although the system has been shown to predict mortality in intensive care units using largely physiologic variables, these variables are not routinely collected on all hospitalised patients; nor had the system been evaluated using a general population of patients. The severity of illness index developed by Horn et al. was designed to measure the severity of illness of hospitalised patients using a generic four-level summary scale constructed by implicitly integrating seven variables that attempt to measure different aspects of severity of illness. Each of the seven variables are also subjectively rated on a 1–4 scale[18, 33].

This review noted problems with each system and found that DRGs represented the only workable case mix measure that was available for a prospective payment system. The important features that were noted included first, that DRGs were defined using available information from the computerised hospital discharge abstract. Second, all patients fall into a DRG. Third, the definitions were developed on a nationally representative sample of data. Fourth, very unusual cases (outliers) can be identified easily

so that they can be reviewed and paid in a different manner. Fifth, given the underlying structure of DRGs which links clinical data and resource data they can serve as a common language to increase interaction between hospital managers and physicians, an increase which may enhance their ability to improve hospital efficiency and effectiveness.[34]

Although the DRGs have undergone many changes in the past 15 years, it is likely that continued improvements will be necessary. Just as with the 383 set problems arise with use and solutions need to be found. It is expected that the DRG definitions will continue to evolve within the context of their four guiding principles used to define clinically coherent iso-resource groups for hospitalised patients.

References

1 Vogel M J. The invention of the modern hospital: Boston 1870–1930. The University of Chicago Press, Chicago, 1980.
2 Thompson J D. The uneasy alliance. In: Dyaggy and Anlyan (eds). Physicians and hospitals. Durham NC, Duke University Press, 1985: 11–19.
3 Starr P. The reconstitution of the hospital. In: The social transformation of American medicine. New York, Basic Books, 1982.
4 Rosner D. A once charitable enterprise. Cambridge, Cambridge University Press, 1982.
5 Berman H and Weeks L. Blue Cross in the financial management of hospitals. Ann Arbor, Health Administration Press, 1982.
6 Freeland M S and Schendler C E. Health spending in the 1980s: integration of clinical practice patterns with management. Health Care Financing Review, 5, 1984: 1–68.
7 Lee M L and Wallace R L. Research report classification of disease for hospital cost analysis. Inquiry, 9, 1972: 69–72
8 Lave J R and Lave L B. The cost and length of a hospital stay. Inquiry, 13, 1976: 327–343.
9 Lave J R and Lave L B. The extent of role differentiation among hospitals. Health Services Research 1971: 15–38.
10 Thompson J D, Fetter R B and Mross C D. Case mix and resource use. Inquiry, 12, 1975: 300–312.
11 Commission on Professional and Hospitals Activities. Hospital adaptation of ICDA. Ann Arbor, Commission on Professional and Hospital Activities, 1976 (2nd ed).
12 Commission on Professional and Hospitals Activities. Length of stay in PAS hospitals, by diagnosis. Ann Arbor, Commission on Professional and Hospital Activities, 1976.
13 Fetter R B, Shin Y, Freeman J L, Averill R F and Thompson J D. Case mix definition by diagnosis related groups. Medical Care, 18, 1980 (supplement): 1–53.
14 Sonquist J A and Morgan J N. The detection of interaction effects. Ann Arbor, University of Michigan, Instute for Social Research, 1964.
15 Mills R, Fetter R B, Riedel D C and Averill R. AUTOGRP: an interactive computer system for the analysis of health care data. Medical Care, 14, 1976: 603–615.
16 Decker B and Bonner P (eds). PSRO: Organization for regional peer review. Cambridge Mass, Ballinger Publishing Company, 1973.
17 Riedel D C, Fetter R B, Mills R E and Pallett P. Basic utilization review program (BURP). New Haven, Institution for Social and Policy Studies, Center for the Study of Health Services, Yale University, 1973.

18 Dunham A B and Morone J A. DRG evaluation volume IV-A: the politics of innovation. Princeton, Health Research and Educational Trust of New Jersey, 1983.
19 Iglehart J K. New Jersey's experiment with DRG-based hospital reimbursement. New England Journal of Medicine, 307, 1982: 1655–1660.
20 Commission on Professional and Hospital Activities. International classification of diseases, 9th revision. Clinical modification. Ann Arbor, Commission on Professional and Hospital Activities, 1978.
21 Health Systems Management Group. The new ICD-9-CM diagnosis related group classification scheme. Final report. United States Department of Health and Human Services, Health care financing administration publication #03167, September 1984.
22 Simborg D W. DRG creep: a new hospital-acquired disease. New England Journal of Medicine, 304, 1981: 1602–1604.
23 Horn S D. Measuring severity of illness: comparisons across institutions. American Journal of Public Health, 73, 1983: 25–31.
24 Smits H L, Fetter R B and McMahon Jr L F. Variation in resource use within DRGs: the severity issue. Health Care Financing Review, annual supplement, 1984.
25 Gornick M, Greenberg J N, Eggers P W and Dobson A. Twenty years of Medicare and Medicaid: covered populations, use of benefits and program expenditures. Health Care Financing Review, annual supplement, 1985: 13–59.
26 Schweiker R S. Hospital prospective payment for Medicare. Report to Congress. Washington, Department of Health and Human Services, December 1982.
27 See 26: iii–iv.
28 Gonella J S, Louis D Z and McCord J J. The staging concept – an approach to the assessment of outcome of ambulatory care. Medical Care, 14, 1976: 13–21.
29 Gonella J S, Hornbrook M C and Louis D Z. Staging of disease: a case mix measurement. Journal of the American Medical Association, 1984: 637–644.
30 Systemetrics Inc. Disease staging: a clinically based approach to measurement of disease severity. Vol. 1, executive summary. Washington, National Technical Information Service, PB 83-254656, August 1983: 1–25.
31 Knaus W A, Zimmerman J E, Wagner D P, Draper E A and Lawrence D E. APACHE – acute physiology and chronic health evaluation: a physiologically based classification system. Critical Care Medicine, 9, 8, 1981: 591–597.
32 Wagner D P, Knaus W A and Draper E A. Statistical validation of a severity of illness measure. American Journal of Public Health, 73, 1983: 878–884.
33 Horn S D, Sharkey P D and Bertram D A. Measuring severity of illness: homogeneous case mix groups. Medical Care, 21, 1983: 14–31.
34 McMahon L F, Fetter R B, Freeman J L and Thompson J D. Hospital matrix management and DRG-based prospective payment. Hospital and Health Services Administration, 31, (1), 1986: 62–74.

3 REIMBURSING HOSPITALS BY DRG

Linda Jenkins

One way in which DRGs stand out from other attempts to describe case mix is that they have seen active service as a determinant of hospital funding. First in 1980 and again in 1983 they played a fundamental role in schemes for reimbursing United States hospitals for patient care. The fact that this new and relatively untried management tool was adopted so widely and so quickly can be explained partly by the ease with which it can be used but perhaps more convincingly, by the financial pressures building up in the US health care industry.

The 'double-digit' inflation experienced in the US in the 1970s was compounded by both internal and external pressures on health care spending. Inside hospitals, charges were based on a fee for service which tended neither to limit the cost of patient care nor provide a competitive market in which prices might be held down. The falling lengths of stay left empty beds creating the incentive to admit more patients. Externally, inflation was already running high on all goods and services. It was felt even more acutely in the hospital sector because the cost of high technology equipment was rising faster than inflation. This coupled with the ageing or greying population led to increased demands on health care funds.

To say the DRG system is easy to use may seem somewhat naïve. What can be done easily is to take brief details of a patient's episode in hospital and assign him or her to the appropriate DRG. Most group definitions use the diagnoses and operations with the variables – age, sex and discharge status – being brought into play where they significantly influence treatment costs.

Establishing the detailed characteristics of these groups in terms of expected resource consumption, and translating these into costs, is considerably more difficult. Questions of how total costs should be apportioned to patients need careful consideration of the nature of the costs incurred – whether they directly or indirectly affect patient care. If costs are not known, can charges be an adequate proxy? Should local variations in cost be allowed (for example a high local wage index) or disallowed (due to inefficient use of resources)?

It was the basic and explicit design of DRGs, as iso-resource groups, that made them an attractive scheme for funding according to case mix. Other ways of paying hospitals such as on a per diem basis failed to recognise differences in the intensity of treatment and provided unwanted incentives to keep patients longer.

This chapter will examine the experimental use of DRGs in the state of New Jersey and, later, for the federally-funded Medicare prospective payment scheme (PPS). In order to compare these schemes, the forces that led up to the development of slightly different solutions to similar problems will be described in parallel. Both schemes have been implemented by legislation and this has inevitably resulted in much comment and criticism.

Evaluations have been made of how closely the reality of prospective reimbursement matched the objectives of the legislators and the objections of the critics. These assessments of success will form a basis for deciding the future of cost containment in the US and are of relevance to many other countries who are considering funding hospitals by case type. Such a decision would then result in detailed cost finding exercises like those described in Chapters 6 to 9.

Background to prospective payment

In its annual report reviewing 1980, 'A prospective reimbursement system based on patient case mix for New Jersey hospitals', the New Jersey State Department of Health claimed that:

> ...the evolution of hospital reimbursement in New Jersey has been marked by constant progression towards greater equity, sophistication and rationality.[1]

The state had a long history of intervention in the health care market, even before allegations of hospitals' bureaucratic malpractice in 1974. Early regulatory plans included putting a ceiling on individual payments to hospitals, introducing controls on capital expenditure and involving hospitals in voluntary budgetary reviews. In 1975, the Standard Hospital Accounting and Rate Evaluation system (SHARE), which set payment rates for similar hospitals and certain types of patient, was implemented. However SHARE did not adequately allow for differences in patient mixes or fully incorporate capital and overhead expenditure. In order to remedy this a new state law (S446) was enacted in 1978 which provided for:

equitable payments by *all* payers;
payments for uncompensated care;
working capital needs including maintenance/replacement of equipment;
establishment of a hospital rate-setting commission.

The DRG scheme was later chosen by the commission as the framework for setting rates – the means of achieving equitable payments to all hospitals by all payers. This was seen as a significant improvement on per diem reimbursement as it provided more appropriate units of output and incentives to achieve the goals of equity, sophistication and rationality. The State Department of Health recognised, however, that not all inequities would be eliminated by this legislative apparatus, and that an appeals process was expected to contribute to further improvements and refinements.

These developments in New Jersey were watched with interest at a national level, and two years later, in December 1982, a report was submitted to Congress by Schweiker[2], the secretary of the Department of Health and Human Services, recommending a hospital prospective payment method to cover Medicare patients in *all* states. A large part of health expenditure draws on federal funds, in particular the Medicare scheme which covers many citizens over 65. Hit by inflation at three times the national levels on other goods, the scheme was expected to go into debt in 1987. With only five years to run before it could no longer provide care for people insured, drastic action

was needed and the high-cost acute hospital was the obvious target. Again, criticisms of former systems of hospital payment highlighted the problem of incentives. In his report[2] Schweiker wrote:

> In cost-based reimbursement, hospitals are paid essentially whatever they spend. There is no incentive for hospitals to operate more efficiently since all allowable costs are fully reimbursed. In fact cost-based reimbursement encourages just the opposite behaviour. The larger a hospital's costs, the larger will be its Medicare reimbursement. Thus, there exists an incentive to spend because the current system provides no incentive to save.

Reviewing experience with different reimbursement systems the Department of Health and Human Resources saw much to recommend the mandatory, fixed-price reimbursement of New Jersey hospitals. The scheme was feasible, it provided hospitals with incentives to control costs, and was thought to be fair to patients, payers, and hospitals. It also showed no real evidence of gaming or 'DRG creep' to receive higher payments, or to increase admissions. The department therefore proposed a prospective payment scheme for Medicare patients based on DRGs.

Congress accepted the department's recommendations and President Reagan signed the Social Security Amendment of 1983. With Medicare covering 40 per cent of all health spending, such an important change was unlikely to leave other payers unaffected. One of the consequences of fixed payment rates and the capping of federal spending was thought to be a shift of costs towards other payers. Thus the hospital that spent more than the reimbursement rate might recoup its losses by increasing charges to non-Medicare patients.

Before describing the way in which these two scenarios developed, it should be noted that three other states used DRGs to a lesser extent in setting reimbursement rates.

The Georgia Alternative Reimbursement Scheme used case mix as one factor influencing an overall reimbursement ceiling. There were also many additional payments to meet physician charges, education programmes, kidney acquisitions, and so on. New York used DRGs as the framework for establishing limits for reimbursement on length of stay and routine and ancillary costs, with allowances made for hospital type. Maryland adopted several variants of the DRG classification as the basis for payment, from which a hospital effectively chose the one that suited it best. These variants were the full DRGs, the major diagnostic categories, DRGs or MDCs by payer, or the primary diagnosis.

Legislation and rate setting

Rate setting in New Jersey

Determining average costs for a DRG, and so deriving a rate of reimbursement for the state of New Jersey, required an enormous data collection exercise which many hospitals found difficult. All patients in all acute general hospitals were to be covered by the state-wide scheme.

Hospital expenditure from all quarters – from inpatients, outpatients and same-day surgery, working capital allowances, depreciation of major

Table 3 New Jersey cost reporting and allocation procedures: direct patient care costs only

S-446 DRG cost centers	*Type and component of costs*	*Basis for allocating cost center costs to DRG*	*Reason(s) used*	*Comments*	*Other units of service used*	*Comments*
I. Direct patient care cost centers						
1 Medical surgical (MSA)	Nursing salaries[1], non-salaries[2] (supplies, contract services, lease costs, depreciation, price level allowance and other expenses)	Patient days	Study results regarding basis for allocation of nursing costs not available at time of implementation	The use of patient days to allocate routine (nursing) costs assumes that : 1) the level of nursing care required for all patients is the same (ie, obstetric vs. ICU patients); 2) the amount of nursing time required per day is the same for each day of the hospitalisation (ie, the same on the first as the Nth day); 3) the age of the patient does not affect the amount of nursing time required (ie, 2 patients with same illness – one 25 the other 60); 4) the kind of nursing care is the same in all hospitals (ie, does not take into account different types: primary, team and so on) The use of charges to allocate costs assumes that the charges are related to the cost which preclude the use of other alternative allocation bases.	1–8 Relative intensity measures (RIMs); Joint nursing analysis pilot study; Joint nursing performance analysis pilot study	1–8 RIMS developed to assess resource consumption of hospital services on a cost per case basis. Study results plan to be available for use in the 1981 rate setting process.
2 Obstetrics (OBS)		Patient days				
3 Paediatrics (PEDS)		Patient days				
4 Psychiatric acute care unit (PSA)		Patient days				
5 Intensive care unit (ICU)		Patient days				
6 Coronary care unit (CCU)		Patient days				
7 Neo-natal intensive care (NNI)		Patient days				
8 Newborn nursery (NBN)		Patient days				
9 Emergency room (EMR)	Nursing salaries, non-salaries, physician fees	Charges			9 Number of visits	9 Determine cost per visit
10 Clinic (CLN)	Nursing salaries, non-salaries	Charges			10 Number of visits	10 Determine cost per visit
11 Home Health Agency (HHA)	Nursing salaries, non-salaries	Charges				
12 Anesthesiology (ANS)	Nursing salaries, non-salaries	Charges			12 Minutes – California Relative Value Units (RVU)	12 More accurate measure available from medical records
13 Blood bank (BBK)	Technician salaries, non-salaries, physician fees	Charges			13 Pints of blood whole (packed)	13 More accurate measure available from medical records
14 Cardiac catheterisation (CCA)	Technician, salaries, non-salaries, physician fees	Charges			14 Number of catheterisations	14 May not be as refined as charges
15 Delivery (DEL)	Nursing salaries, non-salaries	Charges			15 Operative code-weighted by CA-RVU or ANS	15 No additional data to be collected
16 Dialysis (DIA)	Nursing salaries, non-salaries	Charges (PHM)			16 Number of treatments (Hemo or Peritoneal) weighted CA-RVU	16 Information currently available

17 Drugs sold to patients (DRU)	Technician salaries, non-salaries			
18 Electrodiagnosis (EDG)	Nursing salaries, non-salaries, physician fees	Charges	18 Number of EKGs, EEGs – weighted by Statewide avg. cost of each	18 Information available from medical records
19 Laboratory (LAB)	Technician salaries, non-salaries, physician fees	Charges (BBK & LAB)	19 College of American Pathology RVUs	19 More accurate measure of costs
20 Nuclear medicine (NMD)	Technician salaries, non-salaries, physician fees	Charges	20 American College of Radiology RVU	20 More accurate measure of costs
21 Medical surgical supplies (MSS)	Technician salaries, non-salaries	(CSS)	22 ORR minutes or CA Medical Assoc. RVUs	22 More accurate measure no additional data needed
22 Operating room & recovery (ORR)	Nursing salaries, non-salaries	Charges	23 Time in half hour intervals or RVU	23 Likely availability of data due to charge structure
23 Physical therapy (PHT)	Technician salaries, non-salaries, physician fees	Charges	24 American College of Radiology RVUs	24 More accurate measure
24 Radiology (RAD)	Technician salaries, non-salaries, physician fees	Charges	25 RVUs	25 More accurate measure
25 Respiratory therapy (RSP)	Technician salaries, non-salaries	Charges	26 American College of Radiology RVUs	26 More accurate measure
26 Therapeutic radiology (THR)	Technician salaries, non-salaries, physician fees	Charges		

Nursing centers for first years of implementation were combined into Acute (ACU), Intensive care (ICU) and Newborn nursery (NBN).
All components of non-salaries may not always apply to all cost centers.

Source: Economic and Financial Analysis. DRG evaluation, Vol II. Princeton, Health Research and Educational Trust of New Jersey, 1984

equipment and uncompensated care (which refers to bad debts) – was fed into the rate-setting exercise. To these were added increases to cover management charges and inflation in the health care sector. Relatively simple assumptions were made for apportioning costs from cost centres to the diagnosis related groups, as shown in Table 3.

The payment to the hospital was calculated from two components, the state-wide cost and the hospital's actual cost. This was done by weighting standard or geographical area according to how a hospital's actual costs varied. If for a particular hospital the patients in a group were much more heterogeneous than the area pattern, then their payment rate was weighted towards the standard or area cost. If on the other hand the group was homogeneous, or a good group, then the payment was weighted towards the hospital's actual cost. In this way, higher costs incurred by a teaching hospital, for example, were partially met.

In order to be successful the scheme was designed to create incentives for hospitals to manage themselves more efficiently. These incentives were seen simply in terms of profit and loss, and in the competitive US health care market it was thought satisfactory to allow a hospital to decide how it might respond to these incentives – for example, by trading-off short-term losses with long-term improvements in financial stability.

Rate setting for Medicare

When the proposals for Medicare prospective payment[2] were put to Congress at the end of 1982, they were based on several observations about the experience with DRGs in New Jersey. One of these was that

> . . . successful systems require a firm legal basis, strict enforcement and a lack of escape mechanisms (e.g. control of volume, gaming).

The report therefore set about recommending a system which created

> financial incentives that encourage hospitals to restrain the use of resources in providing inpatient care.

The full list of objectives is given in Table 4.

The Tax Equity and Fiscal Responsibility Act of 1982 had already required case mix to be incorporated in Medicare reimbursement, and this was quickly followed by more detailed legislation in 1983, referred to as the revised section 223 limits. The new 467 DRG scheme (as opposed to the earlier version of 383 groups) had emerged from Yale and was used as the basis for setting DRG relative weights. The DRG assignment rules used for Medicare reimbursement differed in that *any* secondary diagnosis was treated as a significant complication or comorbidity, and the first surgical procedure on the patient record, rather than the most resource-consuming one, was used for assignment.

In summary the firm legal basis set out for national prospective payment had the following features.

> All hospitals treating Medicare patients were included in the scheme (except psychiatric, long-term care, rehabilitation and children's hospitals, and excluding those states with an acceptable alternative scheme for

payment by case mix) from 1 October 1983.

Relative cost weights were calculated for each DRG with adjustments for urban/rural areas and with a four year phasing of regional to national rates.

Cost weights were calculated from historical Medicare records and increased by hospital inflation plus one per cent.

Actual rates were calculated from expected workload with a ceiling to achieve 'budget neutrality'.

Capital-related costs, medical salaries, medical education and outpatient costs continued to be paid on a reasonable cost reimbursement.

Cases with exceptional lengths of stay or costs for a DRG benefit from extra reimbursement, up to a maximum of six per cent of total reimbursement.

Hospitals were required to contract with professional review organisations (PROs), who would monitor admissions, re-admissions and quality of care.

How the schemes compared

The national Medicare scheme and the New Jersey state-wide scheme differed in broad scope and in detail. The national scheme used the revised and substantially improved DRG classification, allowed for regional differences in costs such as wages, made allowances for payment of exceptionally high cost patients and established a mechanism for peer review. It was also different from New Jersey's prospective reimbursement in that it did not include all payers, and it excluded outpatient costs, medical salaries, medical education and working capital allowances which were all treated as direct lump sum payments. In the case of medical education, teaching hospitals received up to twice the amount they would have received under the former legislation. This was thought to be in recognition of the fact that teaching hospitals' reimbursement would otherwise be considerably reduced under the new scheme[4]. It also serves as an illustration that rate-setting was not always straightforward but influenced by political compromise.

Initial reactions

As US hospitals realised the extent to which the new legislation would affect them, a wave of comment, criticism and concern built up. The many articles and papers in which these are expressed make up a large part of the DRG bibliography, and are difficult to separate from what hospitals actually experienced under prospective reimbursement.

Voices raised in protest were anticipating ethical problems where the interests of the patient might be subjugated to those of the institution. The preface to Grimaldi and Micheletti's book[5] states that it had been demonstrated that hospitals could survive under illness-specific repayment, but it had not been demonstrated whether the method was cost effective or the best way to measure case mix. The writer added:

Furthermore, the impact of case-based reimbursement on quality and accessibility remains unknown, and marked controversy prevails regarding whether the groups do, as claimed, contain patients with homogeneous resource consumption patterns.

Table 4 Medicare prospective payment scheme

The system must:

- Be easy to understand and simple to administer.
- Be capable of being implemented in the near future.
- Ensure predictability of government outlays.
- Help hospitals gain predictability of their Medicare revenues.
- Establish the Federal government as a prudent buyer of services.
- Assure that Medicare expenditures for inpatient hospital services are no greater than those that would be incurred if the present system of retrospective cost reimbursement with limitations were continued.
- Provide incentives for hospital management flexibility, innovation, planning and control.
- Reduce the cost reporting burden on hospitals.
- Continue to assure beneficiary access to quality care.
- Prohibit hospitals from charging beneficiaries anything for covered services other than statutorily defined coinsurance and deductibles as applied to covered services.

Source: Schweiker RS. Hospital prospective payment for Medicare (report to Congress). Washington, Department of Health and Human Services, 1982.

Pursuing the question of whether appropriate care will be available, Sloan[6] described a study that tested the theory that 'regulation of hospital prices reduces the quality of hospital care and may increase the quantity, which found that although there was some support for the predictions, empirical evidence was still inconclusive.

Observers mostly agreed on the increased demands that would be made on medical records and data processing and the importance of relationships between administrators and clinicians. In Grimaldi and Micheletti's guide[5] these two themes are developed in considerable detail with chapters devoted to the financial planning and billing functions, and the necessity of a team approach with all hospital personnel synchronising their activities.

Three principles of management control against which DRG reimbursement might be judged have been outlined by Young[7]. These are the need for the system to be fair (a clinician is not penalised for making good decisions); for it to maximise the areas of common purpose between the sections of a hospital (maximise goal congruence), and to match responsibility to areas of controllability. It is suggested that a system such as payment by DRG which is imposed on clinicians who have little control of departmental costs, does not satisfy the principles of good management control systems.

The future of surgical practice is called into question by some observers. For example it has been suggested that innovation in high-priced surgical specialties will be attractive[8]. In addition, if such surgery is performed on patients whose condition was previously inoperable there is an additional bonus for placing the patient in a higher-reward DRG. If surgical practice is so influential on hospital revenue a new tension will be created between doctors and hospital managers.

Large variations in the admission threshold for different clinicians have been observed[9]. Up to 3.5-fold differences in rates of admission for hysterectomy were noted and it was concluded that prospective payment by fixed rates would only work if hospitalisation rates were controlled. As discussed in Chapter 1, case mix measures generally assume health status on admission to be the same for all patients in a group; large differences in admitting practice may therefore reduce the homogeneity of the group.

Probably Stern and Epstein[10] summarised the popular view by criticising a system which does not include a range of factors which may affect cost (severity of illness, socio-economic and other patient characteristics); is based on *average* costs across institutions and excludes costs associated with capital and education. They predicted the following effects:

an adverse impact on quality and access;
uncontrolled increases in volume unless marginal costing is introduced;
cost shifting onto the bills of patients in other insurance schemes.

It was also clear that changing the structure of incentives would bring about changes in hospital management and organisation. Management consultants in automated financial planning systems were not slow to produce rate optimisation models and the like, geared towards maximising payment. Their use would enable hospitals to consider the possible rates of reimbursement before deciding whether to admit a patient or how to record his hospital stay.

Multi-dimensional matrices have been enthusiastically constructed within which hospital finances might be managed under the new rules, and health economists have drawn profit-maximising hypothetical marginal cost curves to illustrate how the system can be gamed and the bottom line improved. Quality and completeness of medical records was also regarded as essential, especially when a simple case with no data may attract an average payment which is greater than if the abstract was complete.

Effects of prospective payment

Now the dust is settling after the upheaval in hospital payment, and the new arrangements and their after-effects are slowly coming into view, we can ask: 'Was prospective payment by DRG successful?' A substantial assessment, contained in several volumes, of the early years of the New Jersey scheme is available from the Health Research and Educational Trust (HRET) [3, 11, 12].

New Jersey

The main issues of interest were whether the payment system was fair to patients, hospitals and payers across the different types of institution; what impact it had on the organisation of hospitals; and whether it had helped reduce, or at least slowed down, accelerating health costs.

In order to investigate whether the payment scheme was fair to hospitals, the question of how they gained or lost revenue following the introduction of the scheme is addressed in Volume II of the HRET Report[3]. It was found that there were no overall differences in average DRG costs between hospitals. In other words large hospitals, with high turnover or high occupancy, were no

cheaper, and teaching hospitals or inner city hospitals were not systematically costing more than the New Jersey average. The researchers were quick to point out that although the system appeared to be fair, there may be other patient-specific characteristics which explain higher costs but were not available for analysis.

In this broad negative finding lie a number of interesting discoveries. For instance, there was evidence in a few DRGs of both more expensive specialist treatment in referral centres, which may be described as providing better care, and also cheaper costs for hospitals with high workloads. Another feature was the domination of large hospitals' costs over the statewide averages, and the fact that these were often teaching hospitals.

The second question in HRET's assessment was how prospective payment had affected the organisation of New Jersey hospitals. Increases in computerisation, and the improved status of medical records departments and management information systems, were immediately obvious. The net effect was that hospital operations became more decentralised. The lines of communication were strengthened between medical records and clinical staff. Conversely, links between medical records, finance departments and administration were almost completely satisfied by access to computerised information, and required little direct contact. Medical staff were involved in the need to improve the completion of patient records and in taking advantage of the institutional incentives to lower length of stay and use resources more efficiently. Programmes of clinical and management budgeting were planned but not put in place at the time of HRET's study in 1983.

The hospitals that performed most efficiently (became most profitable) were those which provided hospital-wide training in the roles and responsibilities of departments under the new funding mechanism. They also developed a new post of DRG co-ordinator with the responsibility of providing and interpreting DRG-based information. The most successful hospitals often had a multi-disciplinary committee to monitor the hospital's progress under DRGs and investigate any large differences between costs and reimbursement. They also tended to have effective data processing systems, and were willing to hire new staff to cope with the changes in information.

It was expected that with incentives to discharge patients earlier nursing requirements per patient day would increase. Although it has been noted that patients were often sicker on both admission and discharge[13] the nursing hours and staff levels had not increased. There were, however, reports of a new emphasis on discharge planning and an extra workload being placed on social workers.

Whether the objective of rationalising services and thereby slowing the trend of increasing hospital expenditure has been achieved, is not clear[14]. Certainly hospitals have gained financial solvency under DRGs, and even smaller hospitals have access to a share in the capital allowances. Against this, the increased cost of the data requirements has been heavy and has obscured the assessment of overall cost containment. However, the New Jersey system is a regulatory one and by definition controls the cost per case, if not the total expenditure.

Medicare

If the New Jersey PPS has proved difficult to assess, how much more difficult has it been to isolate the effects of the considerably larger Medicare scheme. Changes in national levels of hospital activity have been noticed, some as predicted, some contrary to expectations, and others which conflict with one another! Inside the hospitals there has been detailed scrutiny of all expenditures which might no longer be covered by prospective pricing. The overall effects of the national scheme will be briefly described under three broad headings.

Activity The New York Times[15] reported the steepest decline in hospital use in at least 20 years, due largely to cost containment policies. Admissions were three to seven per cent down and a day was knocked off the average length of stay for Medicare patients. Accordingly the number of beds and the occupancy rate fell, so that only two in three beds were occupied on average. Shifts from inpatient to cheaper outpatient settings have been reported but have yet to show up in government statistics. Also responsible for cutting costs were the peer review organisations which refused payment for inappropriate care of 2.5 per cent of all admissions in 1984.

Staffing, equipment and use of consumables Staffing levels, for a long time considered as fixed, have been reduced to match falling utilisation, and even the ratio of staff/occupied beds has fallen with the introduction of PPS. Fairly aggressive cost-cutting exercises have also been seen in policies for purchasing supplies at competitive rates, and for operating strict controls on the purchase of capital equipment. The effects of PPS on the speed and extent of adoption of new technologies have been investigated by several researchers[16]. Effects on diffusion of technology in hospitals were found which indicated that an innovative cost-saving technology would be implemented sooner under PPS, but that the *extent* of implementation of new technologies would be constrained. Wage levels, set lower for rural areas, have been a source of debate, particularly if the labour pool is more limited in rural areas. It has been reported[17] that the urban-rural rate differences have proved the most troublesome and are due for elimination.

Administrative functions including medical records Recording of patients' diagnoses has increased markedly, with an apparent increase in complexity of case. This has been encouraged by the fact that the existence of certain complications and concurrent diseases will attract a higher reimbursement for some categories of case. The increase in recording levels is also in line with the greater emphasis now placed on the completion of medical records. Hospitals have realised the effect of statistics on their revenues.

Another knock-on effect of the Medicare reimbursement has been for hospitals to diversify into home nursing, specialised psychiatric services and so on, or to push for an increase in their market share of patients with other health insurance schemes. The latter opens up a range of possible competitive insurance arrangements, such as preferred providers, health maintenance organisations and other forms of innovative pricing.

An assessment of success

The extensive work of the Health Research and Educational Trust to evaluate the payment of DRG in New Jersey has included an assessment of its success. The objectives described earlier were brief and pragmatic, and therefore easier to assess than the list of goals Schweiker had drawn up for Medicare. In summary the New Jersey assessment team found that:

> Payment was equitable for a range of hospital characteristics. HRET was unable to test patient characteristics such as severity of illness or socio-demographic group.
>
> The rates were set to cover bad debts and working capital needs.
>
> Allowances were available for hospitals whose geographical situation, mix of patients, or immediate capital needs made especially heavy demands on their expenditure.
>
> Overall expenditure had not reduced or significantly changed the rate of inflation in the New Jersey hospital sector, which was lower than the national average.

Clearly much progress has been made in regulating and monitoring expenditure in New Jersey, with little evidence of deleterious effects of reductions in quality of care, cost-shifting or refusing to treat patients. However it must be said that these aspects have not been thoroughly studied.

The national Medicare prospective payment scheme was introduced specifically with the aim of cost containment. With a clear limit on total spending, from which prices were fixed to achieve budget neutrality, it was bound to succeed unless hospital utilisation increased dramatically. We have seen the results; utilisation has fallen and, not surprisingly, Medicare has reported its rate of inflation halved in the first year of prospective payment. As a consequence, it has been suggested that five to ten years will be added to the solvency of the Medicare trust funds.

Despite the restriction of funds, the teaching hospitals have been richly rewarded with the lump sum pass-throughs, and have received on average 50 per cent more per patient than other hospitals. On the other hand, payments for capital expenditure also allowed as pass-throughs have now been frozen and will not be subject to inflationary increases.

In assessing how well these schemes have operated, it has to be asked what would have happened without them, and, on a more practical note, what alternative scheme would have done better? It seems clear that the all-payer single-state experiment introduced far fewer tensions and adverse incentives than the single-payer, all-state scheme. The calculation of rates for a limited geographical area was less controversial than those to cover a country the size of the United States, but at the expense of the equal payment for same case type that national rate-setting achieves. The all-payer system also took the pressure out of the financial incentives to refuse admission or refer expensive cases, since other hospitals would be subject to the same rates and all shared the costs of unpaid care.

A possible advantage of the Medicare scheme is the generous teaching allowances which enabled teaching hospitals to subsidise the loss made on

sicker (and usually more costly) patients. If one wishes to encourage this care as appropriate, it will be necessary to quantify more clearly how these patients are more costly, and tailor the reimbursement accordingly. Another area in which Medicare deals more fairly with its hospitals is in the compensation for cases with extreme lengths of stay or high costs. Whereas New Jersey hospitals were automatically paid an average per diem cost for extreme cases in a DRG, the Medicare scheme was more restrictive in its definition and payment for outliers. There were controls on the percentage of such cases; also the peer review organisations had to authorise all such payments as meeting standards of appropriate care.

Both schemes have been subject to annual reviews and rate-setting. In New Jersey a number of factors (for example, the choice of weights, mean vs median costs, inflation factors) have been adjusted to create the desired incentives for hospitals. The Medicare scheme set up a prospective payment assessment commission to report to the US Department of Health and Human Services. In its 1985 report[18] it proposed changing rates by the hospital market-basket inflation less one per cent, plus an allowance for estimated case mix complexity. It also proposed some reassignment and recalibration of weights of high-cost surgery, such as pacemaker implantation, bone marrow transplantation and coronary angioplasty. Finally the commission acknowledged the need for weight recalibration on up-to-date data, and more sensitive regional pricing, influenced rather by the socio-economic characteristics of the population than the wage differences.

The future for prospective payment by DRG

Both methods described here have achieved considerable success in underpinning costs but have attracted critical comment. The president of the American Hospital Association is quoted as saying:

> Before the change a hospital had to spend a dollar before it could get a dollar back. Now in order to get a dollar, you first have to save one.[15]

That is a very important reversal in the motivation of US hospitals, and one in which DRGs will be claimed to have played a vital role.

As Iglehart has observed:

> . . . . looked at another way, hospitals also will have a new incentive to underserve patients – the same incentive that health maintenance organisations have by virtue of their fixed, prospective form of payment.[4]

This is a concern that has not been fully satisfied by the evaluations of PPS. Although quality of care is regulated by the fear of patients filing law suits and the assessments of PROs of appropriate care, it is sometimes traded for the much-vaunted goals of equity, rationality and efficiency. Aspects of patient satisfaction, quality of care and above all outcome of a stay in hospital have not received a great deal of attention, but while doctors' salaries are excluded from the DRG rates it seems unlikely that they will change their standards of practice to the detriment of the patient.

The enchanced feeling of individual responsibility for the financial well-

being of a hospital has been devolving through all departments with the introduction of fixed price reimbursement. At considerable extra expense on computing, but a small increase in staff levels, management information systems and channels of communication between clinicians and nurses, and finance and medical records departments, have become much better developed – mainly to the end of increasing profitability, or at worst avoiding a loss.

Areas of weakness are the difficulty in establishing fair rates for teaching and regional referral centres, for those with particularly heavy loads of low income patients, and for psychiatric and long-term care, where DRGs do not satisfactorily explain costs variations.

However health care services in many countries have cost containment problems and need to introduce ways of regulating expenditure. Clearly there is much to be learnt from these large-scale experiments in the United States with DRGs, and some encouragement is to be derived from the way the industry has responded to the challenge with positive results.

References

1 A prospective reimbursement system based on patient case mix for New Jersey. New Jersey State Department of Health. Third annual report, 1980.
2 Schweiker R S. Hospital prospective payment for Medicare (report to Congress). Washington, Department of Health and Human Services, 1982.
3 Economic and Financial Analysis. DRG evaluation, Vol. II. Princeton, Health Research and Educational Trust of New Jersey, 1984.
4 Iglehart J K. Medicare begins prospective payment of hospitals. The New England Journal of Medicine, 308, 9 June 1983: 14–28.
5 Grimaldi L and Micheletti J A. Diagnosis related groups: a practitioner's guide. Pluribus Press, 1982.
6 Sloan F A. Hospital rate review: a theory and an empirical review. Journal of Health Economics, 3, 1984: 83–86.
7 Young D W and Saltman R B. Medical practice, case mix, and cost containment. Journal of the American Medical Association, 247, 6, 1982: 801–805.
8 Smits H L and Watson R E. DRGs and the future of surgical practice. The New England Journal of Medicine, 311, 25, 20 December 1984: 1612–1615.
9 Wennberg J E, McPherson K and Caper P. Will payment based on diagnosis related groups control hospital costs? The New England Journal of Medicine 311, 5, 2 August 1984: 295–300.
10 Stern R S and Epstein A M. Institutional responses to prospective payment based on diagnosis related groups. The New England Journal of Medicine, 312, 10, 7 March 1985: 621–627.
11 Boerma H. The organizational impact of DRGs. DRG evaluation Vol. IV-B. Princeton, Health Research and Educational Trust of New Jersey, 1983.
12 Case-mix classification, data and management. DRG evaluation Vol. III. Princeton, Health Research and Educational Trust of New Jersey, 1984.
13 Feldman J and Goldhaber F I. Living with DRGs. The Journal of Nursing Administration, 14, 5, 1984: 19–22.
14 May J J and Wasserman J. Selected results from an evaluation of the New Jersey diagnosis-related group system. Health Services Research, 19, 5, 1984: 547–559.
15 Sullivan R. Decline in hospital use tied to new US policies. The New York Times, 16 April 1985.

16 Romeo A A, Wagner J L and Lee R H. Prospective reimbursement and the diffusion of new technologies in hospitals. Journal of Health Economics, 3, 1984: 1–14.
17 First Boston Corporation. Highlights of the Federation of American Hospitals' annual meeting, March 7–9 1985. (Special report).
18 Altman S H. Report and recommendations to the Secretary, US Department of Health and Human Services. Washington, Prospective Payment Assessment Commission, April 1985: 1–83.

DRG developments

4 THE INTERNATIONAL SCENE

Dr Jean-Marie Rodrigues

Introduction

Interest in DRGs and their applications has developed in various western countries outside the US, with different objectives – ranging from utilisation review, reimbursement and budgeting to management and planning.
When considering these approaches it is important to differentiate between DRGs simply as a patient classification scheme and their application to different problems in different countries. This distinction must be borne in mind while reading this chapter in order to understand why so many types of health services are involved in DRG experiments and why this means of measuring hospital performance is increasingly widespread, often in ways that are very different from those in the US.

Review of experiences in different countries

The raw material for this review comes, for the most part, from those countries gathered together by the working group of the coordinated medical research programme, 1985, of the Council of Europe, entitled 'Computerisation of medical data in hospital services including university hospitals[1].' Experiences from other countries have also been included.

France

The first large scale European DRG project was in France. The PMSI (Project for the Medicalisation of Information Systems)[2] was initiated in 1981 by Jean de Kervasdoué of the Ministry of Social Affairs and Health, with Dr J M Rodrigues as project director[3].

Its object was to find a tool which could relate hospital spending to the social objectives of the institution. In practice this meant the application of a classification of case mix which could describe the multiplicity of hospital products in terms of specific groupings of treated patients[4,5,6]. Such groups would recognise not only the patient's condition but also, where necessary, aspects of the treatment process[7].

In order to reach this goal, four major sub-projects were defined:

1 Test and develop a standard discharge abstract, RSS (Résumé de Sortie Standardisé).
2 Examine the variability in average length of stay which could be explained by case mix. The case mix classification, GHM (groupes homogènes de malades), would have to be defined in a way to be compatible with the information systems that were to be implemented.
3 To compute and analyse costs by both cost centre and hospital product, the DRG/GHM.

4 To develop the software needed for data collection, processing and analysis.

The project enlisted the help of more than 300 experts in various disciplines (organisation, epidemiology, statistics, computing, accounting, management and administration) through a system of committees and councils which monitored the work.

In cooperation with the Health Systems Management Group (HSMG) of the Yale School of Organisation and Management, the PMSI team has now completed the first three sub-projects and in 1986 the fourth was underway. The official regulations concerning RSS were published in October 1985. The DRG classification, using the French diagnostic and procedure coding systems, has been shown to explain a large amount of variability in length of stay in French hospitals, and the refined GHM definitions were published in the summer of 1985. The cost-finding model, producing detailed costs by GHM and by cost centre, has been tested in two hospitals with good results. The system is to be implemented, with refined and standardised accounting and activity information, in order to produce uniform cost reports (Guides de comptabilité analytique I, February 1985 and II, February 1986).

For the fourth sub-project, the GROUPER ('groupeur') software was scheduled for summer 1986, and the cost model software by 1987. This software is written in C language and can be used with all types of computer supporting Unix or Xenix operating systems.

Portugal

In 1983 the Portuguese Ministry of Health, with the support of various agencies and universities in the USA, began work on a project to improve the effectiveness and quality of its hospital system[8]. Specifically, the project intended to investigate:

Utilisation review
Quality assurance
Nursing care
Cost accounting by DRG
Budget models
Medical record information systems and data processing

In 1986 the project had reached the first stages of a DRG-based information system. These are:

1 Medical record summaries have been routinely produced in 16 hospitals.
2 The DRGs have been validated after mapping between the Portuguese coding scheme and the ICD-9-CM system used in the US DRG definitions. These groups have been found to be satisfactory in explaining the variability of length of stay in the Portuguese sample.
3 Costs per case have been computed in four hospitals.
4 Data are now being collected for medical record summaries using refined coding schemes for diagnoses and procedures which give the same level of detail as ICD-9-CM.
5 The next step will be to define a standard cost accounting structure for use

in all hospitals and integration of the cost and budgeting model into routine hospital management.

It is planned that all the programs will be fully operational by 1988.

The Netherlands

Some of the earliest work in Europe on DRGs took place in the Netherlands in 1977. However, it was not until 1980 that two research experiments began, in the University Hospital of Leiden and in a Tilburg general hospital, using the medical data from the centre for health care information (named SIG). The work involved extensive case mix analyses using DRGs, and amendments to the classification have been proposed based mainly on experience drawn from the Tilburg study.

Currently, there are two experimental projects on the feasibility and utility of patient classification schemes for internal management purposes only, not for external budgeting. One will study the possibility of adapting DRGs to suit the Dutch situation. The other aims to realise an integrated hospital information system to serve the needs of management and physicians. The project will also consider the suitability of alternative classification schemes and whether they would be better suited to the Dutch health system.

Ireland

In Ireland the problems of determining equitable funding levels have led the Department of Health to initiate a DRG project under the leadership of Dr M Wiley.

The Irish health care system consists, for the most part, of eight regional health boards with responsibility for providing hospital and community services to geographically defined populations. The health boards receive annual budgets on a prospective basis from the Department of Health. For historical reasons, the department also funds voluntary public hospitals, mainly major teaching hospitals which remain outside the health board structure, on a similar prospective basis. In Ireland, the development of resource allocation formulae based on some estimate of need or relative morbidity, for example through the proxy measure provided by standardised mortality ratios (SMRs), is hindered by the concentration of facilities in certain areas.

The first step in the DRG project has been to test patient discharge data to assess the feasibility of assigning records to DRGs. The mapping of the diagnostic and procedure codes to ICD-9-CM was carried out with the help of Yale's HSMG in 1985. The next step, planned for 1986, was to identify one or more pilot hospitals to develop a cost model by DRG. Beyond these stages, the research is concerned with the development of a methodology for budgeting hospitals at national, regional and institutional level. Such budgets would be part of a comprehensive management policy using a variety of measures on which to base resource allocation, the case mix component being dealt with by DRGs.

Belgium

In Belgium, much academic and research work has been done by Professor Blanpain and Dr Roger concerning the uses of DRGs and aspects of data collection, processing, analysis and validation[9]. The Belgian government has recently proposed a new method of financing hospitals based on a prospective budget determined by assessing separately the requirements of support services and medical services.

The support service component will be calculated mainly on bed-days provided, historical costs and the structural characteristics of the hospital. The budget for medical services will be based on a grouping of hospital types according to the similarity in diagnostic groups, procedures and nursing care. The ability to identify these features of individual patients suggests the minimum basic data set has all the information necessary to determine DRGs, although this application is not at present generally used.

Scandinavia

In Sweden, the Swedish Planning and Rationalisation Institute (SPRI) under the direction of Mr Thorsten Thor has initiated a project to evaluate the feasibility and the utility of DRGs in Swedish hospitals, mainly for incorporating in their planning process. They undertook collaborative work with Yale's HSMG in 1985 for grouping, case mix analysis, and DRG validation.

Finland, Norway and Iceland have expressed the desire to proceed along the same lines and were considering initiating projects in 1986.

Australia

On the other side of the world, the Australian health service authorities have embarked on several DRG projects following the successful two day workshop on the 'Potential applications of DRGs' held in Canberra in 1984.

In the state of Victoria, Professor Palmer (from the School of Health Administration at the University of New South Wales) has conducted projects into the mapping and grouping of diagnostic and procedure codes. The resulting DRGs were validated using data from all Victoria hospitals. Yale's HSMG have recently implemented their DRG cost and budget model in three Melbourne hospitals. Their purpose is to use costs by DRG initially to improve the hospital planning process and later to determine prospective budgets for hospitals based on case mix.

The New South Wales health authorities wish to evaluate a DRG based payment scheme (similar to the Medicare funding in the US) for private hospitals. Other states, both at regional (for example Sydney) and hospital level, have shown a great deal of activity and interest in using DRGs for different purposes.

Canada

In Canada, some work on DRGs was carried out in 1983 by the Hospital Medical Records Institute (HMRI) in Ontario on a 3.2 million record data base, and, in 1984, by the Association des Hôpitaux du Quebec (AHQ) in

Quebec. These studies dealt mainly with problems relating to diagnostic and procedure coding schemes which are discussed later in this chapter.

Switzerland

After a debate in the federal parliament in Bern, the Department of Public Health and Planning of the Canton de Vaud and the Department of Preventive and Social Medicine at Lausanne University, have proposed an intercantonal study on DRGs. This project, which is managed by Dr Paccaud, has tested the availability and content of medical record summaries. The collection of relevant management and medical data is being carried out with the aim of validating the DRG groupings by the end of 1986. These are then to be used in developing a hospital cost-finding model. The final report on the applicability of these methods to planning and resource allocation is due by 1987.

Italy

The Italian Laboratoria di Epidemiologia et Biostatistica of the Instituto Superiore di Sanita (National Health Institute), with the guidance of Professor Zampieri, embarked in 1985 on a DRG experiment in the Piemonte (Piedmont) region. Five hospitals (8000 beds in all) with good information systems, agreed to participate in this experiment. The project director, Dr Tarone, wishes to test the applicability of US DRGs in Italy.

Spain

The Spanish Departmento de Trabajo, Sanidad y Seguridad Social in the Basque region is ready to start on a test of DRGs, and possibly a project to look at hospital funding.

West Germany

In the Federal Republic of Germany, the Bundesministerium für Arbeit und Sozialordnung (Ministry of Labour and Social Affairs) plans to carry out a feasibility study on DRGs in some regions that have a wide variety of hospital types. The experiment is planned to start in March 1986 and to run for two years.

Austria

The Krankenanstalten Zusammenarbeitsfonds, within the Austrian Bundesministerium für Gesundheit (Ministry of Health), have agreed to carry out experiments in some regions in order to investigate the value of DRGs for improving productivity and cost control. The experiment is planned to start in 1986.

UK

The experience in the UK is covered in more detail in Chapter 5.

National problems addressed by DRGs

The variety of current interests and experiences with DRGs in different countries indicates that there are many different reasons for studying DRGs and their applications[10].

DRGs can be used as part of more wide-ranging reforms in the management and financing of hospitals. For example, in France the DRG project coincided with the enactment (January 1983) of a new financing law for public and private non-profit hospitals, the 'Dotation Bugètaire Globale'. This law replaces retrospective hospital funding based on computed costs per diem, with a prospective budget determined by historical trends and constrained by a cash containment policy set each year by central government. The law has also initiated a new style of management inside hospitals in order to give them the capability of adapting the resources available, to cope with differences in patient case mix, and changes over time or geographically.

Three levels of management have been established, for the hospital as a whole, for individual service departments (functional cost centres), and for the hospital's products, in terms of DRG/GHMs.

The management of this matrix organisation is based not only on the costs of each DRG/GHM, but on the relative contributions of individual resources – the components which make up the aggregate cost[11]. Thus the matrix can be viewed from a number of different perspectives, depending on the type and responsibilities of management. For example, a clinical manager is concerned with the mix of resources within each patient group. The manager of a service department will look at the unit costs within one cost centre.[12]

In order to achieve this, a patient classification system was required that could identify both clinical and financial aspects of each patient's care. If it was to be used in over 1,500 French hospitals of varying size, it had to have a manageable number of groups and be comprehensible to managers and physicians. The use of DRGs in France is planned to be quite different from the US. The emphasis will not be on individual hospital reimbursement nor an automated budgeting based on case mix. Instead they are intended to be part of a mandatory joint draft for resource allocation drawn up by the hospital, the agency of the central government and the regional agency of national health insurance. In this way both the complexity of cases and the productivity in resource utilisation for different types of case will become an issue in the budgeting process, and a focus for the internal management of the hospital.

The potential to use DRGs as a tool for resource allocation, albeit tempered by other considerations, can be seen in other countries, for example in Ireland. Budgets at a national, regional and institutional level would be in part based on a consideration of case mix; the remainder would be calculated according to different mechanisms.

The use of DRGs in planning is being considered in a number of countries, in Australia, Switzerland, Sweden and other Scandinavian countries. There is also, of course, the possibility of using DRGs as the basis for hospital reimbursement. In New South Wales, Australia, such a scheme is envisaged for private hospitals. Their use in determining hospital budgets is also being emphasised in the Federal Republic of Germany and in Austria.

In contrast to these multi-level approaches it is interesting to consider the situation in the Netherlands, where DRGs are seen as a tool mainly for use within the hospital. The Dutch were the first in Europe to explore the availability of data and the validity of DRGs in the early 1980s. It was at this time that new regulations for planning and financing the health service were being enacted. This legislation was intended to decentralise the planning process within the limits of national guidelines and to establish a central health charges board to oversee financing. Hospitals are paid on a prospective budgeting basis in accordance with national guidelines. The DRG or other patient classification scheme can therefore only be used for internal management of hospitals, clinicial divisions or quality assurance, but not in the determination of a hospital's budget.

Finally, DRGs can be seen as part of an information system. The development of annual reports on medical activity to be used for quality assurance and utilisation review is seen as a by-product of the French experiment. Similarly in Portugal the Ministry of Health has undertaken a comprehensive project to improve quality in acute hospitals and productivity in the whole system. The Portuguese are especially aware of the fact that they have the longest length of stay in acute hospitals in Western Europe and very large differences in the numbers and quality of staff from one hospital to another. However, they have few historical references for studying either patterns of quality or funding differences. They are therefore implementing a new information system for medical records, nursing care and cost accounting to be used as the basis for utilisation and quality reviews. The system will also be able to provide the basis for budgeting by DRG, though it is not intended to implement per case reimbursement at this stage.

Availability of data and comparability of coding schemes for DRG classification

Availability of data: Medical Record Summary (MRS)

One of the main consequences of DRG work has been the increased availability of the MRS in four European countries, Belgium, France, Norway and Portugal.

The situation in 17 countries is summarised in Table 5. In some countries, for example, Australia, Canada, the Netherlands, Ireland and the United Kingdom, data for the MRS have been successfully and progressively collected on a voluntary basis for various reasons – such as epidemiology, clinical research and the monitoring of services. In Norway, efforts to expand the MRS to all discharged inpatients seem to be related more to the objectives that were typically claimed by other Scandinavian countries, epidemiology and planning.

In Belgium, Portugal and France their ministries of social affairs are modifying hospital financing schemes to take into account diagnostic categories. Developments in the hospital discharge abstract systems based on the MRS are necessary in order to measure case mix.

This new use of medical record summaries might prove to be a strong

Table 5 Degree of availability of the MRS

Nationwide		*Not Nationwide*
In 1985	Beginning in 1986	In 1986
Australia	Belgium	Austria
Canada	France	F.R. Germany
Denmark	Norway	Italy
Finland	Portugal	Spain
Ireland		Switzerland
The Netherlands		
Sweden		
UK		

incentive to give greater coverage to all inpatients in those countries where summaries are not yet widely available.

The fact that a MRS is not available in a country does not necessarily imply that all its regions are without it. For example data are available in several regions of Northern Italy (among them Lombardy), the Federal Republic of Germany (such as Schleswig-Holstein) and in Spain (for example Catalonia). A nationwide agreement appears to be difficult to obtain in federal states (such as Austria, Germany and Switzerland), where regions have a strong autonomy.

Comparability of coding schemes

Table 6 shows that the International Classification of Diseases (ICD) is widely used for diagnostic coding, mainly in its ninth revision. There are however some alternative systems in use. Northern countries publish common international statistics (NOMESCO) and still use a Scandinavian version of the ICD-8 code. There are plans to move to a ICD-9-SC (Scandinavia extension) but dates vary (1987 in Finland, 1987 in Sweden and Norway, no date for Denmark).

Belgium and the Netherlands have adopted the ICD-9-CM version used by the US CPHA, though extensions of the ICD-9-CM have been implemented with additional digits to take into account new innovations and classification problems.

A major finding of the recent inquiry of the Council of Europe was the persistent difficulty in comparing operations and surgical techniques when most countries have their own coding scheme. This problem is largely due to the absence of the equivalent of ICD codes for operations and procedures, and to the inadequacy of the present experimental classification provided by the World Health Organization. Only Belgium uses the ICD-9-CM code. The Netherlands uses the HICDA code issued by the CPHA, but with Dutch modifications.

In those five countries where a hospital MRS has not yet achieved national coverage (at the bottom of Table 6), variations in individual coding schemes and standards were observed. For example, a hospital might use the ICD-9-CM for diagnoses, although this is not obligatory; hence the description 'recommended or suggested' in the table.

Table 6 Diagnostic and surgical procedure coding schemes in 1985

	Diagnostic Coding schemes	*Surgical Coding schemes*
Australia	ICD-9	I 9
Belgium	ICD-9-CM	ICD-9-CM
Canada	ICD-9	Canadian
Denmark	ICD-8-SC	Danish
Finland	ICD-8-SC	Finnish
France	ICD-9	Veska, I9, I9-CM, C
Ireland	ICD-9	OPCS (UK)
The Netherlands	ICD-9-CM	Dutch
Norway	ICD-8-SC	Norwegian
Portugal	ICD-9 +ICD-9-CM	ICD-9-CM + regional
Sweden	ICD-8-SC	Swedish
United Kingdom	ICD-9	OPCS (UK)
Austria	ICD-9 (suggested)	Austrian
F. R. Germany	ICD-9 (recommended)	VESKA, KDS, GMDS
Italy	ICD-8 and 9	ISTAT (Italian)
Spain	ICD-9 (recommended)	Spanish
Switzerland	ICD-9 (suggested)	VESKA

I 9: ICPM International Classification of Procedures in Medicine WHO
C: Catalogues des actes medicaux, Ministère des affaires Sociales (Paris)
OPCS: Office of Population Censuses and Surveys (London)

Using codes in DRG classification

Application of DRGs is being tried out or planned in the majority of western countries – 16 to 18 European countries answered the questionnaire of the Council of Europe enquiry – plus Australia and Canada outside the US.

The classification developed at Yale was based on the ICD-9-CM diagnostic coding scheme.[13] This contains an extra fifth digit to ICD-9 in order to give additional information about the degree and localisation of a condition. In accordance with WHO recommendations, most European states and Australia and Canada use, or are going to use, ICD-9 (the exceptions are Denmark, which uses ICD-8 with an extension, and Belgium and the Netherlands which use ICD-9-CM). For a limited number of items the ICD-9 code is less precise than ICD-9-CM; therefore the software currently available for assigning DRGs cannot be used without first modifying the data or the program. Different levels of precision for coding in different countries make international comparisons impossible and may restrict the wider implementation of this new type of approach.

The coding obstacles can be summed up as either the use of ICD-9 for coding of diagnoses, or the use of international, national or subnational classifications for procedures. Three kinds of solution have been adopted and a fourth can be envisaged. The strategies available and the advantages and disadvantages of each are shown in Table 7.

Table 7 Coping with incompatible diagnostic codes

1 *Map to ICD-9-CM*	
Examples:	France, Portugal (early stages), Ireland, Australia
Advantages:	The grouper is used, affording a maximum of comparisons, and the manner in which data collection is organised remains the same.
Drawbacks:	The validity of comparisons with groups using ICD-9-CM is reduced in some areas. On the other hand, if an international mapping table was to be drawn up, this would no longer be a disadvantage for all states using it.
2 *Translate DRG definitions into national codes*	
Examples:	United Kingdom, Canada
Advantages:	The manner in which data collection is organised remains the same.
Drawbacks:	If every state does this independently a great deal of work is involved and no international comparisons are possible. If redefinition is done at the international level, there is less work and comparisons are possible.
3 *Use ICD-9-CM*	
Examples:	Belgium, Netherlands, Portugal (later stages).
Advantages:	Direct use of the grouper, which affords a maximum of comparisons (including US).
Drawbacks:	The manner in which data collection is organised is changed in all but 3 countries, and makes translation necessary because ICD-9-CM is available only in English, French and Dutch.
4 *Adaptation of the 10th ICD Revision*	
The 10th ICD revision is scheduled to appear in 1990 and will raise the question of redefining the grouper on the basis of the new diagnosis codes. It could be of great interest for countries interested in DRGs to embark on a cooperative study to agree on one adaptation of ICD-10 for DRGs.	

For procedure codes, which are of paramount importance in the classification process, the study revealed a great diversity of schemes. Table 8 summarises the possible solutions to this problem.

These coding problems are of great importance for the application of DRGs. It has been shown that the national problems to be addressed by DRGs are very different from one country to another, and it is not surprising that the countries most actively involved in experimenting are the ones that have compatible diagnosis and procedure codes. These are either ICD-9-CM (Belgium, Netherlands and now Portugal) or schemes that allow a mapping of codes. On the other hand, countries still using ICD-8 for diagnosis have been delayed and most of them have decided to move to ICD-9 (Scandinavian countries except Denmark, Federal Republic of Germany, Austria, Italy). No country without a procedure coding scheme has embarked on such a project.

It looks as though the degree of involvement in DRG experiments is as much related to the availability of appropriate coding schemes as the potential aims of utilisation.

Table 8 Coping with incompatible surgical procedure codes

1 *Mapping to ICD-9-CM*	
Examples:	France, Portugal, Ireland, Netherlands, Australia.
Advantages:	Use of the grouper, which permits a maximum of comparisons and does not entail reorganising the manner of data collection.
Drawbacks:	There is not necessarily a great deal of work but there is a problem of the validity of comparisons. Most correspondences are simple but there are more difficult cases where ambiguities may occur either due to the nature of the procedure or its place in the hierarchy of procedures in the major diagnostic category.
2 *Translate DRG definitions into national codes*	
Examples:	United Kingdom, Canada.
Advantages:	No reorganisation of the manner of data collection.
Drawbacks:	For all states the work is considerable, and virtually no comparison can be made validly.
3 *Use of ICD-9-CM*	
Examples:	Belgium
Advantages:	Use of the grouper, which permits the maximum of comparisons, and procedure classification is clear in relation to the DRGs.
Drawbacks:	Reorganisation of data collection in all states except one, and the need for translation into most languages (since ICD-9-CM is currently available only in English and French).
4 *International classification of procedures for western countries*	
Advantages:	Such a procedure classification would make DRG assignment easier since the same grouper should be used for all countries.
Drawbacks:	Substantial work would be needed in harmonisation between different languages and technical practices. WHO appears at present to have decided not to include procedure classification in its 10th ICD revision.

Conclusion

This chapter has shown the interest in and experiences with DRGs in countries other than the US, where the classification scheme was born. The reasons behind these experiments and the intended uses of DRGs vary widely from one health service to another. DRGs appear to be a new tool for the scientific investigation of the process of providing acute hospital care. The knowledge that DRGs can provide is important to achieving the differing goals of hospitals – specifically improving the quality of care and ensuring optimal productivity.

The general availability of the appropriate data is increasing worldwide and DRGs are often the reason for this increase. The variations in diagnostic and procedures coding schemes are the principal obstacles to international comparison and the widespread implementation of DRGs. This must be considered an important issue for western health services in the coming years.

It is possible to say that although the uses of the DRG scheme vary from country to country, DRGs as a patient classification can be considered a universal tool for measuring case mix in acute care facilities.

References

1 Council of Europe. Computerisation of medical data in hospital services including university hospitals. Strasbourg, Coordinated Medical Research Programme, 1985.
2 Rodrigues J M. Le PMSI: Histoire et état d'advancement. Paris, XIè Assises Nationales de l'hospitalisation publique, 1985.
3 Fetter R B. Le projet Français sa spécificité par rapport aux principaux projets DRGs dans le monde. Paris, XIè Assises Nationales de l'hospitalisation publique, 1985.
4 Hornbrook M C. Techniques for assessing hospital case mix. Annual Review of Public Health, 6, 1985: 295–324.
5 Lave J R and Lave L B. Hospital cost functions. American Economic Review, 60, 1970: 379–395.
6 Luke R D. Dimensions in hospital case-mix measurement. Inquiry, 16, 1979: 38–49.
7 Fetter R B, Shin Y, Freeman J L, Averill R F and Thompson J D. Case mix definition by diagnosis related groups. Medical Care, 18, 2, (supplement), December 1980: 1–53.
8 Freeman J L, Fetter R B, Newbold R C and Mullin R L. Development of a DRG based payment system for hospitals in Portugal. Final report. School of Organisation and Management, Yale University, 1985.
9 Roger R F. Standardisation and harmonisation of medical data classification systems in Europe. Brussels, Commission of the European Communities, 1986.
10 Rodrigues J M. Utilisation des systèmes d'information médicale dan le monde et en Europe. Information hospitalières, numero special, Mai/Jun 1986.
11 Neuhauser D. The hospital as a matrix organisation. Hospital Administration, 17, 4, 1972: 8–25.
12 Fetter R B and Freeman J L. Diagnosis related groups: product line management within hospitals. The Academy of Management Review, 11, 1, 1986: 41–54.
13 The new ICD-9-CM diagnosis related groups classification scheme. HCFA Pub. No. 03167. Baltimore Md, US Department of Health and Human Services, 1983.

5 DRGs IN THE NHS

Dr Hugh Sanderson

Over the last fifteen to twenty years, major changes have occurred in the availability of information technology in the health services and in the needs of management for information. During this period a number of major influences have been at work to bring us to our present state, where we have not only a requirement for the measurement of case mix but also the technology to do it. This chapter explains the background to these developments in the NHS and describes the first stages in using DRGs in the UK.

Providing the technology

The last two decades have seen staggering developments in the power and availability of computing. Twenty years ago computers in the British health service were restricted to regional computer centres, universities and a handful of experimental projects funded by the then Ministry of Health. All too often this limited processing power was jealously guarded by data processing professionals, and patient related data came a poor second to payroll requirements. Since the information revolution, the computing power of those mainframes of twenty years ago now sits on the desks of many health service managers, together with the user friendly software needed to manipulate data files and present information flexibly and attractively.

Computing is of course no exception to Parkinson's Law and it can be confidently stated that programs will expand to fill the memory available, and data requirements will always be greater than the available mass storage. There is no doubt, however, that the storage and processing facilities available for the sole use of managers have expanded enormously, and to a point at which sensible manipulations of individual patient data are available at the touch of a few keys. With this kind of power readily available, the complex processing of large numbers of cases required to provide useful measurements of case mix has become a reality.

Developing a requirement:

a) *Resource distribution*

Up to the early 1970s resource distribution in the NHS was based on steady incremental growth from a historical starting point. This process served to entrench inequalities in resource allocation inherited from the creation of the NHS in 1948. At that time regional budgets were set on the basis of the cost of providing the inherited services. Consequently, regions around London, with ample provision of hospitals and expensive teaching facilities, attracted a much larger share of the resources than the less well-provided regions further north, resulting in substantial differences in per capita funding. Partly as a

result of the reorganisation of the health service in 1974, but also as a consequence of the development of regionalised specialty services and new medical schools, there was by the mid 1970s a growing appreciation of the degree of inequality and the impossibility of it being corrected with the existing funding technique.

These growing complaints from regions with low per capita funding[1] forced the DHSS to look for a new method of funding. The Resource Allocation Working Party (RAWP) report in 1976 devised a new method of allocation based on an attempt to estimate the needs of the population. A major part of this new formula used standardised mortality ratios (SMR) as a proxy estimate of morbidity (and hence need). Although the use of such measures attracted a great deal of criticism[2], at regional level it was widely accepted that any errors would be compensated for by the large populations involved. In any case there were substantial logistical problems in collecting data more closely related to morbidity, which have still not been resolved. At the local district level significant problems seem more likely to persist.

One potential source of error might occur in districts which have an excess of morbidity due to a specific cause, but which rarely leads to death (arthritis for example). It was argued that these cases could consume considerable amounts of health service resources, but not be adequately represented in the SMR figures used in resource distribution. A more substantial anxiety was the effect of inflows of patients into districts with teaching hospitals or other special services[3]. It has often been argued that patients flowing into teaching districts are more complex than those they export. These complex patients consume more hospital resources and the teaching district loses revenue if calculations are based on simple net flows of patients. This situation requires an ability to measure the case mix of the flows (and hence the expected resource consumption by case type) both in and out of the district, and use this measure to adjust the resource distribution.

b) *Management accountability*

A second strand of the developing need for case mix measures has been the increasing trend towards greater managerial accountability which was initiated by the reorganisation of 1974 and the development of the planning cycle, but was reinforced by the Public Accounts Committee's stinging criticism of the department's lack of management control of the NHS.[4] This attack led to the introduction of the review process in which regions are held to account by the department on an annual basis for their performance. In turn, districts are held to account by region in a similar annual review process. However, the initial reviews were conducted with little reliable and comparable information about how well regions or districts were performing. The need for such information spurred on the development of performance indicators as the basis for comparison of activity and output.[5] As with the RAWP exercise ten years before, there was a good deal of criticism of the indicators used and the way in which they were compiled.

One major problem was the concentration upon ratios of activity to resources (discharges per bed and so on) without reference to the quality or outcome of that activity. It can be argued that in many cases fewer items of

high quality activity might be preferable to more items of low quality. The problem of course is the lack of suitable measures of quality. Lack of suitable information was not so problematic for a second major criticism of the early performance indicators, which was the perceived absence of sensitivity to case mix. This meant that it was impossible for the indicators to make allowances for districts with more difficult or complex cases than the average. A partial solution came in the second generation of indicators, where it became possible to account for differences in the age and diagnostic mix of cases. However, other variables influence the resource consumption of individual patients and there is still a need for a more accurate characterisation of case mix in the standardisation process.

c) *Körner data sets and information strategies*

A third strand in the development of interest in case mix measurement was the review of health service information chaired by Mrs Körner[6]. This working party has identified the minimum data sets which could be considered sufficient to manage health services in a district. The collection and transmission of these data sets will be required by the DHSS within the next two years. The spirit of the reports contains an understanding of the importance of linking data about patient activity to costs, and thereby making sensible comparisons of efficiency. Although not explicitly stated, the identification of specialties in information on resource use should assist in more accurate costing. Furthermore, by using a measure of case mix it should be possible to compare the actual and expected costs of groups of patients.

Going beyond the Körner proposals there are strategic information plans by several regions[7] which again put much emphasis upon the ability to relate costs to activities. Ultimately, integrated information systems imply the ability to collect all the elements of care within an individual episode. Relating these items, be they drugs, procedures, investigations and so on, to the underlying stock control, accounting, manpower, and other resource management information systems, will allow actual costs to be built up for individual patients, not only in hospital care, but later on in the community as well. Measures of case mix in which patients are allocated to similar groups will then be required to compare the expected and actual costs for individual patients. Such refinements of the information may be helpful in examining variations in expected costs for particular units or patient groups, as well as being closely related to the next issue, that of clinical budgeting.

d) *Clinical budgeting*

Finally, a fourth strand in the development of case mix methods has been the growing awareness of a need to involve and motivate clinicians in the management of the resources which they control. For many years, clinicians have had few incentives to work more efficiently and they have had neither responsibility nor information on the resource implications of their decisions. Rewards through the merit award system are more likely to come about through participation in the committee structure than through productivity of clinical work, and peer pressure to be more productive is only exerted in the most extreme cases of inactivity. In many ways, the only motivation for

efficiency has been the clinician's self esteem, but without a feedback of useful information, even this has been of doubtful effectiveness.

In order to change this state of affairs, systems of clinical budgeting have been experimentally developed to give units and clinicians annual budgets[8]. In some of these experiments, incentives in the form of rewards for savings (or opportunities to redeploy underspending) on the annual budget have been incorporated as a natural consequence of the budgeting process. The rewards may include the provision of items of equipment, or additional clerical help. Setting budgets may prove difficult, for the historic budget can owe more to inefficiency than industry or complexity, and vice versa. Clearly, it is necessary to be equitable in setting budgets, otherwise clinicians who have been efficient in the past may be provided with smaller budgets than their less efficient colleagues. The system would then tend to favour the inefficient who had been provided with larger initial budgets which would be relatively easy to trim. A more equitable way of setting budgets is to examine the mix of cases dealt with in the past year (this is considered further in chapter 10) and, on the basis of that case mix, establish an appropriate budget. Clearly this does not deal with instances where the mix of cases is changing, but if the case mix measurement is sufficiently sensitive it should resolve the more obvious inequities.

Available measures of case mix

For these various reasons, interest in case mix has been developing in the UK for a number of years. Because of the commercial nature of its medical care and the need for case mix in making comparisons, interest in monitoring hospital efficiency in the US developed earlier than in the UK. The resulting case mix work, which started in the mid to late 1970s, was of two major types. One sought explanations for the variation in hospital costs by using regression techniques, with 'case mix' as one of the explanatory variables. Several ways of broadly describing case mix for this purpose were developed but in general, these broad descriptions did not take into account the clinical conditions of the patients, and their success was limited. The second type took a different starting point and attempted to build up a classification of similar cases based on medical judgement and the costs (or length of stay) of individual cases. Of all the projects that attempted this path, only the DRG team was even partly successful in reconciling the statistical and clinical requirements. The survival of DRGs was, of course, greatly enhanced by their adoption as a tool for prospective reimbursement of hospitals, despite the considerable controversy this has caused and the detailed criticisms levelled at them[9].

Given the considerable work already put into developing DRGs, it seemed sensible to use it as a first step in examining case mix in the UK; in particular, to determine to what extent US groupings were statistically and clinically sensible in the UK. Even to do this required a considerable amount of work in adapting the program for use in this country.

Adaptation

Although patients can be assigned to DRGs by manual coding, a computer is required for large numbers. The allocation program developed in the US

needed changes in its software for use in the UK. First, it had been written to run on IBM computers in IBM assembler code. IBM computers are not widely used in the UK health service and in university computer centres, so it was decided to rewrite the allocation logic in a widely available high level language. FORTRAN was chosen. Second, there were differences between the two countries in the types of information available on discharged patients. In nearly all instances, UK hospital discharge data, whether held as hospital activity analysis (HAA) at the local level or in the national hospital inpatient enquiry (HIPE), could provide most of the information required by the DRG allocation program. However, availability is less of a problem than differences in coding for diagnosis and operative procedures.

Diagnosis

Although the ninth revision of the International Classification of Diseases (ICD-9) has been in general use internationally since 1979, hospital discharges in the US are coded using ICD-9-CM, a clinical modification of ICD-9 containing a fifth digit extension which allows a greater specificity in allocating diagnoses. Usually, this is the only change, but in a few cases ICD-9-CM codes have been created which do not relate to ICD-9. For DRG purposes a problem arises where an allocation decision rests upon the fifth digit. If that is not present it is impossible to determine which of two or more DRGs a case should be allocated to. This is illustrated in the case of haemangioma in Table 9. What appears as a single condition at the four digit level is broken up into six categories, which map into four distinct DRGs, when the fifth digit is added. Particular problems were also encountered with coma and maternity codes.

Table 9 Fifth digit specificity – the case of haemangioma

ICD-9-CM		DRG	
228.02	Haemangioma intracranial	034	Nervous system
228.03	Haemangioma retinal	046	Other eye disease
228.00	Haemangioma NOS	144	Other circulatory disorders
228.04	Haemangioma intra-abdominal		
228.09	Haemangioma NEC		
228.01	Haemangioma skin	283	Minor skin disorders

Operative procedures

An acceptable international standard for classifying operative procedures has proved much more difficult to develop. Although an International Classification of Procedures in Medicine (ICPM) has been published by WHO, it is used neither by the UK or US. In the UK, a coding devised by the Office of Population, Censuses and Surveys (OPCS) is used (and is under extensive revision at present) whilst in the US a different classification forming part of the ICD-9-CM has been developed. Although there are similarities in the

structures of the two systems, there are great differences in the codings used and we in the UK have to translate the OPCS procedure codes in order to achieve correct DRG assignments. Since the procedures are similar across the two systems, translation is usually possible, although one area, orthopaedics, is particularly difficult. This is due to a difference of emphasis, with ICD-9-CM being more anatomically precise while the OPCS code focusses on the type of surgery. It has been necessary to rearrange some of the orthopaedic DRGs in order to reflect these differences.

Coding sequence

In the DRG allocation program, the sequence of operative procedures and diagnostic codes (if more than one) are crucial. For instance, a patient with a peptic ulcer and hypertension will be allocated to the gastrointestinal major diagnostic category (MDC) if peptic ulcer is listed above hypertension, but to the cardiovascular MDC if hypertension is listed above peptic ulcer. If the patient then has a gastrectomy, combining gastrectomy with a cardiovascular diagnosis would appear to be inappropriate. In order to overcome this problem, the program could look at all diagnoses and operative procedures and select the most appropriate pair. Since this would result in using secondary diagnoses on the patient record for assignment, it was not attempted.

A further requirement of the program is to select the most significant operation if more than one is recorded. At present this can be performed by looking at the DRG hierarchy, although the OPCS intend to incorporate such a refinement in the revision of their procedure codes.

Non-operative procedures

Not all procedures which can be coded have important resource implications; for example, there is a code for venepuncture. Patients having these minor procedures must not be allocated the same DRG as patients having more major procedures, and a set of 'non-operating room procedures' has been identified. The UK list is not exactly the same as it is in the US where, for example, '-oscopies' are done on an outpatient or office basis to a greater extent than in the UK. Consequently, the US list of non-operative procedures has been slightly shortened in the UK.

In summary, it was possible to adapt the DRG allocation program for use in the UK. Some changes had to be made, but these were relatively minor overall.

Testing the program with UK data

Once a UK version of the program had been created, it was possible to use it to assign DRGs to patients and test the resulting groups to see if they were homogeneous. For this purpose a 10 per cent sample of all discharges in England in 1979 was used (the HIPE sample), a file containing approximately 415,000 records. In carrying out this test, a number of specific problems came to light:

1 *Unassignable cases* About 1.5 per cent of the cases could not have a DRG code assigned to them because the fifth digit specificity is lacking in ICD-9. Examination of secondary diagnoses was not helpful in the vast majority of these cases and, unless arbitrary decisions on allocation to a DRG are made, they will have to be excluded. For the purposes of this study, they were excluded.

2 *Inappropriate surgical procedures* In the original version of the UK program, incorrectly sequenced diagnoses or operative procedures could not be allowed for, so a proportion of cases were assigned to the 'inappropriate operative procedure' groups (2.2 percent of the HIPE sample). This error was only detectable in surgical cases where the operative procedure served as a cross check on the diagnosis (and vice versa). This cross check is not possible in medical cases, even though transposition of primary and secondary diagnoses is likely to happen (perhaps more likely). Thus an error rate of five per cent over all medical and surgical cases would not be surprising. Indeed, it would be expected given the literature on coding errors in computer abstract data[10].

3 *Missing data* Where items of information required by the DRG allocation program were not available, cases could not be assigned. Missing diagnostic or operative procedure data is not a problem with HIPE data because special efforts are made to ensure completion of all fields. In the case of HAA data, however, much more serious problems might be expected to arise in that some cases (usually a non-random sample) will have no diagnosis or operative procedure attached. Only slightly less problematic is the allocation of the non-specific code 799.9, which indicates insufficient information in the patient's record to provide a proper ICD code.

Homogeneity of DRGs

Having assigned most cases to DRGs it was then possible to determine how homogeneous the DRGs were in relation to length of stay. The results of this exercise showed that in some DRGs a high degree of homogeneity was achieved, but this was not the case for others. In about 10 per cent of DRGs the distribution of length of stay was very unusual, being either extensively skewed or bimodal.

Examination of these DRGs showed that removal of some cases with specific diagnoses or procedures could improve the homogeneity in about half the groups. In many instances, however, there was no obvious DRG to move the subgroup to without adversely affecting the homogeneity of that group. In the other half, no subgroup could be identified.

These observations supported the general validity of the DRG concept in the UK, even though the statistical performance of the DRGs was less satisfactory than in the US. To an extent this must be due to differences in medical practice and tradition. The experience of US clinicians of concurrent review over the last few years is likely to have made them much more conformist in their practice than their UK counterparts, who have had very little experience of reviews or even available statistics. As discussed elsewhere[11], the lack of statistical homogeneity is not a major disadvantage in the

use of DRGs to monitor length of stay. The central limit theorem enables useful comparisons of important DRGs even where the distribution of the DRG is abnormal.

Description of teaching hospital case mix

As a first step in the use of DRGs however, it is useful to look at the mix of cases purely descriptively. Using the 1979 HIPE material as a base line, it has been possible to examine the case mix of some districts, the most detailed comparison having been carried out with one of the inner London teaching districts.

In this study for an inner London teaching district, hospital activity analysis data for four years (1979–82) was aggregated in order to provide sufficient cases for analysis.

Case mix was examined in four specialties (general medicine, urology, ENT and orthopaedics) and compared with the case mix of the national HIPE sample. It was easy to identify differences in the pattern of case mix in the teaching hospital and the examples in Table 10 are typical in showing that unusual/complex problems tend to be more common than expected in the teaching district. On the other hand, 'routine' DRGs are seen relatively less frequently than expected from the national averages.

Table 10 Mix of cases in general medicine

DRG		PERCENTAGE OF ALL CASES IN GENERAL MEDICINE	
		Teaching district	National
155	Stomach, oesophageal and duodenal procedures	7.5	1.5
395	Red blood cell disorders	4	1.5
403	Lymphoma and leukaemia	6	2
122	Circulatory disease with acute myocardial infarction	2.5	6.5
127	Heart failure and shock	2.5	4.2
14	Specific cerebrovascular disease	1.7	5.2

This confirms the intuitive impression that the teaching hospitals have more than their fair share of complex and difficult cases and reflects the fact that academic units gather reputations and patients in highly specialised areas. In the district studied, it could be demonstrated that the over-representation was largely due to flows in from outside the district. Table 11 shows the percentage of local cases for a number of selected diagnoses only a quarter of cases for the more complex DRGs (such as red blood cell disorders) came from local districts although for general medicine as a whole half the cases were local. For the more common DRGs, such as heart failure and shock, 80 per cent of the cases were local and there was very little inflow from outside.

Again, this confirms the intuitive impression of the way in which teaching districts operate and provides a useful quantification of the degree of specialisation and inflow of cases that a teaching hospital experiences.

Moving from this descriptive level to an analysis of length of stay and resource use will help to explain the implications of these differences in case mix. From there we can examine the effects of changing the case mix in various ways in order to increase or decrease the special activities of the teaching district.

Table 11 Flows of patients in general medicine

DRG		PERCENTAGE OF CASES FROM LOCAL AREAS
395	Red blood cell disorders	25
122	Circulatory disease with myocardial infarction	65
127	Heart failure and shock	80
	All DRGs	50

Conclusions

The techniques of using DRGs as a tool for describing and measuring case mix are now being put into effect and have gone some way to meeting the requirement for a routine measure of case mix. Considerably more work and experience is required, however, in order to develop DRGs as a robust tool and to gain experience and confidence in carrying out analyses with them. Since a case mix analysis is only a marginal extra cost in processing expensively collected data, not to use some method of case mix measurement is potentially very wasteful. DRGs offer benefits in combining classifications of both diagnosis and procedures, as well as including, where necessary, additional patient variables. No other system of case mix is now as available or as widely validated as the DRG technique. It would appear to be the logical line to develop.

References

1 Griffiths D A T. Inequalities and management in the NHS. The Hospital, 67, 1971: 229–233.
2 Forster D P. Mortality morbidity and resource allocation. Lancet, I, 1977: 997–998.
3 Akehurst R L and Johnson K W. Cross boundary flows of patients. Hospital and Health Services Review, 76, 1980: 334–336.
4 House of Commons. Committee of Public Accounts. Seventeenth report. Financial control and accountability in the National Health Service. London, HMSO, 1981.
5 Department of Health and Social Security. Performance indicators for the National Health Service. London, DHSS, 1984.
6 Department of Health and Social Security. Steering Group on Health Services Information. First report. London, HMSO, 1982.

7 Jackson F. Making it happen. British Journal of Health Care Computing. 3, 2, 1986: 13–15.
8 Wickings H I, Coles J M, Flux R and Howard L. Review of clinical budgeting and costing experiments. British Medical Journal, 286, 1983: 575–578.
9 Horn S D, Sharkey P D and Bertram D A. Measuring severity of illness: homogeneous case mix groups. Medical Care, 21, 1983: 14–31.
10 Butts M S and Williams D R R. Accuracy of hospital activity analysis data. British Medical Journal, 285, 1982: 506–507.
11 Sanderson H, Craig M, Winyard G and Bevan G. Using diagnosis related groups in the NHS. Community Medicine, 8, 1986: 37–47.

6 ATTRIBUTING COSTS AND RESOURCE USE TO CASE TYPES

James Coles

Introduction

Since its inception the DRG classification system has become inextricably linked with methods of cost control within health care systems and many commentators seem unable to separate the intrinsic features of the classification from those of the US Medicare prospective payment system[1] which has been instrumental in spreading the fame of DRGs; but DRGs do exist by themselves. The initial purpose for their construction lay outside a formal financial set up. It was intended they should assist clinicians and other hospital professionals by describing a variable set of 'products' of a health care system as well as the resources required in terms of nursing hours, theatre minutes and so on. Such details would aid self-audit by clinicians, utilisation review and other approaches to quality assurance, and provide a useful basis for strategic and operational decision making.

It is fair to say though that a prime attraction of DRGs to health service managers lies in the claim that they are 'iso-resource' groupings of patient types, although the distinction between 'iso-resource' and 'iso-cost' groups must be firmly maintained. In their initial development of DRGs the researchers at Yale only sought to produce groups that 'had statistically stable distributions of resource use'[2]. They did not claim to have attached cost values to each group. It is often this last step – from a protocol of resource use attached to case types to the production of a monetary value for each DRG – that produces ambiguity, uncertainty about validity, and conflict between various disciplines or levels within a health care setting. Greater flexibility in the application of costs and more attention to the behavioural effects of particular costing methodologies could assuage some of these difficulties.

This chapter, while reviewing some of the issues around the attaching of costs to DRGs as well as the various approaches that have been made in the US and elsewhere, seeks to highlight the benefits that can occur if management firstly separates the resource use associated with case type from the costing methodology and then carefully considers the appropriateness of the costing methodology used.

Issues of costing

As has been mentioned in earlier chapters, and above, one of the criteria used in the development of DRGs was the concept that patients falling into a particular group should be fairly homogeneous with regard to resource use. Initially it was assumed that length of stay was a reasonable proxy for resource use and homogeneous groups were formed with this as the dependent variable. The Yale group then repeated their work using financial

data and found that the previous groupings provided reasonably satisfactory results when patient charges in the US were used as the dependent variable in place of length of stay. However, it is readily acknowledged[3, 4] that charges do not correlate well with the true cost of particular case types. The construction of charges is not always based on a 'cost-plus' approach, but on relating to competitors within the limits that a particular market will bear. This tends to have a differential effect across case types with the more 'marketable' types of care, such as maternity, effectively subsidising some of the less attractive types. With the introduction of prospective rate-setting, hospitals in the US are becoming more cost conscious since they are less able to influence the 'price' of a particular case. Much attention is now being paid to accurate accounting for costs.

However, is it costs by DRG that we are really interested in or, as in the original definition, some measurement of resource use by case type? The answer will often depend on the position from which the question is being asked. Possessing a resource profile across all DRGs itemising the demands each case type will put on pathology, operating theatres and so on, will enable the heads of these services to determine the effect of case mix changes on the use of fairly fixed resources, such as large items of equipment and established staffing. It would only be necessary to translate this to cost when demand exceeds available resources and one type of resource has to be traded off against another in the same budget, or with other budget holders. On the other hand, at regional or central government level strategic decisions cannot take account of the minutiae of local effects and are, rightly or wrongly, more often concerned with aggregate financial figures. Here too, it is because of the need to consider together a number of disparate resourcing issues that financial elements such as costs are required. Assessments of efficiency, for example, patients per bed or prescriptions dispensed per pharmacist, do not need to introduce a £ or $ sign until the issue is one of trading off resources between different interest groups – be they professional, specialty or geographically focused.

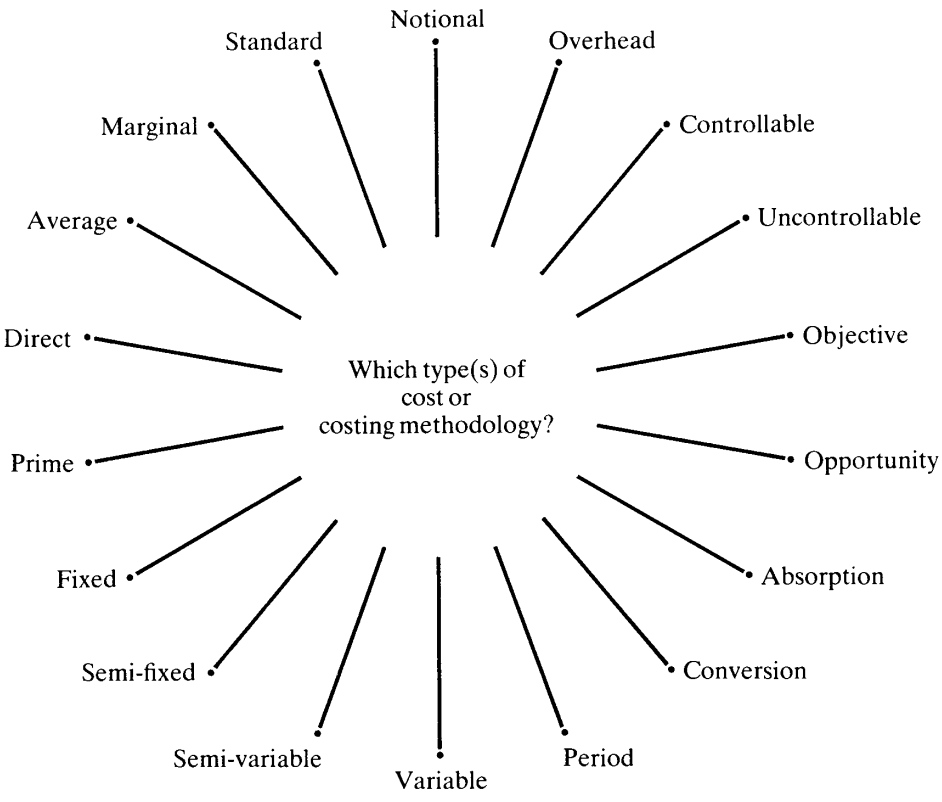
Of course, having said that it is unrealistic to imagine health care managers in the late 1980s not being interested in the cost consequences of changes in the case mix treated. In many instances their interest will be fully justified in order to assess what opportunities are open to them to provide the best possible care within the resources available. Before starting out, though, on a somewhat mechanical approach to costing 'everything that moves', managers need to consider what sort of action they are expecting to take or what decision they will be making on the basis of this information. Addressing this question will often focus managers' attention on whether there is a need for cost information and the level of detail required. Gwyn Bevan examines this question further in chapter 8 with particular reference to developing a local decision support system. The next four paragraphs set out what needs to be considered before embarking on costing DRGs.

Figure 4 shows a long (though not exhaustive) list of cost types and serves to highlight the truth of the phrase that there is no one true cost. Depending on whether you are interested in knowing the additional cost of treating another patient of case type X or retrospectively assessing the cost of one of a group of previously treated patients, will determine whether you are

interested in marginal or average costs. The importance of such differences is gradually becoming accepted in the health service but there remains a tendency to be somewhat offhand about it. Top down cost allocation models which apportion costs incurred in an aggregate cost centre to 'lower' levels according to a formula, or on some fairly arbitrary basis, may be appropriate for strategic decision making but will not necessarily be acceptable at operational level. In the NHS it would seem unlikely that DRG cost estimates arrived at by a top down approach could be used satisfactorily in management budgeting discussions with clinicians at local level.

The sophistication with which DRGs are costed must again depend on intended applications. An aggregate cost for a case type will be sufficient for strategic choices but might well be valueless at the local operational level where managers, including clinicians, will wish to switch resources either between case types or between cost components, for example, nursing, x-ray, and drugs within a single DRG. At local level too, the cost-volume relationship becomes much more important. Managers not only need to know the marginal cost of an increasing volume but also the behaviour of stepped costs, such as the volume at which staffing costs will change by a sizeable amount. The behaviour of cost functions at micro and macro levels will also vary[6] according to the appropriateness of the assumptions made about the

Figure 4 Different types of 'cost'



utilisation of the various departments. Again, the application of broad-based estimates at a local level may be discredited.

Some proponents of costing DRGs[7, 8] have advocated the introduction of a costing system at patient level. In the UK, as in other socialised health services, costs are currently collected at a level higher than that of the individual patient, and to achieve patient-based costing would require a marked increase in information gathering and computerisation. While it is clear that improved DRG estimates might be obtained from patient-costing studies related to specific departments, large elements of costs – nursing, medical staff, estate management and so on – typically accounting for about 72 per cent of all costs cannot, at present, be readily attributed to individual patients. Although dependency measures, apportionment and recording systems can be devised, anything less than a very detailed system may give a seemingly legitimate, but actually spurious, level of accuracy to individual patient costs, and hence to case type, which might mislead managers rather than assist them. Currently attributing costs to individual patients is likely to prove of most benefit to managerial decision-making in areas such as drugs and dressings, surgical supplies and other consumables which can be directly costed to the patient, or in the diagnostic and therapeutic departments where the activity associated with individual patient care is now routinely recorded. It is interesting to notice that even in the US, where detailed patient billing systems already exist, current developments in case mix cost accounting initially collect many costs at a level higher than the individual patient rather than attempt to measure all cost inputs directly at patient level.[9, 10]

Table 12 Comparing retrospective and prospective systems

	Retrospective Costing	*Prospective Budgeting*
1	Snapshot, based on particular volume or workload for a single period.	Continuous; open to influence by changes in volume or methods of working.
2	Atomistic; as a result of 1, can only examine the influence of external forces in a limited way.	Holistic; as a result of 1.
3	Results need not be agreed by those costed.	Results need to be agreed by budget holders.
4	Feedback of information uni-directional; tends to be punitive in outlook.	Feedback of information is two way and can therefore be encouraging with standards set having commitment from both parties.
5	Aids analysis of previous plans but primarily at a macro-level.	Complementary to planning at both micro- and macro-level.

Finally in this section the behavioural aspect of estimating the cost of care needs to be addressed. Costs are generally estimated using historic information and are applied to past or future behaviour. Table 12 shows the differences between using costs in a retrospective manner and using them, through a budgeting system, prospectively. The major difference lies between the uni-directional nature of retrospective costing and the two-way participa-

tion of true budgeting. This is expanded upon later in chapter 10. Of course, not all prospective systems necessarily have these desirable features. Most notably the US Medicare prospective payment system in effect takes historic cost estimates and translates them into future 'prices' that the federal agency is prepared to pay for a particular type of case. Clearly this centralist and rather autocratic approach does not exhibit the qualities instanced in Table 12 and it is hardly surprising that lack of commitment to these prices has reportedly resulted in disfunctional behaviour harmful to better patient care [11, 12]. While not suggesting that the US government was unaware of these possibilities, the example shows the very different effect that single cost estimates can have on behaviour, depending upon the managerial environment in which they are introduced and the perspectives from which they are viewed by those concerned.

Approaches to costing DRGs

The previous paragraph has mentioned, and chapter 3 described at some length, the two best known methodologies in the US for arriving at DRG payment rates, namely the Medicare prospective payment system and the New Jersey DRG experiment. Table 13 gives some of the important features of their construction and shows their differences. However, this is not the end of the story. When setting up the Medicare system the US Congress established a commission, 'Propac', charged with analysing the new system and advising Congress on ways of improving it. The first reports[13, 14] focused on what the inflation update should be overall and whether there should be any adjustments in the DRG classifications or weights. Among the clinical areas so far considered by the commission are changes in practice in the treatment of cardiac pacemaker implantation, coronary angioplasty and cataract extraction, as well as new therapeutic procedures such as use of cyclosporine, magnetic resonance imaging and extracorporeal lithotripsy. These considerations have resulted in recommendations to recalibrate the 'DRG weights'. Together with political considerations, such as retaining budget neutrality, they clearly indicate the pricing (as opposed to costing) nature of the Medicare system.

Apart from the governmental production, at both national and state level, of relative weights by DRGs, many hospitals in the US are developing protocols to determine their own DRG costs, either as a first step towards internal budgeting for case mix or merely as a monitoring system to ensure continued financial viability. Figure 5 shows a typical 'cost-finding' process in US hospitals whereby specific cost centres, associated with discrete services – for example laundry, x-ray, ITU – are defined to cover the whole of the services provided by the hospital. They are used to determine 'final cost centres' that are more directly related to patient treatment and receive a proportion of overhead costs, as well as attributed or apportioned costs from the other initial cost centres. For British readers this is not dissimilar to the processes currently being adopted in management budgeting exercises where the consultant (or consultants) might be thought of as a 'final cost centre'! Once adjustments have been made to remove non-inpatient costs (not covered by DRGs), costs are assessed for each patient within a DRG using

Table 13 Comparison of Medicare and New Jersey prospective payment systems

	Medicare (effective 1 10 1983)	*New Jersey* (effective 1980–83)
Differential rates	For 9 census divisions of the US, urban and rural rates for labour and non-labour components were established. After 3 years a national urban and rural rate will apply.	
Historical cost used	Average Medicare operating cost per discharge, adjusted forward and weighted by a DRG price index ("DRG weight").	Actual historic cost of DRG adjusted forward but tempered by a weighted combination of hospital's own cost and the average of others in the area. Rates adjusted to ensure recovery of 'fixed' costs.
Exclusions in assessment of operating costs	Capital related costs, medical education, outpatient and physician payments, but including malpractice insurance costs.	Physician payments and indirect patient care costs, capital costs and other working cash allowances that are fixed for the rate year
Outlier adjustments	Average amounts decreased to offset additional payments for unusually long stay cases e.g. > 20 days more than expected stay for DRG or beyond mean + 1.94 sd, or for particularly expensive cases.	Outliers defined on LOS by narrower trim-points: low cut-off at mean −3sd, high cut-off at mean + 3sd on logged data, and reimbursed at average DRG costs (as defined by the rate setting authority).
Coverage	Medicare patients accounting for only about 30 per cent of patients nationwide allowing cost shifting between payers.	100 per cent of hospital in-patients and across all payers.
Exemptions	Many; initially these were psychiatric, long term care, children's and rehabilitation hospitals, some 'sole community' hospitals and cancer research and treatment centres.	Many since only in-patients and out-patients in short-term general hospitals included.
Controls	Hospitals required to contract with professional review organisations (PROs) to monitor quality.	Market forces assumed to act as quality control.

resource consumption statistics. Once again, though, it needs to be stated that many of the so-called 'cost' components that eventually become part of a DRG cost estimate are in fact patient charges that have either been passed straight through, for example emergency room or anaesthesia, or have been adjusted by a cost-to-charge ratio. In addition, many of the general services costs are allocated on the basis of patient days, or some function thereof; largely for want of a better methodology.

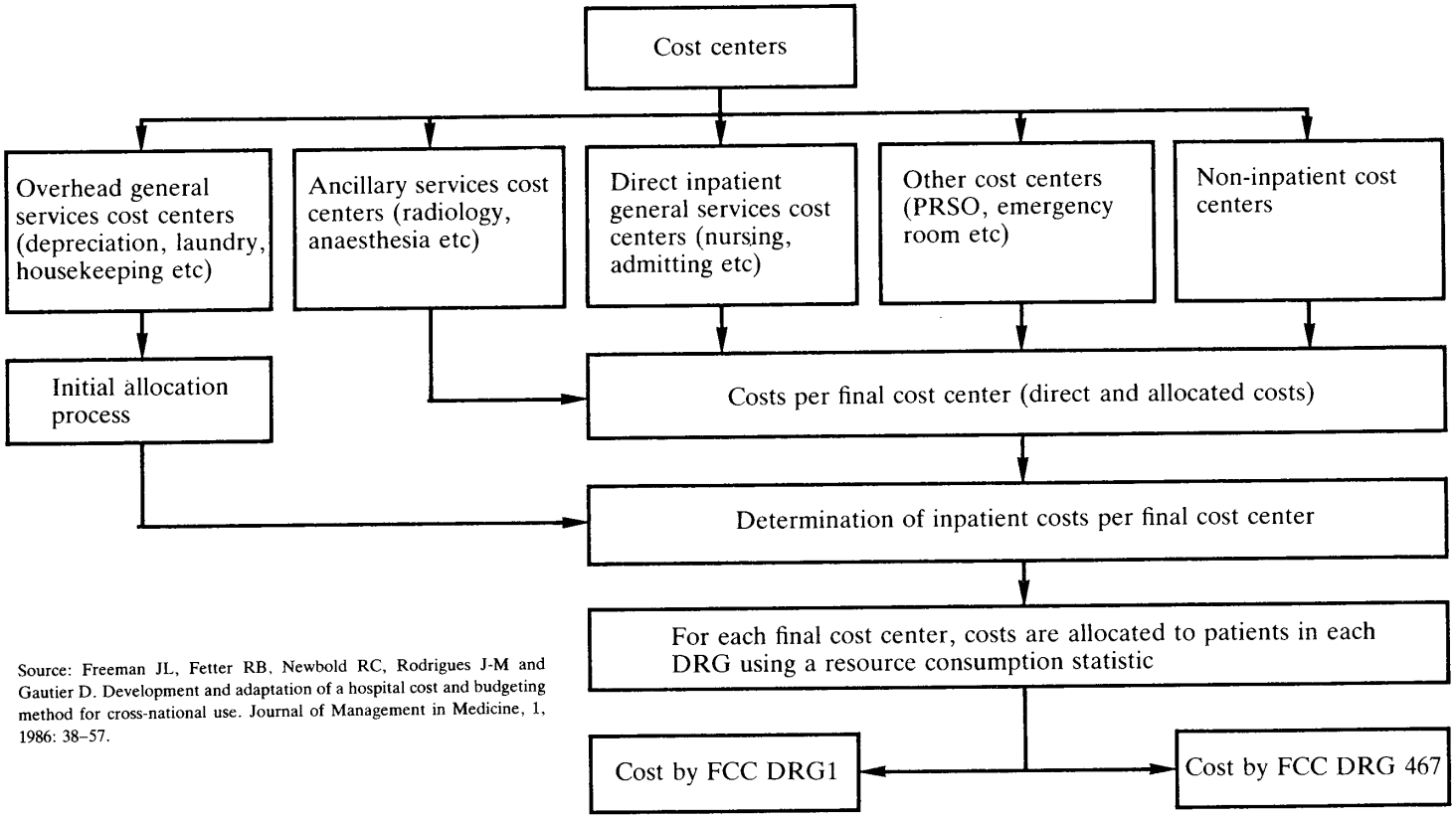
Such costing systems are now being examined by other countries and in some cases[15] are being adopted with little or no change. Where little information on patient activity is available, this may be appropriate as a starting point. But there must be reservations about its applicability – including the adoption of US relative weights and hence clinical practice – where the intention is to use the results for localised internal budgeting. The particular difficulties would seem to be threefold. Firstly, the definition of the original cost centres will differ across countries. For example, the US experience does not include the use of some medical staff resources, either directly or within the various service departments, including theatres and x-ray; while the cost of interns and residents, when included, are considered as overheads. Secondly, it would appear necessary to determine local methods for attributing resource use between DRGs. As has been said, many cost components in the US models are passed through charges. In other cases the adoption of the American statistic may be inappropriate, for example, the use of drugs. Thirdly, the fixed and variable nature of the different cost components is often only acknowledged at the budget-setting stage in the accounting system. While it is here that such issues are of direct relevance to managers, variation in costs with changes in volume should be identified at an earlier stage so that misleading costs are not estimated. At the very least, acknowledgement needs to be given to the possible sensitivity of the various cost components to such changes.

So what are the opportunities for the UK, and indeed for other countries, in using DRGs and attaching cost estimates to them? Depending upon the purpose for which the estimates are required, an approach which brings together the best elements of the following four methods should provide the most flexible and locally relevant system of costing:

1 *American information; cost component weights or relative length of stay by DRG.*
While the use of such information must be handled with care because, as mentioned above, there will be many differences in clinical practice and in the handling of certain costs (medical salaries, education and so on), it can be useful in providing comparisons with local findings.

2 *Undertake patient costing and other patient-specific studies.*
Patient costing exercises covering selected specialties[8] or particular locations within a hospital, for example, certain wards[16], have provided much useful information about activity at the patient level. While the caveats about costs mentioned earlier in the chapter remain and selective studies have their own problems, such as what proportion of shared costs are attributable to a particular specialty, a knowledge data base of the likely resource implications of particular patient types could be established to be

Figure 5 Overview of the cost finding process in US hospitals



Source: Freeman JL, Fetter RB, Newbold RC, Rodrigues J-M and Gautier D. Development and adaptation of a hospital cost and budgeting method for cross-national use. Journal of Management in Medicine, 1, 1986: 38–57.

built upon as techniques develop. Beside costing studies, patient-specific information is available from workbooks in many individual resource areas, much of it seldom used. Closer analysis could also usefully contribute to the knowledge base about current practice for each case type.

3 *Obtain consultant and other opinions about the resources required for a particular DRG.*
Intuitively appealing, a limited experiment using this approach has been reported elsewhere[17]. However to adopt this approach comprehensively would seem to require participation by respected and august bodies, such as the Royal Colleges, to determine the standard protocol centrally. Establishing these profiles locally can lead to compromises which please no one or to clinician specific profiles for each case type which serve little or no managerial purpose. Further problems arise when the 'professional opinions' are obtained too close to where the information will ultimately be used, since the assumptions underlying these opinions are likely to be coloured by proximity. For example, are the resource estimates based on the way the consultant practises within present constraints; or the way current best practice says the patient should be treated (irrespective of cost); or has the task of estimation been seen as an opportunity to bid for resources from management with all the political connotations that this carries.

It is obvious, though, that any attempt to attach resource use estimates to case types is going to require input from professional staff of a variety of disciplines; and if progress is to be made locally in advance of a national system, local input will be needed. Probably the best way to avoid the problems mentioned is to adopt a policy similar to that used in the original development of DRGs, namely to temper statistical findings with professional opinions in order to ensure the clinical probity of the results.

4 *Use all available information sources on resources used and case types treated as a basis for estimation.*
This clearly draws on the other three methods and its strength probably lies in the fact that it does not rely on a single approach. At the minimum level, the best use possible should be made of information routinely collected on a national basis. In the UK this will shortly include detailed information on the use by consultants of various diagnostic and therapeutic departments[18]. It will typically be collected in aggregate form, not related to individual patients or case types; nevertheless it will be usable. Detailed information available locally, either through patient specific studies or through other initiatives such as management budgeting, can be used to enhance the estimates. In some areas, though, it is likely that written information will not be available. In these cases, other studies or professional opinions may provide the best estimates to be found initially. Of course, once these estimates are in use, it will quickly become apparent whether or not they are reasonably accurate. This sort of approach might be required for estimating medical staff time across case types.

An example of this rather pragmatic use of available information has been used on the British specialty costing[19] data which was a precursor of the current management budgeting initiatives. Between 1980 and 1982,

seven centres in England and Wales used broadly similar costing methodologies to translate cost records which had previously been kept on a functional basis (nursing, pharmacy and so on) to a clinical specialty base (such as general medicine). Using reported data from these centres the CASPE DRG team:

assessed which cost components were best considered as one-off per case costs, and which were closer related to length of stay;

calculated specialty specific case and per diem costs for each centre;

standardised these across centres (to remove the effect of an 'expensive' centre) and produced specialty weights;

tested the specialty weights for stability;

used specialty weights to predict hospital costs for 22 hospitals in one UK region;

applied specialty weights to each DRG for each centre separately to arrive at a DRG weight; and standardised DRG weights using a 'basket' of 300 DRGs common to all seven data sets.

It was found that the production of fairly stable specialty weights was possible, although some specialties showed more variability than others e.g. special care baby unit (where different centres used different definitions) and plastic surgery (where relatively few cases were seen). The weights were found to be better predictors of cost variation across a region than either the number of cases treated or the number of patient days. When translated to DRG weights the results were again broadly encouraging in that the unexplained difference in cost between centres was reduced by 53 per cent.

Although it uses only fairly limited and crude data, this approach has provided some interesting insights into cost variation by case type. It would be possible to refine the methodology further, but even at this undeveloped stage the approach could be useful to managers at a strategic planning level.

Resource use before costs

Towards the beginning of this chapter I argued the case for separating the issues surrounding the costing of DRGs from those surrounding the assessment of resource use by particular case types. Not to do so seems, firstly, to risk introducing all the ambiguities surrounding the appropriateness of the costing methodology at the expense of identifying the true efficiency of resource use; while, secondly, the reduction of information to a single cost per DRG (or even to a number of cost components) limits the choices for action that are open to managers, including clinicians.

The approach I have been advocating draws on the four methods itemised in the previous section, and in particular method 4, in order to build up a matrix of the relative intensity of resource use by particular case types (referred to as 'the Relative Intensity Matrix' or RIM). The relative intensity,

which can be divided into two further components, the relative frequency of use and the service intensity of use, can be defined as follows.

Relative intensity	=	*Relative frequency*	×	*Service intensity*
(Relative use of resource type j by a typical patient in DRG type i)		(Proportion of patients in DRG type i expected to use resource type j)		(Intensity of resource use of type j by a typical 'resource using' patient in DRG i)

Among the points to note in this equation are first that the resource type j can be defined at a variety of levels of sophistication: for example, at department level; as labour and consumable elements within a department; or by type of labour and so on (enabling marginal costs to be eventually estimated). Secondly, intensity of resource use incorporates not only concepts of volume but also of gradation within the resource type. For example, different grades of staff undertaking procedures of the same duration in the pathology laboratory will clearly have differing resource implications. Thirdly, the relativities between two DRGs will remain unchanged in any particular resource type, x-ray for example, if the proportion of patients using x-ray varies in a reciprocal manner to variations in the intensity of use by 'resource using' patients.

Using the pragmatic approach of method 4, at any point in time the elements of a RIM might be assessed as follows:

Medical staff (direct only)	No knowledge: assumed equal across all cases, hence equal weight = 1.
Nursing staff (on ward)	Assumed directly related to length of stay.
Diagnostic services	Determined through management budgeting activity data, with some cases being investigated as outpatient and hence attracting a zero weight.
Theatres	Zero weight for medical DRGs; minutes of theatre time weighted by specialty for surgical DRGs.
Drugs	Relative weightings across DRGs assessed by the District's pharmacy and therapeutic sub committee and agreed with the District Pharmaceutical Officer.
Overheads	Assumed directly related to length of stay.

Such assessments might be incorporated into the matrix as shown in Table 14 where, it should be stressed, the figures are purely illustrative.

Although some elements are based on absolute figures like length of stay, to avoid confusion when applying a costing methodology they should all be considered as relative weights within resource type: for example, DRG 39 uses 8.8/8.5 = 1.035 as much nursing resource as DRG 38. In this way it would be possible to make various comparisons of the relativities across

districts; for example, DRG 39 is more resource intensive in relation to other case types in District A than in District B. To make comparisons in absolute terms – DRG 39 uses more resources in District A than in District B – one would also need to consider the amount of resource to which a unit weight referred. For example, in Table 14, the unit of nursing resource might be described as the amount of basic nursing care given within a single day of stay; theatre usage might relate to the average resources employed in one minute of theatre time on a general surgical case: while in the diagnostic services, the unit weight within the RIM might refer to resources employed on a basic profile of laboratory tests, x-rays and so on.

Knowing the basis on which the weights are established and possessing a comprehensive matrix of resource weights (however imperfect in its construction), it is then open to local choice to determine the costing methodology to be used. Two possibilities, out of many, would be a zero-based approach or an attribution of historic aggregate costs to case types. In the zero-based approach the cost of a basic unit, for example, a theatre minute, might be established and from this the cost of a planned case mix could be calculated.

A significant difference between this approach and the adoption of a US costing model would be that the nature of the fixed and variable costs – as well as knowledge of other local cost functions – could be built into the weighting system rather than producing costs on the assumption of total variability and then having to disentangle the nature of the various cost elements.

Attribution of historic costs starts from knowledge of the aggregate cost, either as an absolute or in terms of a per case or per diem average. These costs can then be passed through to case types based on the relative weights and the number of cases treated in each DRG. In Table 14, average per case figures for diagnostic services and overheads of say £80 and £200 per case would be attributed to the DRGs as follows:

	Diagnostic Services	Overheads
	£	£
DRG 38	63.88	232.03
39	88.43	240.22
40	0	57.32
.	.	.
.	.	.
.	.	.
45	132.65	210.19
46	103.17	169.24
47	93.35	92.81
Average cost per case	£80	£200

The emphasis of this section, though, has been primarily at local level where it is hoped that knowledge of the resource implications of particular case types will assist managers in their decision making. As case mix measures become more widely accepted throughout the UK, in Europe and other parts of the world, clinicians will inevitably become more involved in planning the use of resources together with the output of their service. The availability of the RIM within a clinical service planning framework offers doctors the opportunity to review their practice and to model the effect of changes on

Table 14 Example of resource use weightings by case type

DRG	38	39	40	 45	46	47
Number of cases	48	417	55	3	10	95
Length of stay	8.5	8.8	2.1	7.7	6.2	3.4
Relative Weights						
Medical	1	1	1	1	1	1
Nursing	8.5	8.8	2.1	7.7	6.2	3.4
Diagnostic services	1.3	1.8	0	2.7	2.1	1.9
Theatres	3.3	29	19	0	0	0
Drugs	1.5	1.2	1.1	1.6	1.4	1.2
Overhead	8.5	8.8	2.1	7.7	6.2	3.4

NB Figures are purely illustrative

their use of resources in at least five different ways. They can examine the effect of:

1 changing patient numbers across DRGs or altering the length of stay within a DRG;
2 changing the frequency of use of service departments by patients in a particular DRG;
3 changing the intensity of use of service departments by patients in a particular DRG;
4 reducing the overall unit cost of individual cost components;
5 changing a combination of 1–4.

Although dealing in the main with average resource use and average cost within case type, the approach could be adjusted to cope with variability within DRG where it was felt that this was significant. Such adjustments might include assessing the differential effect across case types of those patients who remain 'blocking beds' once their medical treatment has ended or the effect of the additional cost of secondary referrals to specialist centres. The sensitivity of overall costs to changes in these and other factors would provide much useful information in responding flexibly to demands placed on health services and would lead to more appropriate planning, as well as more appropriate assessment of performance.

Summary

This chapter has reviewed a variety of approaches to attaching costs to DRG case types. It has attempted to show the catch-all nature of the term 'cost', caused partly by the confusion surrounding the many different types of cost which through common usage, implies a simple, unambiguous concept. In the US literature, for example, 'cost' has often been synonymous with 'charge' or some function of it.

Moves to attach costs to DRGs without sufficiently examining the underlying activity ignore much useful information. However, a necessary condition for the success of any information system is that the recipient is motivated to act (or to decide not to act). DRGs and their associated resource use offer a greatly enhanced opportunity to managers and clinicians to plan

the use of clinical resources more efficiently. But they do require both an appropriate framework within which to operate and the managerial will to allow them to fulfil their potential.

References

1 Schweiker R S. Report to Congress. Hospital prospective payment for Medicare. Washington, Department of Health and Human Services, 1982.
2 Thompson J D, Averill R F and Fetter R B. Planning, budgeting and controlling – one look at the future: case-mix cost accounting. Health Services Research, 14, 1979: 111–125.
3 Finkler S A. The distinction between costs and charges. Annals of Internal Medicine, 96, 1982: 102–109.
4 Lave J R. Is compression occurring in DRG prices? Inquiry, 22, 1985: 142–147.
5 Transition Systems. Transition 1: an integrated planning and management control system of hospitals. Boston, 1985.
6 Evans R G. Behavioural cost functions for hospitals. Canadian Journal of Economics, 4, 1971: 198–215.
7 Poynton D. Difficult questions: no easy answers. The Health Service Journal, 19, June 1986: 822–823.
8 Newman A. DRG based financial information systems – a case study in ophthalmology. In: Diagnosis related groups and resource management. CASPE Research, 1985: 33–42.
9 Newbold R C. UHS cost/budget model: overview. New Haven, Health Systems Management Group, School of Organisation and Management, Yale University, 1984.
10 Williams S V, Finkler S A, Murphy C M and Eisenberg J M. Improved cost allocation in case-mix accounting. Medical Care, 20, 1982: 450–459.
11 Stern R S and Epstein A M. Institutional responses to prospective payment based on diagnosis-related groups. New England Journal of Medicine, 312, 1985: 621–627.
12 Grimaldi P L and Micheletti J A. Utilization and quality review under the prospective rate system. Quality Review Bulletin, February 1984, 30–37.
13 Prospective Payment Assessment Commission. Report and recommendations to the Secretary, US Department of Health and Human Services. April 1985.
14 Prospective Payment Assessment Commission. Report and recommendations to the Secretary, US Department of Health and Human Services. April 1986.
15 Freeman J L, Fetter R B, Newbold R C, Rodrigues J-M and Gautier D. Development and adaptation of a hospital cost and budgeting method for cross-national use. Journal of Management in Medicine, 1, 1986: 38–57.
16 Financial Information Project. Patient based costing in the hospital – field trial report. Birmingham, West Midlands Regional Health Authority, May 1984.
17 Scott T and Sherwood D. DRG treatment profiles. Public Finance and Accountancy. December 1984: 34–36.
18 Department of Health and Social Security. Steering group on health services information (Chairman Mrs E Körner) First report. London, HMSO, 1982.
19 Hillman R L and Nix G R. DHSS funded research into specialty costing 1980–1982. London, DHSS, 1983.

DRGs in management

7 CASE MIX AND MANAGEMENT ISSUES

Martin Bardsley and James Coles

Previous chapters have discussed the ideas behind case mix measures and have described some situations in which they have been used. These might be thought somewhat removed from the current experience of health care managers. This chapter changes the emphasis by first examining the tasks facing health service managers before considering how case mix information can help them. Using a number of examples it identifies how a knowledge of case mix can bring a fresh perspective to issues, both old and new. While political and other constraints will undoubtedly influence the managerial action locally, the benefits that accrue from a measure of case mix are widely applicable across the service.

One of the main tasks of health care managers is to determine the strategic direction to be taken in terms of the nature and volume of particular care programmes that are to be offered. Here, case mix classifications at regional or district level will act chiefly as descriptors of the activities of the organisation and facilitate planning and management that are more responsive to the needs of the locality. Management is also concerned with making plans operational, monitoring performance against desired objectives and taking action where differences from the plan arise. As is discussed later, a variety of factors can cause such differences, only some of which will be in the manager's control.

What is good health service management?

Specifying what constitutes good health service management is not an easy task, especially in a limited space; however it is possible to determine some characteristic goals that a good manager will seek to fulfil. At the most basic level, a health district or its hospital will seek to provide health care services for its catchment population, defined on geographic, historical or financial grounds. These services should be as effective and efficient as possible, and to the highest level of quality that can be achieved realistically. It is the intangibility of some of these ideals and the absence of unambiguous indicators of good management that make the process of decision making in health care so difficult and, to many people, so challenging.

The rational approach to planning and management[1] outlined in the previous section is appealing, but even in the private sector, where it is argued that less ambiguous indicators of success exist, it has been criticised as being over-simplistic. Its applicability to health services is also open to question simply because of the difficulties of identifying with any certainty the complex and often conflicting objectives of the enterprise[2]. Alternative models emphasise a somewhat less structured role for management within the political resource market which controls social policy[3].

The successful health care manager is someone in the unfortunate position

of having to marshall limited resources to achieve the greatest social benefit; that is, to provide a 'good' service. To do this, the manager will need to arbitrate between the competing claims of professional staff in an institution while coping with the demands passed down through the health service bureaucracy. At the local level, effective management will entail trade-offs within and across professional and patient groups, a process which becomes more pressing with finite, and often diminishing, budgets. For example, the British NHS accepts that certain types of care cannot (or will not) be provided because resources have to be moved to other areas of provision. In the past, the recognition of these often unpleasant choices has tended to be implicit in the decision making process. Now there is pressure to bring these choices into the open in order to make judgments more explicit and more open to questioning. An example is clinical budgeting (discussed in more detail in chapter 10). Many health services have decided that clinicians should be more accountable for the cost consequences of their clinical decisions, but in order to make appropriate choices on resource deployment a greater knowledge is needed of the effect of case mix.

The process of distinguishing patient types is often used, though the specificity of descriptions can vary enormously – from the level of individual procedures (for example, the number of coronary vein grafts to be performed) to that of a care programme, such as the provision for geriatric patients. DRGs differ from many of these descriptions in that the classification can provide routine cost details across all inpatient case types. This can give a description of hospital activity which is closer to an approximation of a health service output than any other currently available on a routine basis. Such an ability to describe more accurately the services a hospital provides – and in terms that are more closely related to patient care – can form the basis for more equitable decision making. This is a potentially important shift for health service management.

Determination of appropriate service provision

In the past, evaluations of the level of services provided have had to be expressed either as structural aspects of care made available, like the number of acute beds, the location and capacity of community health centres and the range of diagnostic facilities, or more directly as crude patient statistics, such as cases treated, visits to GPs, x-rays performed and so on. DRGs offer a more precise and relevant input to these evaluations, albeit only for the hospital inpatient sector. A description of gynaecological inpatient care, for example, can be broken down into the specific numbers of cases in individual DRGs, together with estimates of their associated resource use. An example from one district for one year is shown in Table 15.

Having this information on a routine basis gives a much better understanding of overall provision and of the trade-offs that may occur between the groups of patients hidden within aggregate headings. Provision for a district might be considered in terms of the numbers of cases treated in a given DRG, per head of population. This would be especially relevant when considering, for example, the rates of certain types of elective surgery and the waiting lists they cause. If a health authority wishes to increase the number of hip

Table 15 Describing the workload in gynaecology

DRG	Number of cases	Average daily bed use
353 S Pelvic evisceration, radical hysterectomy and vulvectomy	49	1.5
354 S Non-radical hysterectomy age>=70 +/− cc	26	0.8
355 S Non-radical hysterectomy age<70 w/o cc	256	6.8
356 S Female reproductive system reconstructive procedure	83	2.0
359 S Uterus and adenexa procedure for non-malignancy excluding tubal interruption	123	2.2
360 S Vagina, cervix and vulva procedure	91	1.0
361 S Laparoscopy and endoscopy excluding tubal interruption	40	0.3
363 S D&C, conization and radio implant for malignancy	37	0.4
364 S D&C, conization except for malignancy	346	2.5
365 S Other female reproductive system OR procedure	73	1.8
368 M Infections, female reproductive system	33	0.4
369 M Menstrual and other female reproductive system disorders	117	0.7
379 M Threatened abortion	78	0.6
381 M Abortion with D&C	553	3.1

cc = complications or co-morbidity
S = Surgical DRG; M = Medical DRG
Seventy seven per cent of cases and 81 per cent of inpatient days are covered by these DRGs

replacements performed, case mix information could identify the resource implications and also the effect on services likely to be denied resources. If more hip replacements are to be performed, should other forms of elective surgery, in terms of specific DRGs, become lower priorities or should different care groups share the burden? Which are the case types that have a high proportion of emergency admissions where there may be little opportunity for change?

Another advantage of DRGs is that they provide a comprehensive description of inpatient types which permits the relative frequency of admissions to be compared between districts. Why does a DRG show relatively high admission rates? It may be due to lack of facilities in the primary or ambulatory care sector; it may reflect patients coming from other districts; or there may be over-investigation of specific complaints at the expense of other types of care.

Similar information can be used to determine the provision required when new facilities are being planned, or to review the needs of a particular service. Hospitalisation rates and the length of waiting lists vary dramatically across the country. Even in non-elective conditions, a large discretionary element in

hospital admission practice has been demonstrated[4]. DRGs, with their greater specificity, enable the relative provision of facilities to be assessed in the light of the needs of the local population and for planning future requirements. Such evaluations need to accommodate likely changes in the pattern of care provision. While DRGs define inpatient care, to identify the benefits and trade-offs between inpatient and outpatient, acute and long term care settings, additional systems of classification need to be developed.

Planning for changing populations and medical technologies

Case mix measures offer further opportunities for managers through their ability to monitor the resource implications of patients being treated in a hospital who come from outside its immediate catchment area. England, like many countries, provides specialist services, such as renal care and neurosurgery, from centres serving large geographical regions. Historically, funding mechanisms are deemed to have provided these centres with the financial resources sufficient to care for all patients in the region that present for treatment. However, the receiving referral centres have felt unable to control the volume of cases or to satisfy themselves that income from them equals expenditure.

It is not only in these specialist areas that health authorities have been concerned about the financial burden of patients 'imported' across administrative boundaries. Under the methodology used to allocate resources to health authorities in England, each patient brings with him a 'per capita' assessment of the resources required for his care, but since this is an average calculated at the specialty level there will be a shortfall if the more resource-intensive cases within a specialty predominate. In both examples, a DRG-specific assessment of costs would make the systems more equitable, without necessarily increasing overall expenditure. In the terminology of the US, it could be designed to retain 'budget neutrality'. The 'pricing' of these individual cases would need to reflect the overall constraints affecting local patient care but would have the dual advantages of removing the incentive to 'export' particular case types and of making explicit the volume of cases that can be afforded within any resourcing level.

As implied above, the value to managers of DRGs lies in the predictability of resource use associated with a particular case type. The ability to predict the likely effect of case mix changes, or changes in the pattern of care for a given case type, affords managers the opportunity to plan for demographic or technological change. For example, the implications of an ageing population can be examined in terms of its effect on case mix and hence on all aspects of resource use in the hospital. Table 16 shows the very different age profiles of patients treated across a range of DRGs in a single specialty. By identifying the case types or DRGs associated with certain age groups and then modelling the changing age structure of the population, the relative importance of these groups within the total case mix can be assessed. Though adjustments may be needed to cope with changes in technology and styles of provision, it should be possible to predict the likely future demand on resources in a more specific and comprehensive way than at present.

Developments in medical technology can have important implications for

Table 16 Number of cases by age for DRGs in ophthalmology

DRG	<20	20–29	40–59	60–79	>80
36	5	7	7	13	1
37	0	0	1	2	1
38	6	1	5	20	7
39	6	5	14	73	21
40	2	8	24	32	5
41	257	0	0	0	0
42	4	5	5	5	0
43	1	0	2	1	2
44	1	0	2	1	2
45	4	2	2	5	1
46	0	0	1	7	5
47	0	9	2	22	10
48	42	0	0	0	0

hospital financing and managing. Questions such as should the hospital invest in new imaging equipment, can be considered in terms of the changes in inpatient case types treated, the quality of care given to the individual patient and any resultant change in resource use. Which DRGs will be admitted less frequently; which DRGs will require additional resources for diagnosis or a longer length of stay; and, perhaps most importantly, are the benefits to patient care sufficient to justify any additional resources needed and, if so, in what areas of case mix will savings have to be made to pay for them? Once a decision to invest has been made, and the consequences of introducing the new technology agreed with the clinicians concerned, DRGs can be used to examine whether the desired results have been achieved.

Implementing policy

The vocabulary that DRGs provide is important in the process of turning the policy requirements of either the health authority, or higher levels of bureaucracy, into practical change. Such policy requirements are often couched in the most general terms, either through a deference to local autonomy and responsibility or through an unwillingness to specify practical changes that must be made at the grass roots level. Managers need a useful way of interpreting policy and expressing it in readily understandable terms that can provide the basis for subsequent action. For example, there is currently pressure on acute hospitals in the UK to increase the amount of day surgery performed[5]. Routine information on case mix can identify those patient types where day surgery is appropriate, indicate the level of extra day care provision required, and any resources that could be expected to be released for other inpatient services. Changes in the style of provision are probably best agreed through a negotiating process between managers and clinicians. DRGs can assist these discussions both within a particular clinical area and between specialties – the wider perspective required by hospital managers. In the day surgery example, the additional resources of theatre time, specialised day wards, changes in staffing and so on must be addressed at the hospital level since they have implications for more than one specialty

and service department.

To expect a single system to be appropriate for all clinical and managerial aspects of inpatient care is, of course, too optimistic. DRGs can be seen as crude descriptors in clinical terms, while for some managerial purposes the 467 groups will be far too many. In some applications it may be that the DRG classification will require refinement. For example, established measures of nurse dependency by individual patient may be more useful when planning day-to-day nursing requirements than the knowledge that over a year a particular case mix is likely to require *x* amount of nursing hours. DRGs are a compromise between the specificity at the individual patient level demanded by clinical practice and the practical requirements of general management. They do, though, offer an increased knowledge of the relationship between the input of resources and output in clinical terms as well as identifying opportunities for change where this relationship is judged to be unsatisfactory.

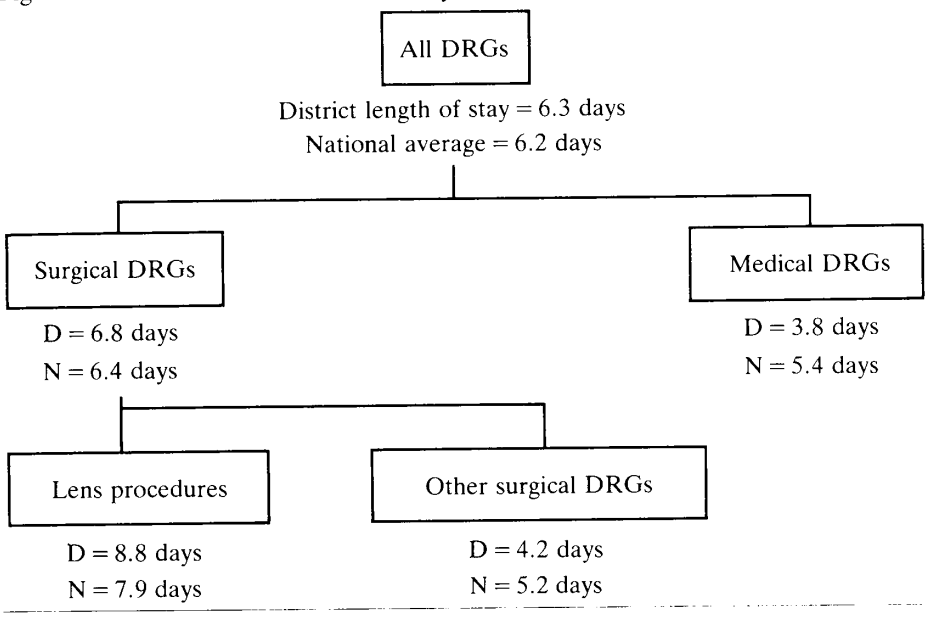
Monitoring performance

A prime responsibility of managers is to monitor progress against agreed plans and to decide whether action is necessary should variation from the plan be identified. They might either relate to differences in the resources consumed in providing a given level of care, or to the outputs obtained, usually expressed as the volume of care provided, although of course quality and other factors should be included. So what part can DRGs play in assisting managers with this aspect of their responsibilities?

Earlier sections touched on some of the issues surrounding variations in the volume of different types of cases treated and the difficulties this causes when resources are not available to treat additional cases, such as the funding of regional specialties and cross-boundary flows. It is not only in these select examples that changes in the balance of case types treated will cause concern to managers, yet until DRGs were developed there were few practical ways of approaching such problems. A further example, shown in Figure 6, may be helpful.

In this instance, a health authority is reviewing its performance in the specialty of ophthalmology. Examining the length of stay of ophthalmic patients, it finds a figure similar to the national average (district (D) = 6.3 days; national average (N) = 6.2 days). Without case mix information, no action is thought to be warranted since overall volume is expected to remain static. If that district had case mix information it would have shown that while medical cases (DRGs 43–48) are efficient in their use of beds (D = 3.8 days; N = 5. 4 days), the surgical cases (DRGs 36–42) accounting for 85 per cent of the cases, were less so (D = 6.8 days; N = 6.4 days). Furthermore, on looking at individual DRGs one finds that a single case type accounting for 48 per cent of the district's ophthalmic cases has a length of stay 10 per cent greater than the national figure. Managerial judgment must decide what action to take about such a difference since there could be many possible causes, such as difficulties in theatre scheduling, obtaining prostheses and so on. What is certain, though, is that if this particular case type increases its share of the specialty workload without a change in the overall volume of cases treated, more bed days and associated resources will be required; a fact hidden when

Figure 6 Diseases and disorders of the eye



examining only the aggregate figure.

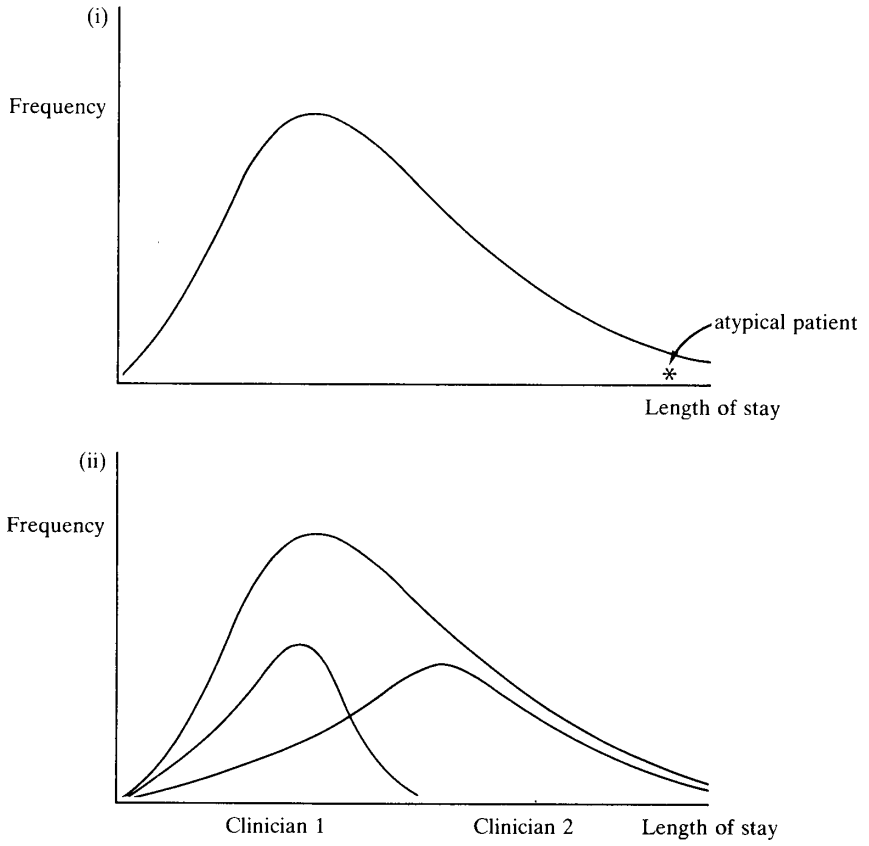
It is not only the balance of care between patient types that needs monitoring in order to achieve the plan. While a single DRG groups patients who are determined to be medically similar and to use approximately similar resources, an individual patient within a DRG will have a unique resource consumption. Any population of patients in one DRG will exhibit a distribution of resource use, so that there is inherent variability within a DRG.

Figure 7 shows the distribution of length of stay of two separate DRGs. Figure 7 (i) is the more typical distribution – a fairly homogeneous group but with a number of patients, at the right hand side of the distribution, with a long length of stay. Figure 7 (ii) shows a distribution with a greater variation. This is a less satisfactory DRG for planning purposes than that in 2 (i) and might contain two or more distinct patient types. For example, the group might be found to contain two diagnoses which are treated very differently from each other, or two different patterns of clinical practice related to a single diagnosis.

It is important to know whether variation within a DRG is due to a number of isolated cases or to systematic variation from the plan or target. In the case of the former, the action needed is to identify these individual cases and examine them separately. It may well be that they are inappropriately grouped in the DRG in question, not necessarily through coding errors but merely through the way their cases have developed. For example, a patient admitted for a cataract operation may have a heart attack during his stay. Since the reason for admission was ophthalmic he will be assigned to an ophthalmic DRG but clearly his use of resources would, we hope, be atypical from the normal cataract care pattern. Such auditing of individual cases can often provide valuable insights into the quality of care and in to latent problems in admission and discharge procedures.

In order to screen out such atypical cases from individual DRGs, the Medicare scheme in the US adopted a technique known as 'trimming' the distribution (to avoid distorting the federal government's system of 'prospective payment'). While the detailed calculations need not bother us (see Appendix III), the method determines points along the length of stay axis for each DRG beyond which a patient would be considered atypical for that group and the hospital would be reimbursed differently for that case. A variety of trimming methods have been used ranging from the complex US formula[6] to the more usual 'mean ± 2 standard deviations' carried out on transformed data. Statistically derived trim points can also be modified in the light of medical judgments about the resource use of patients 'typical' of the group. In the UK an additional form of trimming is sometimes necessary to separate 'social' care from the acute treatment days, in other words to identify those days at the end of a patient stay which are needed while arrangements are made with other support agencies. This mostly occurs in the care of the elderly and it is often worth looking at geriatric cases within a DRG separately from other specialties. To examine them in aggregate can cloud messages that could be obtained from either group individually. Before leaving this discussion, it should be noted that the choice of trimming methods is a political one, not only in the US where such decisions affect the

Figure 7 Length of stay distributions



methods is a political one, not only in the US where such decisions affect the resourcing levels of hospitals, but in other health care systems where assessment of performance is likely to be based on the population remaining within the trim points.

Identifying causes of systematic variation

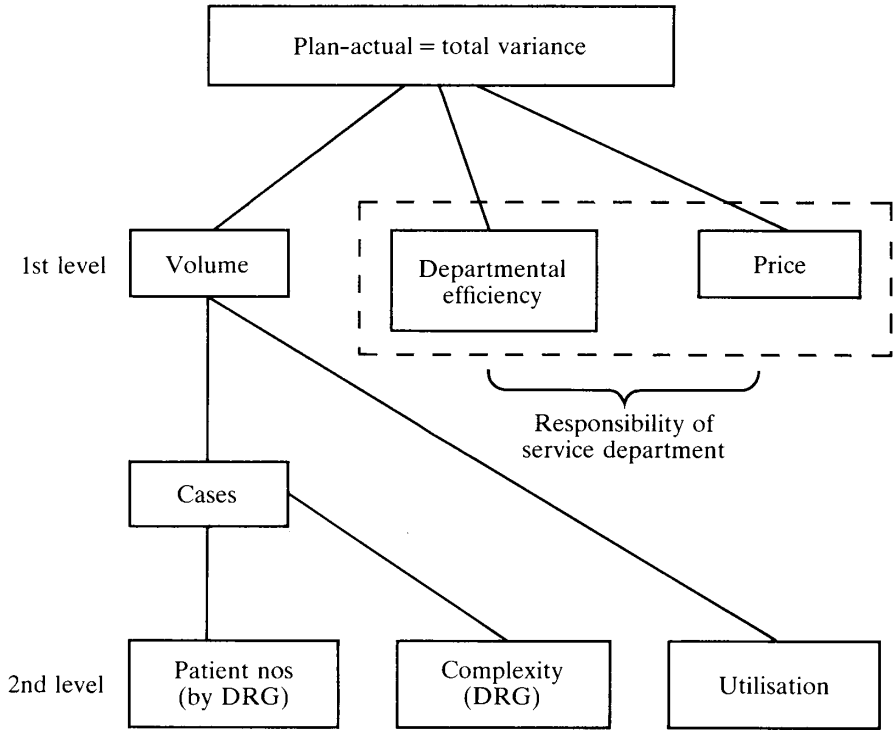
Having separated out the atypical cases, attention should turn to any residual variation exhibited within a DRG. For example if it was expected that on average patients in DRG 36 'Retinal procedures' would have a length of stay of 8.2 days or require 62 minutes of operating theatre time and the actual outturn was significantly greater, then managers would wish to determine the cause of this variation. One proposal that is often put forward for explaining such differences is that cases in a DRG present with differing levels of severity of illness. Consequently the comparison to the reference value may be unfair if it does not take this into account[7]. This was more fully discussed in Chapter 1. For such arguments to be valid, it needs to be demonstrated that more severe cases are distributed on a non-random basis within the case type under consideration. Furthermore, in order that greater severity should command additional resources from the total available it needs to be shown that such measures are based on patient attributes on admission, rather than attributes identified or occasioned during their stay. Work on this difficult topic is continuing.

Returning to our consideration of residual variation within a DRG, treasurers have for many years been familiar with the techniques of variance analysis. Although often used in financial circles this method of disaggregation, which examines the contribution made by a number of independent factors to a single indicator, is seldom applied to managerial statistics[8]. Figure 6 is an example of disaggregation.

Figure 8 shows diagrammatically the value to managers of examining variances in a structured manner, taking as an example the use of x-ray film on the total caseload of a single consultant. The overall difference between planned and actual usage is a fairly bald figure of little help to managers. First level differences can identify whether the volume of requests made to the x-ray department by the consultant was different from that planned, or whether the efficiency of the department – the number of films used to achieve a reported x-ray – had varied. If the cost of the x-ray film was under consideration then the price of film would be another important variance at this level, with the responsibility for purchasing perhaps also lying with the head of the service department.

If the variance is shown to be due to the volume of requests made, a further level of disaggregation is of interest. Is this variance due to the number of cases treated or the number of requests made per patient? Under a case mix information system, the volume of cases treated can be separately identified, each case type having an expected level of demand on the radiology department. Variation from the overall planned volume of requests would then be identified as having been caused by one or more of the following:

Figure 8 Variance analysis



a change in patient numbers (by individual DRG);
a shift between DRGs (a change in complexity);
a change in requests made per patient (by DRG).

Variation in any of these factors will provoke a different response from management. Increased knowledge about the effect of each of them on resource use should offer improved opportunities to meet the needs of the service.

So far these examples have largely focussed on comparison of workload or performance against a national figure, a norm, or a plan agreed with the individual clinician. Comparisons between clinicians at the local level have been avoided, although this may be an area which managers will wish to address. Two further examples will serve to illustrate issues that have significant resource implications for a health authority but which without a case mix classification are not readily explored or easily debated.

First, comparison of DRGs by discharging specialty consistently shows (across a wide range of groupings) that cases discharged from a specialty seemingly inappropriate to the diagnostic group under which they are classified have a longer length of stay than similar cases discharged from the more appropriate specialty. Table 17 gives examples of this for selected DRGs in the major diagnostic category relating to the ear, nose and throat.

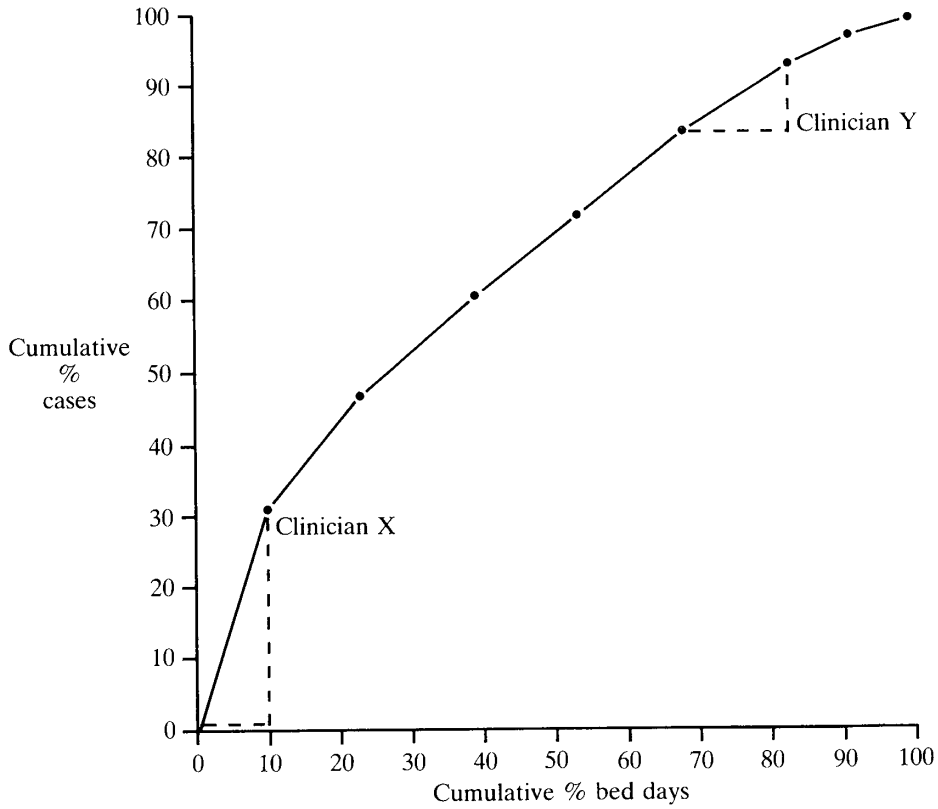
Secondly, it appears that clinicians who have a particular interest or specialism covered by a subset of DRGs are more parsimonious in their use of

Table 17 Length of stay by discharging specialty

DRG	ENT	Paediatric	*Discharging specialty* General surgery	General medicine	Infectious diseases
50 Sialoadenectomy	5.6		6.7		
63 Other ENT operating room procedures	3.4		4.3		
65 Dysequilibrium	3.6	2		7.2	
68, 69, 70 Otitis Media + URI. (defined by age etc)	2.7	3.3		4.7	3.4

resources than clinicians who only treat a small number of patients within the same subset. This is shown diagrammatically in Figure 9. Clinician X admits nearly 30 per cent of the total workload for this DRG yet uses less than 15 per cent of the bed days. Clinician Y uses approximately the same number of bed

Figure 9 Clinician variation within DRGs



days but admits many fewer cases. Although one might include caveats concerning quality of care, differing severity of cases, differences between on-take and elective work and so on, this finding is so consistent as to belie these suggestions. It is also largely independent of clinician, since clinician Y is not necessarily 'inefficient' across all case types.

These two examples would suggest that the referral process, either to individual clinicians or to particular specialties, may not make the best use of available resources. An improved system with patients being routed appropriately, where possible, would result in a better use of resources and patients being treated faster. It would also permit doctors to concentrate on areas of special interest or expertise with a probable, although as yet unquantified, improvement in patient care. While this radical approach might be unacceptable to those who wish to receive a broad mix of patients, the analysis by case types does at least permit these possibilities to be considered.

Conclusion

The issues raised in this chapter represent only some of the problems facing health service managers. While many have acknowledged that case mix measures would assist in tackling these problems, measures have rarely been used in other than their crudest forms, such as at specialty level.

DRGs, like other management information systems, will not remove the need for sound managerial judgment. They will assist in setting appropriate objectives and formulating plans but they will not tell whether it is the best time to pursue a particular issue. Nor will they provide a ready-made solution to difficult problems. DRGs provide a tool in the form of a simple classification of patient types that can address a large range of questions currently facing health service managers and they bring together demographic, clinical and resource databases in an easily accessible language. It remains to be seen whether the managerial will exists to use such a tool and how effective it is in helping to provide appropriate and high quality care.

References

1 Koontz H, O'Donnell C. Essentials of management. New York, McGraw Hill, 1978.
2 Titmuss R M. Social policy – an introduction. London, Allen & Unwin, 1974.
3 Glennerster H. Social service budgets and social policy: British and American experience. London, Allen & Unwin, 1975.
4 Wennberg J E, McPherson K and Caper P. Will payment based on diagnosis-related groups control hospital costs? New England Journal of Medicine, 311, 1984: 295–300.
5 Royal College of Surgeons of England. Guidelines for day case surgery. Commission on the Provision of Surgical Services. London, Royal College of Surgeons, 1985.
6 Health Systems Management Group. The new ICD-9-CM diagnosis related groups classification scheme. Section IV. Final Report. 1982.
7 Horn S D. Measuring severity of illness: comparisons across institutions. American Journal of Public Health, 73, 1983: 25–31.
8 Finkler S A. Flexible budget variance analysis extended to patient acuity and DRGs. Health Care Management Review, 10, 1985: 21–34.

8 USING DRGs TO PLAN IMPROVED HEALTH DISTRICT PERFORMANCE

Gwyn Bevan

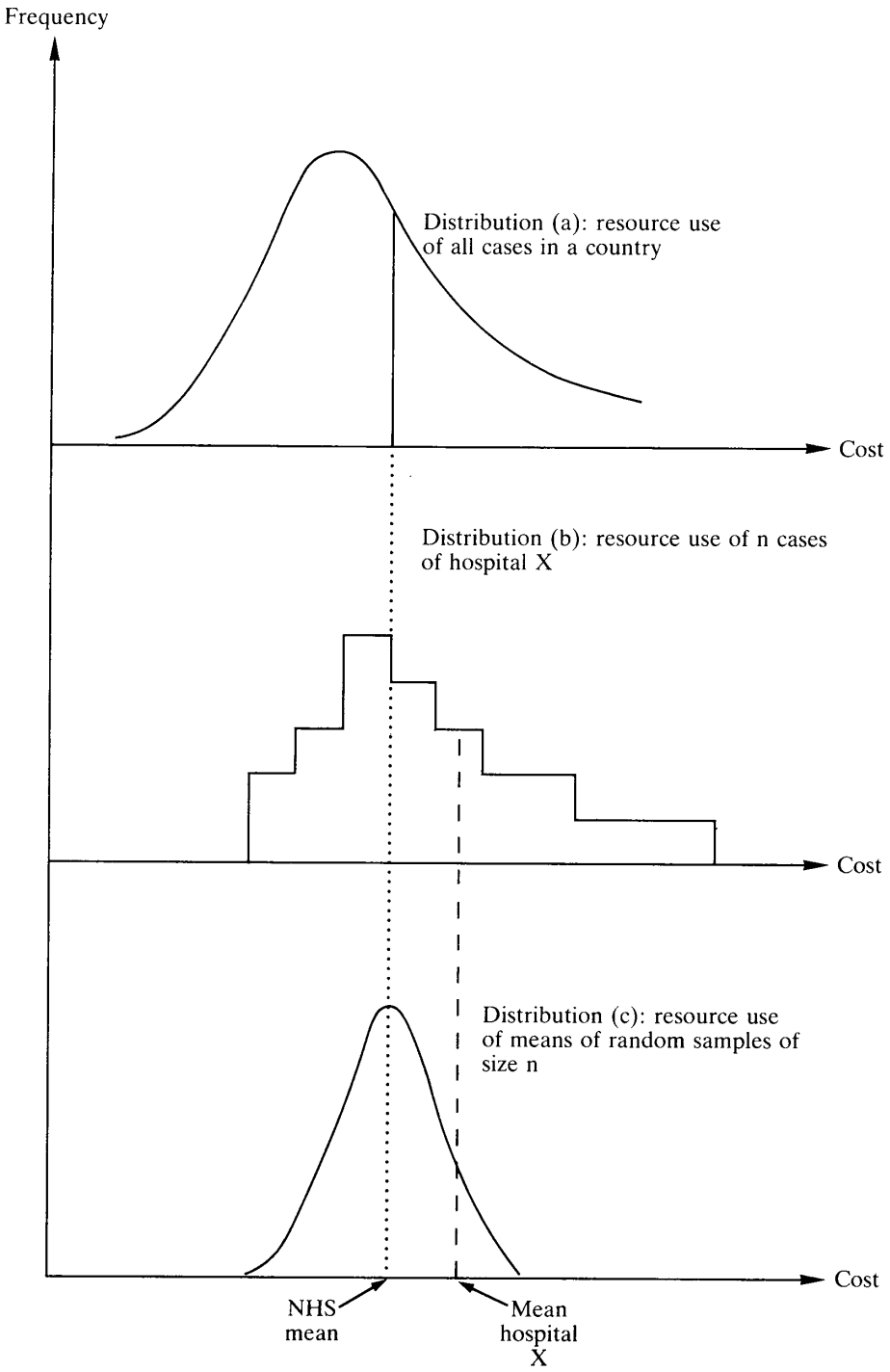
Introduction

Both the US and the UK have abandoned methods of financing health services which now appear to providers to have been relatively undemanding. Prior to 1976, the NHS had become accustomed to real annual increases in revenue of about four per cent; since 1976, total planned growth has been reduced, subjected to cash limits, and distributed to health authorities following methods based on the Resource Allocation Working Party[1,2]. US hospitals used generally to be reimbursed for the full costs incurred in treatment; now, their prospective rates for payment are set by DRG for Medicare patients, one of various initiatives introduced in the 1980s to contain costs of health care. The short time lag between the two countries' response indicated how pervasive and profound the problems of health service finance have become; even the richest country in the world cannot afford to ignore them. Action is, of course, being taken by virtually all developed countries – see Abel-Smith[3] for a review of cost containment in the EEC.

Although they face common problems, the initial responses of the UK and US governments have been markedly different. In the UK, resource allocation (RAWP) methods are intended to divide a cash-limited budget equitably according to relative need of populations. Prospective payment by DRG is intended to promote efficiency by rewarding admissions which cost less than the rate of reimbursement. Thus the US and UK governments have pioneered initiatives directed at important complementary problems, and each has had the opportunity learn from the other. Although the literature on RAWP and DRGs is mostly critical, such revolutionary changes are only implemented when they offer significant improvements on what they replace. It is now incumbent upon critics to propose better ways of measuring need and case mix: merely demonstrating imperfections of existing methods is of little value since perfection will never be achieved.

Criticism of DRGs in the US is naturally directed at their use for prospective payment. As hospital finance in the NHS is largely determined by capitation, the main value of DRGs here is likely to be in reviewing performance (the purpose for which the original DRGs were developed). This chapter draws on the debate in the US about what constitutes a homogeneous measure of case mix to propose how homogeneity can be most meaningfully defined; and uses results from recent DHSS performance indicators (PIs) to argue that specialties are inadequate for reviewing clinical performance. Criteria for choosing good measures of case mix for local reviews of performance are proposed; and ways of using more detailed measures of case mix in decision support systems to respond to performance reviews are discussed. The chapter points out that although DRGs appear to

Figure 10 Distributions of resource use of a typical DRG



be a significant advance on what is available for performance review in the NHS, their limitations need to be respected. There is still a need for research to develop further measures of output.

What is a homogeneous measure of case mix?

An important design feature of DRGs was the requirement that each group should be homogeneous with regard to resource use. Once this had been achieved and validated on a large data base, variability within a DRG became predictable and comparisons across institutions were possible.

Using DRGs for prospective payment in the US, though, has stimulated intense debate about the nature of a homogeneous measure of case mix and has shown how the definition of homoeogeneity can be affected by the purpose for which the chosen measure is used.

Using a case mix measure for prospective payment may encourage fallacious interpretations of the properties of a homogeneous measure. Distribution (a) of Figure 10 is typical of resource use of a population of cases with in a DRG: the distribution is skewed – the mean is greater than the median and a small proportion of cases accounts for a large proportion of resources. Prospective payment puts pressure on hospitals and physicians to try to ensure that each case treated costs less than the mean, and may even result in the mean becoming the upper limit for acceptable expenditure. Initially, these pressures may not cause serious problems in the US because there is likely to be considerable scope for reducing costs from the previous era of full cost reimbursement. In future, problems will become more serious if the mean is recalculated from data derived from the new era of prospective payment. Eventually, hospitals and physicians will be faced with certain admissions resulting in heavy losses. The for-profit hospitals and their physicians will want to avoid, or at worst minimise, the number of 'DRG losers'; for example, distal bypass for limb salvage[4].

Now consider using the DRG of Figure 10 to monitor hospital performance. Distribution (b) gives a histogram of resource use of the number of cases (n) in that DRG at a hospital (x) selected at random. For each DRG, a hospital's distribution of resource use would have a similar shape to that of the population if the cases treated and resources expended were a random sample from the population. Hospital x has a mean resource use slightly higher than the national average. Distribution (c) gives the expected distribution of mean resource use of random samples of size n from the population of distribution (a). The central limit theorem justifies the assumption that distribution (c) has a quite different shape from distributions (a) and (b): it will tend to be distributed normally (and not be skewed); and have a smaller variance than (a) or (b). Distribution (c) enables the performance of hospital x to be examined. What proportion of random samples of size n have a mean resource at or above the level of hospital x? A small proportion suggests the performance of hospital x ought to be closely examined for this DRG. Such an examination would be less justifiable for a large proportion.

There are three lessons for the NHS from the US debate about how

homogeneity can be most usefully defined. The first concerns the level and nature of the DRG's claim to be homogeneous. Greenhalgh and Todd[5] criticised DRGs for assuming that consumption of resources is the same for all patients in the same DRG, and remains the same throughout a patient's length of stay; and implying that patients are treated according to their diagnosis, regardless of their individual characteristics or treatment. They argued that these assumptions may be reasonable in longer term plans, but in the short term, and in decision making relating to small units, wards, or groups of patients, averaging out is insensitive. While some of these criticisms misinterpret the claims made for DRGs it is fair to note that their homogeneity is indeed based on averages and should be interpreted as implying that mean resource use by DRG by different hospitals ought to be the same. Homogeneity must be defined in terms of averages, because, as argued above, the concept becomes meaningless at the level of the individual patient[6]. Averaging does not deny the existence of variation, but allows for it by drawing on the power of the central limit theorem. Whether averaging over small numbers is insensitive is an empirical question which is explored below for DRGs.

The second lesson follows from the first, and concerns the level at which information is required on the use of resources in order to support performance review by homogeneous case mix measures. This could be done by hospital, specialty, physician or patient. Greenhalgh and Todd[5] report results from an experiment in patient-based costing in the NHS which pointed to marked differences in the use of resources per day between patients with exactly the same diagnosis and between patients requiring the same treatment procedure. They inferred that this demonstrated the need for patient-based costing to be introduced into the NHS. The availability of daily charges by each patient, however, has hardly helped the US manage its use of resources. To make sense of these data it has been necessary to find ways of grouping cases, and to focus not at the level of the individual patient, but on mean resource use by case mix group. This point is developed in our examination of the value of patient-based costing for the NHS.

The final lesson encompasses the purpose for which case mix groups are to be used. DRGs' claim to be homogeneous has been contested by showing that there has been variation in resources used for the same DRG between different hospitals[7]. Such evidence does not refute that claim, however, since variation in mean resource use by DRG by hospital could be accounted for by difference in efficiency in treating the same mix of cases. The index of 'severity' used by Horn et al does not necessarily provide an objective measure of case mix that is independent of resource use[8]. A problem that will trouble the US, however, is that even if DRGs were homogeneous across US hospitals in the past, the claim will not be true if hospitals respond differently to the incentives of using DRGs for prospective payment. Some hospitals, for example, would cease to operate on distal bypass for limb salvage. For the NHS, where finance is by capitation, using DRGs for retrospective review of performance lessens such problems.

Are specialties homogeneous measures of case mix?

Before the development of DRGs, clinical specialty had generally been the basic level at which case mix analyses took place. Indeed, on a routine basis this is still the case in the UK where acute hospitals are required to report costs by specialty, these costs being used for making comparisons. Thus it is important to know whether specialties are homogeneous and whether such comparisons will be valid; if they are, there may be no need for more detailed measures like DRGs.

For specialties to be homogeneous measures of case mix, mean resource use by specialty in different hospitals ought to be similar and predictable; but this has been disputed because teaching hospitals, for example, are likely to have a more complex case mix within specialties.

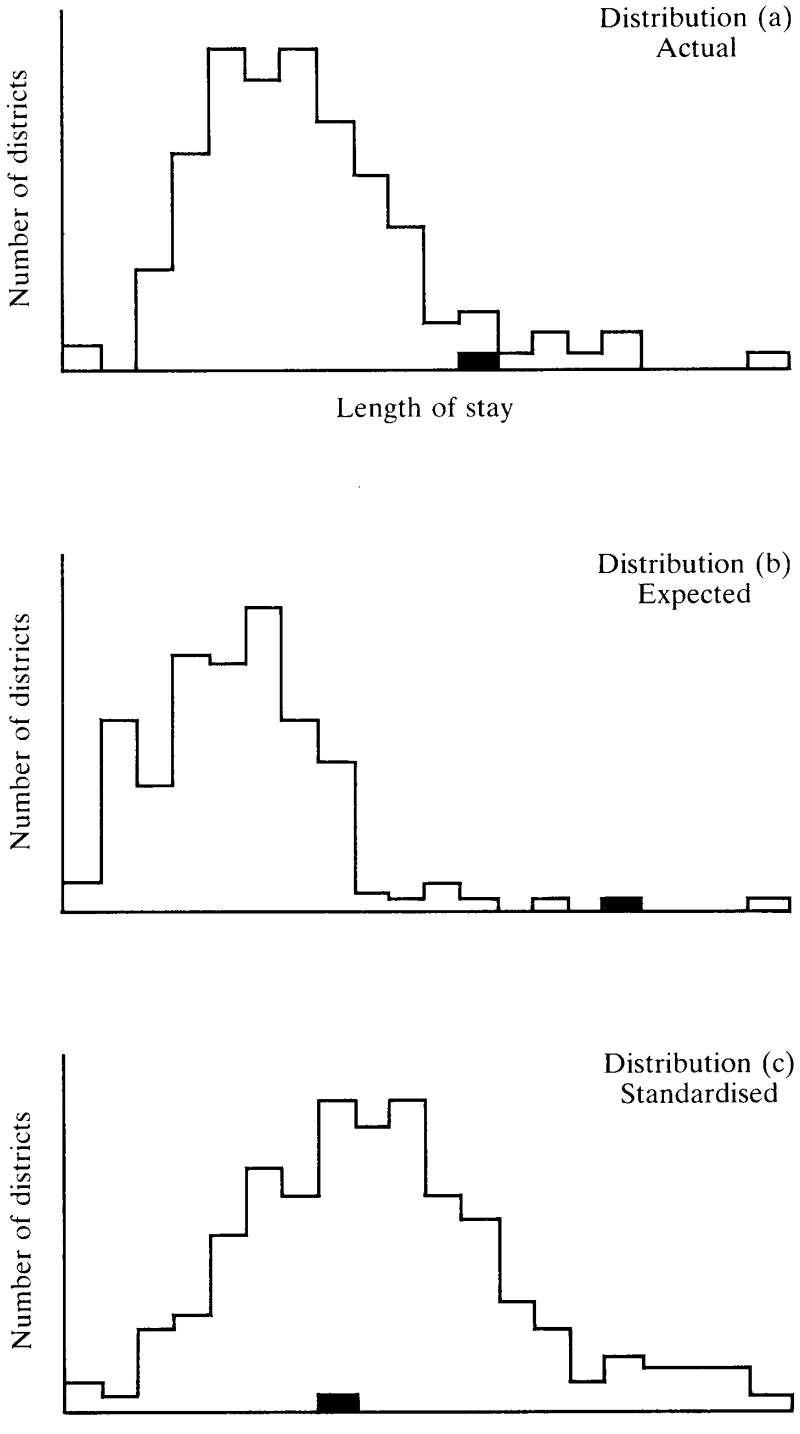
The introduction of specialty-costing was recommended by the Körner working group on financial information systems[9]. Research[10] showed substantial variations between hospitals in mean annual costs and lengths of stay by specialty. Thus the reported differences were *not* between *individual patients* in the same specialty, but between *hospitals' mean annual statistics*. Because differences may be due to variations in efficiency, it is necessary to examine case mix within them in order to judge whether specialties are homogeneous.

Using DRGs to compare case mix within specialties is complicated since many DRGs span a number of specialties. Fortunately, DHSS PIs for 1984 enable each hospital's length of stay by main specialty to be compared with the national average length of stay standardised for case mix. The measure of case mix used in these PIs is based on nearly 3,000 categories (referred to here as 'case mix 3000'): seven age groups; the two sexes; and the first three digits of ICD-9 classification (excluding psychiatric and maternity cases) which give several hundred diagnostic groups. Using case mix 3000 may well alter the impression given of a hospital's performance by specialty.

Figure 11 shows the national distribution of mean length of stay in 1983 from DHSS performance indicators for the specialty of general surgery (including urology) and highlights the position of a London teaching district. Distribution (a) shows the actual figures; the district's statistic is towards the right hand side of the distribution (with the obvious suggestion of relative inefficiency). Distribution (b) gives the distribution of expected mean lengths of stay by English health districts based on case mix 3000; that is, the number of cases actually admitted in each cell (of age, sex, first three ICD digits) multiplied by national average length of stay of that cell and averaged over all cells by the total number treated. This shows the district's statistic to be on the extreme right hand side of the distribution – greater than it actually is. Distribution (c) gives the ratios of actual to expected length of stay and shows the district's performance to be (just) on the left hand side (suggesting that it is slightly more efficient than the national average performance).

Thus distributions (a) and (c) give quite different impressions of efficiency in resource use for one specialty in a teaching district. This is true for most, but not all, of this district's specialties. This suggests that specialties do not provide homogeneous measures of case mix for hospital comparisons, and that the reported substantial differences in specialty costs may be due, at least in part, to differences in case mix.

Figure 11 Length of stay – general surgery including urology: West Lambeth and England (1983)



Choosing good measures of case mix for performance review

DHSS PIs give substance to objections that comparisons between the same specialty in different hospitals, and between different consultants in the same specialty in the same hospital, may be meaningless; they may not be comparing like with like. These objections can only be met by recourse to measures of case mix that are more detailed than specialties. Could case mix 3000 be used by staff in a English health district as an adequate supplementary description of specialties for local performance review instead of DRGs? Or does local use of case mix measures have different emphases and requirements from measures of national indicators of performance?

Whereas comprehensive measures are necessary for description by PIs of total hospital performance, scope for local improvements can be identified by choosing the most promising areas for investigation. Although this may leave a disturbing sense of incompleteness, it will serve to identify areas for detailed examination which would be missed by using a comprehensive index. A hospital's total caseload may well include cases of type A with high volume and low variance, and type B with low volume and high variance. If a hospital's mean length of stay for type A cases was greater than would be expected from random variation it would be identified by selective examination. In a comprehensive index for the same hospital, the 'noise' from type B cases could swamp a clear signal from type A. This problem can also be obscured by the way the national PIs are displayed. The distribution of mean statistics by district (for example, length of stay by specialty) does not reveal the degree of acceptable variation in each district's performance. The significance of variation would be obvious if the DHSS followed statistical practice and displayed for each district how its mean statistic compares with the distribution of sample means of the same size drawn at random from the NHS population.

How would case mix 3000 compare with DRGs if it were used for local selective review? Performance could be reviewed at three levels: hospital, specialty, and physician. The crucial managers of clinical resources are physicians, but the general problem with examination at this level is that the numbers of cases treated will often be too small for valid statistical comparisons – the problem of the insensitivity of averaging over small numbers, raised by Greenhalgh and Todd. This is illustrated by simulations of random samples of cases drawn from the populations of two different DRGs: DRG 26, seizure and headache, age 0 to 17, and DRG 127, heart failure and shock. The distribution of DRG 26 was described by Sanderson and Andrews[11] as lognormal and that of DRG 127 as 'odd' (both distributions being from the 1979 Hospital In-Patient Enquiry, HIPE). Experience shows that the majority of a consultant's cases in a year will fall within twenty DRGs, with between eight and 50 cases in each. Figures 12 and 13 give for both DRGs the distributions of populations (12a and 13a) and means of 100 random samples of size 10 (12b and 13b) and size 50 (12c and 13c). They illustrate the effects of the central limit theorem, even on samples of size 10. These and other simulations suggest that it is reasonable to assume that approximately 95 per cent of sample means lie within two standard errors of the mean for both sample sizes. On this basis 95 per cent (two-sided)

Figure 12 Distributions of lengths of stay for DRG 26: seizure and headache Age 0–17

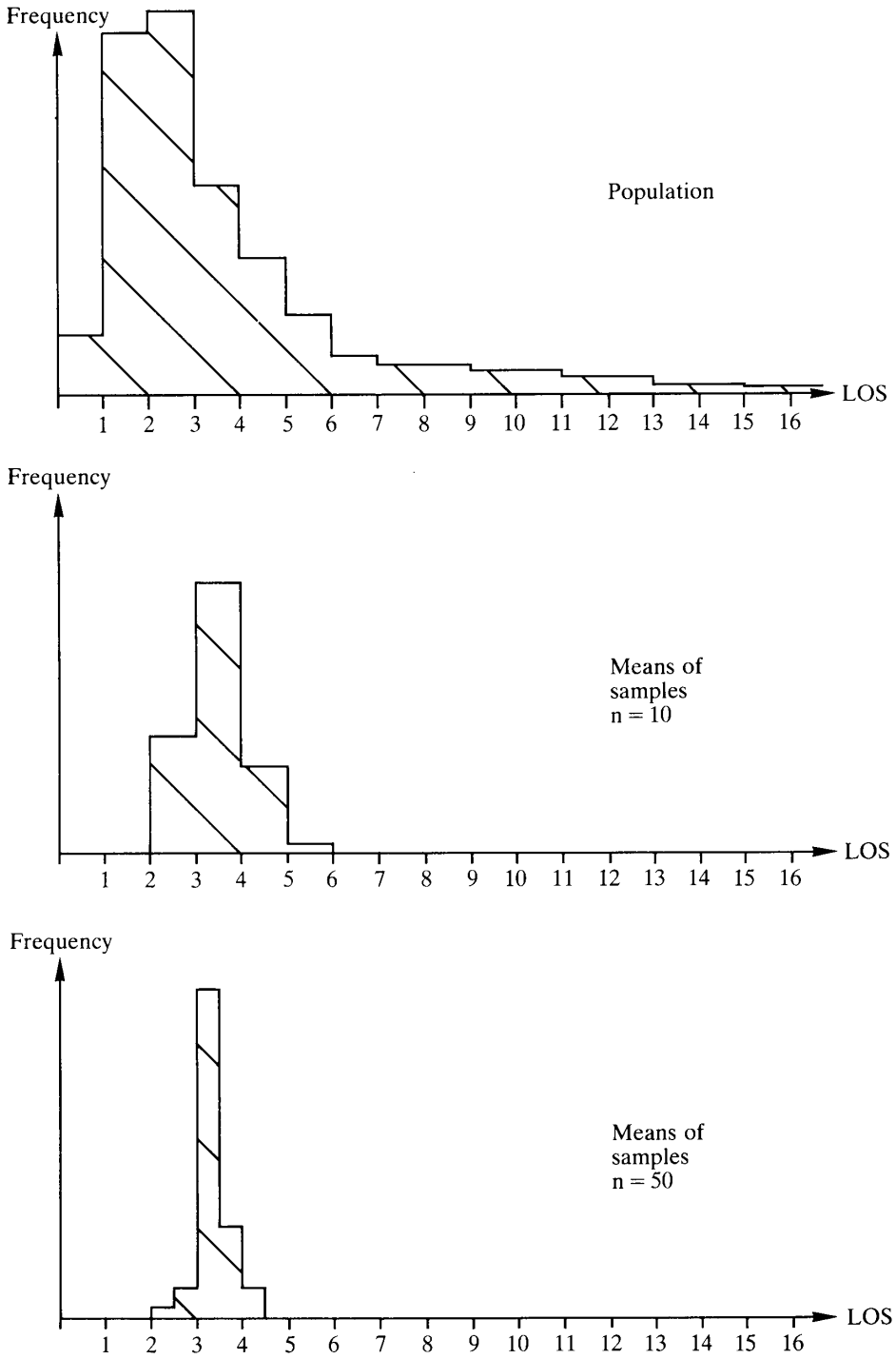
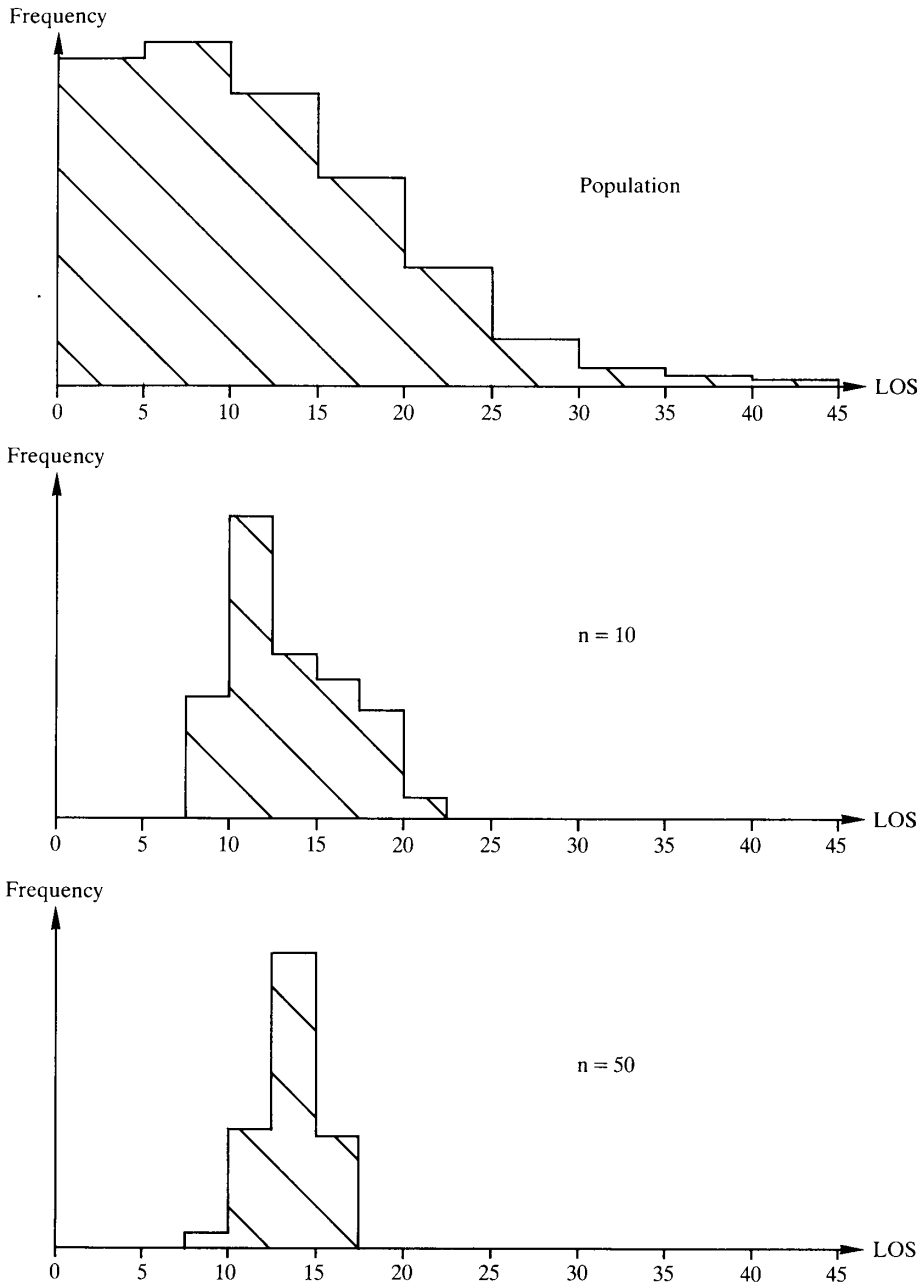


Figure 13 Distributions of length of stay for DRG 127: heart failure and shock



confidence intervals are given in Table 18 for the 20 DRGs which accounted for most inpatient days in the analysis by Sanderson and Andrews. Table 18 makes clear the problems of making inferences about performance from small samples, a particular problem in DRGs with a large variance. Consider, for example, a physician with ten cases in DRG 236. Acceptable performance could range from eight to 38 days. As examination of performance by homogeneous case mix groups will be more reliable the larger the number of cases in each group, more circumspection is necessary for examination at the level of the individual physician as compared with specialty or hospital.

The problems of making statistically valid inferences from the individual categories of case mix 3000 are much more severe than from DRGs, simply because case mix 3000 has at least six times as many categories. To overcome this handicap, case mix 3000 would have to be grouped, which requires a procedure similar to that underlying the generation of DRGs. But it is questionable whether case mix 3000 provides an adequate starting point for clinically meaningful groups. The original development of DRGs used diagnoses only while the revised set used for prospective payment partitions the major diagnostic categories into medical and surgical groupings, and uses diagnoses for medical groups and operations for surgical groups. While there remains a potential disadvantage in measuring case mix for prospective payment by operating procedures instead of diagnoses because it provides financial incentives to operate, when finance is by capitation and DRGs are used for performance review, the disadvantage is considerably lessened. Since surgeons are more likely to accept reviews of their performance based on procedures than on diagnoses, this modification reinforces the advantages of DRGs over case mix 3000 (or groupings of its categories) for performance review.

This comparison of case mix 3000 with DRGs has highlighted the importance of criteria which governed the development of the most recent set of DRGs: that there should be a manageable number (there are less than 500); they should be statistically stable (with a significant number assigned to each DRG); and they should be homogeneous and clinically meaningful (as discussed above). Not all DRGs are equally successful in satisfying these demanding criteria. Experience in using them for performance review in the NHS, in addition to reports from the US, will show which DRGs are reliable and which represent diverse clinical practice and need modification before comparisons are valid. Using DRGs for performance review has one advantage – it is not necessary to use all of them, a luxury which is not permissible when DRGs are used for prospective payment or, indeed, other financing systems which require full coverage of all cases. DRGs which reward investigation for performance review are likely to have these characteristics:

> They account for a significant volume of hospital resources. In different hospitals in different countries using cost data or inpatient days, 50 DRGs are consistently found to account for about 50 per cent of resources used.
> They include a reasonable number of cases so that the central limit theorem may be assumed to apply. DRGs which have the first characteristic will tend to have this one.

Table 18 DRGs ranked in order of resource use (total inpatient days) for 1979 HIPE

Rank	DRG	Percentage[1] cumulative freq. of total days	HIPE[1] mean length of stay (days)	95 per cent confidence intervals[2] n = 50	n = 10
1	467	5.31	7.84	5.39–10.29	2.56–13.32
2	14	8.45	19.28	12.21–26.35	3.45–35.11
3	209	10.62	22.48	18.94–26.02	14.56–30.40
4	234	12.60	8.08	5.69–10.47	2.74–13.42
5	122	14.58	11.33	9.88–12.78	8.08–14.58
6	355	16.38	11.30	10.46–12.14	9.43–13.17
7	127	18.14	13.20	9.88–16.52	5.78–20.62
8	233	19.56	15.72	10.87–20.57	4.87–26.58
9	243	20.85	11.74	8.89–14.59	5.37–18.11
10	82	22.11	11.41	8.21–14.61	4.26–18.56
11	210	23.35	30.98	23.38–38.58	13.97–47.99
12	88	24.51	12.89	10.07–15.72	6.57–19.21
13	236	25.57	23.30	16.47–30.13	8.01–38.59
14	270	26.61	3.47	2.48– 4.45	1.27– 5.66
15	39	27.61	7.92	6.98– 8.86	5.82–10.02
16	167	28.59	5.63	5.04– 6.22	4.31– 6.94
17	12	29.56	20.23	14.07–26.39	6.45–34.00
18	294	30.44	11.91	9.01–14.81	5.42–18.40
19	89	31.30	14.11	9.93–18.29	4.75–23.47
20	154	32.15	13.57	10.04–17.10	5.66–21.47

1 From Sanderson and Andrews (1984)
2 Assumes that 95 per cent of sampling distribution of means is within two standard errors of the population mean

> They set narrow confidence intervals for mean length of stay by hospital. This will be a matter of judgment about the acceptable size of the standard error of the estimate of the mean. It is not solely dependent on the variation in the distribution but also the number of cases admitted to the hospital chosen for examination. Thus 'good' DRGs will vary between hospitals, and it is not possible to give a universal list of 'good' DRGs, although certain groups will be found to meet this criterion in many hospitals.

This outline shows how flexible a hospital's approach to performance review needs to be. At any given time, different hospitals will find different subsets of their case types matching these characteristics. In future, it might be expected that revised groupings will become available which may enable reviews to be extended into important sectors of resource use where DRGs are now less homogeneous. However, hospitals will still need to develop methods of approaching performance review in 'poor' DRGs that they deem to be important, while paying due regard to their statistical limitations.

In examining the performance by a hospital of its own 30–50 'good' DRGs, two criteria can be used.

> One compares the hospital's actual length of stay for each DRG with a 95 per cent upper confidence interval limit from national performance. This would identify those DRGs where length of stay might be reduced. Performance by physician for these DRGs can then be examined on a similar basis, making due allowance for the smaller number of cases treated.
>
> The second is more direct; it is simply to examine performance for high volume DRGs individually, irrespective of whether the hospital's overall performance appears to be better or worse than the national average.

Enthoven[12] has described an encouraging example from the introduction of DRGs in the US. An orthopaedic surgeon saw that the DRG reimbursement rate for hip replacements was based on a mean national length of stay of 18 days, the same as the actual length of stay in his hospital. Work began to see how length of stay could be reduced.

> An industrial engineering analysis of the procedure enabled [the orthopaedic team] to reduce the time of the operation by adopting the techniques of the most efficient surgeons. This reduced blood loss and the need for transfusions. They decided to initiate the use of a continuous passive motion machine (for flexion and extension) sooner. This reduced post-operative pain and enabled patients to regain the full range of motion sooner. They started exercise on the first post-operative day. They introduced autotransfusion: patients deposited a unit of their own blood 4 and 2 weeks in advance of surgery. This substantially reduced complications from transfusions. And they did patient education preoperatively.

The average length of stay was reduced from 18 days to nine. Not only hospital costs were reduced; patients suffered less pain, had fewer complications and enjoyed faster recovery and earlier return to work and normal activities. Thus being more economical can also increase quality of care

(although the two are often said to be in conflict). As Enthoven points out, the reason for these changes not being made previously appears to have been a lack of incentive in a system of full-cost reimbursement. In the NHS the stimulus could come from generating enthusiasm through an equally close scrutiny of performance. Perversely it may be easier to generate enthusiasm when performance appears reasonable (that is, at or better than national average levels) and to challenge staff to do significantly better than the national average, than where performance appears to be poor.

Information requirements for implementing improvements in performance

Suppose critical examination of hospital performance identifies dramatic potential for reducing length of stay. How the health district chooses to improve performance may well be influenced by expectations about the hospital's future resources – whether they will increase or decrease. Treating more cases in the same number of beds will usually result in increased total expenditure for the hospital (although the average cost per case will typically decrease). Thus districts expecting more resources will tend to be interested in being more efficient by treating more cases, while districts with declining resources will try to treat the same number of cases with fewer resources.

The influence of a district's financial position on its response might suggest that the first requirement from information is detailed costs. Before costs are assigned, however, it is essential to know whether the proposed changes are feasible. If, for example, a district with increasing resources decides to treat more cases, a number of obvious questions follow: will there be enough beds of the right kind, enough sessions in theatres, and enough anaesthetic and recovery rooms; and will the proposed increase in cases treated alter training opportunities of learner nurses? Where extra capacity is necessary it needs to be identified, and then the cost consequences can be estimated.

Therefore, systems designed to report information for decision making about more efficient care ought to

be flexible in choice of case mix measures so that each hospital can choose its own 'good' DRGs and change its measures of case mix (as better categories become available) for its internal use, while retaining the ability to provide comparable information to other levels in the organisation for use in a multi district context;

be able to check the feasibility of proposed changes on relatively fixed resources, beds, theatres and so on to ensure that adequate capacity is available;

be able to establish what is necessary for changes to be feasible and estimate their costs by identifying at what point the proposed changes cause 'fixed' costs (such as staff) to vary;

and be able to estimate the directly variable costs of consumables associated with the proposed change.

Comprehensive patient-based costing might be seen as a flexible way of costing changes in case mix. But costing systems alone do not indicate how staffing is likely to change when the number and mix of patients treated is changed. It may be argued that average patient-based costs can be used to give estimates of marginal costs, and that all costs are variable in the long run. However, people in the NHS trying to keep spending within cash limits are unlikely to be convinced by this argument. As Keynes observed, in the long run we are all dead. Ways are needed of realising benefits in the short run by using DRGs, so that they are not discredited as another management fad, while developing more sophisticated approaches.

The argument here is not against patient-based costing as such, but rather that patient-based costing must not be seen as a panacea, and that progress can be made with less detailed costs. As we have argued, to make sense of costs by patient it is necessary to be able to group patients into useful categories (for example, in 'good' DRGs). Given useful categories, it may be possible to estimate costs at a more aggregated level than by patient for many types of resources. Providing comprehensive patient-based costs for heterogeneous categories in the expectation that these data might prove useful, is to squander effort unnecessarily. The NHS has enough information systems damaged by inaccuracy because they were developed for some unspecified purpose in the future rather than for well-defined, immediate requirements.

So if, as has been argued, patient-based costing systems are not the correct response to changing categories of case mix, what information structure should be used for assisting decision making? DHSS PIs show that specialties may not be homogeneous measures of case-mix. What then is the value of systems committed to reporting resource use by specialties? Should priority be given to ensuring that the categories to which costs are to be assigned are homogeneous? The difficulty is that perfect measures of case mix will never be available, and whatever categories are chosen they will be open to criticism.

One approach might be to regard specialties as cost centres for which DRGs provide a means of measuring their product mix, or of unbundling specialties into more homogeneous groups. Specialty-specific DRGs (SSDRGs) could be used for performance review to generate potential for increased efficiency. Options for change could first be specified as numbers of cases and mean length of stay by SSDRGs, and then translated into revised performance and associated demands by specialty.

An example is currently under consideration at St Thomas' Hospital, London, where it is intended to make use of a decision-support system, originally designed for hospital planning by specialty, which assesses the feasibility and costs of options[13]. For nursing and other ward-based resources, the system is given, for each option, data on planned cases, length of stay, and occupancy rates by specialty. It then checks whether there will be enough beds in the available wards and assesses the impact on the nurse training programme. The system estimates resource use on existing practice, which decides staffing by ward, by calculating how many wards are required to make the necessary number of beds available, and the staffing of each ward. The requirements for nurse training are first taken as a constraint that has in some way to be satisfied, although the system can be used to investigate the impact

of varying the assumptions about training. Surgical specialties are provided with estimates of the demands of options on theatres and anaesthetic and recovery rooms, and whether or not the sessions planned will be adequate. The change in total hospital costs is estimated by summing the changes in fixed costs – which enable the planned number and mix of cases by specialty to be treated – and the changes in costs of consumables. For example, costs of heating and cleaning are included, but the effects of changes in staffing of diagnostic departments are excluded, in the belief that they are most fruitfully explored through budgeting systems.

At St Thomas' the interest is in the information that can be produced from the specialty-based planning system, specialty costs and clinical budgeting to support performance review by using more detailed measures of case mix than specialties. Thus for options developed by performance review by SSDRGs which are intended to alter case mix within specialties, it is necessary to estimate the effects of these changes on the resources of each specialty. Since length of stay is routinely reported for each patient, the effects on length of stay by specialty of proposals which alter case mix and performance by SSDRG can be estimated directly. No data are routinely reported which would provide revised estimates by specialty for theatres and consumables. However, as demonstrated in chapter 6, the introduction of specialty costing and clinical budgeting will provide an improved basis for estimating these effects. Unbundling specialty data to the level of DRGs to enable revised estimates of each specialty's demands to be produced, is likely to provide adequate information for immediate action. Waiting until patient-based costing of all resources is in place means deferring action indefinitely: and experience with approximate methods will highlight where more exact information will be useful. Seeking to produce accurate, comprehensive data for each patient without specifying their purpose means that it is impossible to know either the data's appropriate form or the degree of accuracy required.

The scope of DRGs

This chapter has pointed out the advantages of using DRGs compared with other measures of case mix available to the NHS, but to balance the account, it is important to recognise their limitations.

DRGs apply to acute inpatient care only. The efficiency of day or outpatient care cannot be examined by using them. Neuhauser and Pine[14] have illustrated how using DRGs for prospective payment for inguinal hernia may generate incentives against substituting day and outpatient care for inpatient care, and lead to inefficiency in and poorer quality of care. It would be interesting to know whether using DRGs for reviewing performance would enable an assessment to be made of the potential to substitute other forms of care for inpatient care.

The use of DRGs for psychiatric and maternity services has not been considered. US experience suggests that psychiatric DRGs are less reliable[15, 16]. Maternity cases have only been excluded because they are not recorded by Hospital Activity Analysis – the basis for pioneering work on DRGs.

DRGs do not, of course, measure genuine efficiency which requires

measurement of output in terms of change in health. Unfortunately routine measures are not available for this purpose, although mortality can be applied to some conditions[17]. This is a particular problem in using DRGs for prospective payment, as finance by activity (as opposed to capitation) generates incentives for medical intervention. Wennberg et al[18] argued that the ubiquity of large variations in hospital admission rates by DRG provides scope for responding to the system of prospective payment by reducing hospital stay and increasing admission rates, which is only a superficial increase in efficiency: for example, removing healthy appendices risks life and wastes resources, regardless of its cost. There is an urgent need for research to develop methods of assessing appropriate levels of care – probably through the measurement of outcomes – in order to supplement the current work on measuring the inputs used by specific case types.

Thus, the scope of measuring performance has been limited in this chapter to examining resources consumed by acute non-psychiatric inpatient episodes, excluding maternity care.

Conclusions

This chapter has argued that DRGs, although they have attracted considerable criticism in the US, ought to improve our capacity to review NHS performance. The US criticism is in part because DRGs are used to contain costs, which inevitably results in objections. They are also criticised because of their use for prospective payment. No measure of case mix can be perfect, yet of those available DRGs currently appear to be the most promising. If an obviously superior measure were available it would probably be in use now in the US.

As variability is intrinsic to the treatment of patients, it is important that performance is reviewed by examining average levels over significant numbers of patients in the same case mix groups. The purpose of review is to improve performance and because physicians are crucial in realising improvements it matters that the groupings which are used for comparison are meaningful to them. Grouping cases into specialties ensures that large numbers of cases are assigned to each group, but the groupings are not clinically meaningful. The categorisation currently used by the English performance indicators results in few cases in a large number of categories and is probably weak in reviewing surgical performance since it is based on diagnoses, and not operations. DRGs, on the other hand, were designed to be clinically meaningful and result in significant numbers of cases incurring significant resources being assigned to at least 50 DRGs in any given hospital.

One obstacle, though, to using DRGs for performance review lies in the estimation of their costs. To wait for comprehensive patient-based costing is to postpone action indefinitely. Specialty-based approaches may be adequate when coupled with methods which unbundle demands of specialties by DRG. Furthermore, the first step in appraising options for improved efficiency is assessing their feasibility rather than estimating detailed costs.

DRGs look to be better than existing alternatives of measuring case mix for the purpose of performance review. Those interested in making immediate use of them ought not to be deterred by what may be transient difficulties.

References

1 Bevan G, Copeman H, Perrin J and Rosser R. Health care: priorities and management. London, Croom Helm, 1980.
2 Bosanquet N. Public expenditure on the NHS: recent trends and outlook. London, Institute of Health Services Management, 1985.
3 Abel-Smith B. Cost containment in health care. London, Bedford Square Press, 1984.
4 Rhodes R S, Krasniak C L and Jones P K. Factors affecting length of hospital stay for femoropopliteal bypass. New England Journal of Medicine, 314, 1986: 153–157.
5 Greenhalgh C A and Todd J N. Financial information project: message for the NHS. British Medical Journal, 290, 1985: 410–411.
6 Jencks S F, Dobson A, Willis P and Feinstein P H. Evaluating and improving the measurement of hospital case mix. Health Care Financing Review, 1984, annual supplement: 1–11.
7 Horn S D, Bulkley G, Sharkey P D, Chambers A F, Horn R A and Schramm C J. Interhospital differences in severity of illness. New England Journal of Medicine, 313, 1, 1985: 20–24.
8 Lakhani A, Mulrow C, Prochaska A and Bevan G. Case severity: a new measure? London, Department of Community Medicine, UMDS, St Thomas' Hospital, 1986.
9 DHSS. Steering Group on Health Services Information. Sixth report. A report on the collection and use of financial information in the National Health Service. (Chairman, Mrs Edith Körner). London, HMSO, 1984.
10 Hillman R L and Nix G R. DHSS-funded research into specialty costing 1980–1982. London, DHSS, 1982.
11 Sanderson H F and Andrews V. Monitoring hospital services: an application of diagnosis related groups to hospital discharge data in England and Wales. London, Department of Community Health, London School of Hygiene and Tropical Medicine, 1984.
12 Enthoven A C. Reflections on the management of the NHS. London, Nuffield Provincial Hospitals Trust, 1985.
13 Beech R, Brough R and Shaw M. Aiding and abetting the planners. Health and Social Service Journal, 25 July 1985: 930–931.
14 Neuhauser D and Pine R. DRGs and elective surgery. Medical Care, 23, 2, 1985: 183–188.
15 Taube C, Lee E S and Forthofer R N. DRGs in psychiatry. Medical Care, 22, 7, 1984: 597–610.
16 Davis G C and Breslau N. DRGs and the practice of psychiatry. Medical Care, 22, 7, 1984: 595–596.
17 Charlton J R H, Hartley R M, Silver R and Holland W W. Geographical variation in mortality from conditions amenable to medical intervention in England and Wales. Lancet, 26 March, 1983: 691–696.
18 Wennberg J E, McPherson K and Caper P. Will payment based on diagnosis-related groups control hospital costs? New England Journal of Medicine, 311, 1984: 295–300.

9 DESCRIBING CLINICAL WORK IN OPHTHALMOLOGY

Antoinette B Newman

Introduction

It is against the background of a developing need for output costing that this study was conceived. In the early 1970s, a working party of the King Edward's Hospital Fund for London, chaired by Professor Brian Abel-Smith, was asked 'to examine how economic principles could best be employed within an integrated health service to secure a better use of resources made available for the service at both the planning and operational levels'[1]. Their 1973 report, *Accounting for Health,* concluded that the main obstacle to improved decision-making about the use of available resources was the information systems. Specific to financial information, the report concluded that its form at the time

> .... does not lend itself to the discussion of medical priorities, because breakdowns of expenditure and costs are generally given for broad categories which are not subdivided in such as way as to be of use in costing particular objectives of service. The essence of the problem is, therefore, to determine a new basis for subdividing total expenditure on the health service[2].

This new basis for organising financial information would have two features: it would separately identify and cost the activity components (inputs) which relate to specific outputs of the system; and the objectives or outputs of the system to which activity units are aggregated would include classification by medical condition or diagnostic group. '... if units of activity are to be medically useful and outcome-oriented, classification by operational units (hospital or general practice) or specialty, is not sufficiently precise. Classification by the condition or problem of the patient seems essential'[3]. The report went on to recommend the development of a disease classification system.

Contemporary with the working party, other theoretical and practical exercises in the United Kingdom were reaching similar conclusions. The common theme in these studies was the inadequacy of the existing input-focused accounting system and the need for the development of costed output measures which would allow meaningful analysis and comparison in order to improve the effectiveness with which scarce resources were allocated.

In the introduction to his *Patient Costing Study,* Russell observes '... an ever increasing concern for using the limited resources of the health service to the best effect for the health of the patients has aroused much interest in methods of measuring the use of services by individual patients ... One of the major obstacles in such evaluation has been the difficulty of obtaining a realistic costing of the treatment given to a single patient. The present system

of hospital costing . . . gives no indication whatsoever of the use by individual patients and therefore can give little guide to the financial implications of any change in policy or practice at ward level'[4]. The practical costing studies carried out were congruent with *Accounting for Health* in focusing on disease groups as the descriptor of output but Babson observed in his 1973 report, *Disease Costing,* that '. . . although the importance of case mix on hospital costliness has been demonstrated, this is in no way compensated for in the data generated by the present costing system'[5].

However, none of the work carried out in this period developed a comprehensive disease classification system as recommended by Professor Abel-Smith's working party. Rather, each study selected a limited number of particular diseases to cost.

The work done on diagnostic group costing in the early 1970s went into abeyance for a variety of reasons which together resulted in the effort being perceived as greater than its worth. The need for an output as well as input focused accounting system did not abate and has subsequently led to further work, some with different output measures such as costs by specialty[6] and patient costing[7].

The need for output costing has been re-emphasised by recent developments in the National Health Service, most notably following the report of the Griffiths inquiry[8]. In the years since the early reports already mentioned, the expertise required for the development of such an accounting system has expanded markedly. Activity measurement in service departments is becoming more common while a comprehensive yet manageable case mix classification system became a reality with the development of the diagnosis related groups (DRGs) system in the United States.

Overview of the diagnostic group costing study

The study was conceived in 1983 with the aim of developing and applying a financial information system which would relate costs to health outputs described in terms of diagnostic groups. The system was intended to be sufficiently robust and flexible to be useful to both management and clinicians in a range of decision-making activities.

Diagnostic groups were selected as the definition of output for two reasons. First, a level of description was sought between the specialty level and the individual patient which would provide greater opportunity for looking at variation in resource use than specialty data without running the risk of an unmanageable quantity of information from patient level data. Secondly, the study sought to define output in a vocabulary that was helpful to clinicians when describing the way they plan and manage care, thereby facilitating dialogue on resource management. Diagnostic groups were the primary level of categorisation but the study sought to maintain flexibility by allowing other groupings that could be appropriate to particular needs. For example, aggregating diagnostic groups within a specialty might be useful in order to make them coincide with clinical organisation and its decision-taking structure.

The study intended to utilise a disease-based scheme of patient classification but had no predetermined ideas. After analysis of existing options and a

study of the potential of a new system, it was decided that the DRGs developed at Yale were the most thoroughly researched, best tested and most widely accepted. Consequently they were selected for use in the inpatient part of the study.

The study's emphasis was on the integration of the costing methodology in the overall management systems of the health district. It planned to utilise cost accounting techniques in order to produce costs as accurately as possible, and to develop a methodology that could form part of the on-going accounting system. Prominence was given to the production of costed information by patient groups, but detail of activity and resource use were also collected to enhance the scope of the analysis. A major objective was to provide an accurate assessment of the total and the component costs of treating types of patients, information that would allow assessment of the factors contributing to cost variation and identification of opportunities for changes in practice.

The study was based on a single specialty. It was thought that this limited framework would facilitate data management while providing sufficient scope to investigate the utility of sub-dividing a specialty into case mix groups. An eye hospital, the Western Ophthalmic Hospital, was selected as the study site. It had the advantage of being a relatively small, self-contained unit with a single specialty.

The principle data collection was carried out prospectively from 1 October 1983 to 29 February 1984. Data was collected for all inpatients (627) and day cases (50) admitted and discharged during the period and for all direct patient care services. Outpatient clinic data obtained from a two week study in August 1983 produced a sample of 495 patient visits. Costing the resultant activity data involved the development of weighted units of resource use and the identification of costs from the existing accounting system. The final assignment of costs to activity is by and large complete and a report on the whole study is expected in 1987.

This study will produce results of two kinds. First, costed inpatient DRGs which are descriptive of the hospital's workload over a given period of time. This will encompass the total cost per case by patient as well as a breakdown of activity and cost by service and by day of stay. These data can then be summarised as the average cost per DRG and as the distribution within a DRG. The second result will be an assessment of the feasibility of implementing a diagnostic group costing system on a routine basis.

The data bases produced will provide a description of variation in services used, associated resources required and resultant cost by and within case type. This information provides a refined and sensitive measure of resource use with which to plan, set budgets, manage operations and monitor performance. It must be emphasised that it is only a means toward improving managerial decision-making and not an end in itself. To be effective, it must be taken up with other types of managerial inputs and then acted upon.

The following sections describe in more detail selected applications of costed diagnostic groups, using preliminary study data to illustrate the analyses.

Regional and inter-district comparisons of case mix variations

One reason for interest in case mix measures is the extent to which variation in the types of cases treated in different geographical areas accounts for variation in resource use and cost. At present, the United Kingdom inter-area variation is analysed by groupings of clinical specialties to provide comparisons relating to activity, costs and funding. There is some concern that data at the specialty level is too broad and does not permit control over variation in resource use caused by case mix differences.

The first analysis conducted with study data was a preliminary exploration of this issue of case mix variation in a specialty. The distribution of inpatient and day cases by DRG within the specialty of ophthalmology in the study district was compared to similar distributional data from other districts within the same region, and with regional and national data. Comparisons were based on activity data, including the number of cases and length of stay. Some costing information will be incorporated later although comparable data for other districts will not be immediately available. The principal data base used for the analysis was 1982 hospital activity analysis data from the North West Thames Regional Health Authority (NWTRHA) for all the districts in the Authority, including Paddington and North Kensington the district responsible for the Western Ophthalmic Hospital.

A prime consideration in the use of DRGs was their association with clinical specialty-based information. The majority of eye-related conditions are covered by DRGs in major diagnostic category MDC 02 – Diseases and disorders of the eye. Similarly, there is a high correlation between MDC 02 and the specialty of ophthalmology. In NWTRHA 91.5 per cent of the cases in MDC 02 were discharged by the specialty of ophthalmology. Conversely, 94.4 per cent of ophthalmology discharges fell within MDC 02.

Another aspect of the value that DRGs have in comparative analyses is the extent to which they highlight variations in case type distribution between districts. While MDC 02 is consistently dominated by several DRGs, the study data showed considerable proportional variation in DRGs across districts, the dominant DRGs being 039 Lens procedures, and 040 and 041 Other extraocular procedures. The range in a DRG's percentage of total cases by district was broad for these dominant DRGs as well as the balance as evidenced in the examples in Table 19.

Table 19 Highest and lowest percentages of cases in selected eye disorders

	Percentage of total cases in MDC 02 by health district	
	LOWEST	HIGHEST
	(of 15 districts)	
DRG 039 Lens procedures	18	54
DRG 040 Other extraocular procedures, 18 yrs +	6	20
DRG 041 Other extraocular procedures, 0–17 yrs	10	25
DRG 036 Retinal procedures	1	8
DRG 042 Other intraocular procedures	1	8

One observation from these data is that the degree of variation in DRG distribution increases as the size of the geographical divisions decreases, suggesting that they are more useful on a small area basis. In this case district comparisons showed considerable variation, whereas data at the regional level showed less variation with local differences being masked in the aggregation.

Possible factors responsible for variations in the distribution by DRG were suggested by the data. The existence within a district of a specialised, regional referral centre for ophthalmology services was reflected in a greater proportion of the more specialised surgical DRGs, namely retinal (036), orbital (037) and other intraocular procedures (042), than might have been expected. In other districts, particular areas of specialisation for example by certain types of procedure, condition or age group, was a factor and explained a higher proportion of cases in DRGs relating to retinal procedures, neurological conditions and paediatric cases respectively. Variations in these practice patterns will affect the number of cases in certain DRGs. The DRG data analysed here only include acute inpatient and day cases. If certain procedures such as minor operations or laser treatment are performed on an outpatient basis in some districts, they would not be reflected in the associated DRGs. Sometimes this may affect the volume of inpatient work in a particular district, which also appears to influence the distribution of cases across DRGs. However, volume is much more influential on individual DRGs and the distribution of resource use within them. There are factors other than the type and pattern of service provision which would promote differences in the case mix seen across districts. Demographic patterns, epidemiologic characteristics of the population and doctor referral patterns would all be expected to cause differences. It should also be noted that at this early stage of application of DRGs in the UK variations in the coding practices of districts may affect the reported incidence of individual DRGs.

Several approaches to interpreting the impact of distributional variations were explored. One was the analysis of select or 'indicator' DRGs which were thought to represent more or less specialised activity. Within MDC 02, more specialised work was thought to be represented in the following DRGs:

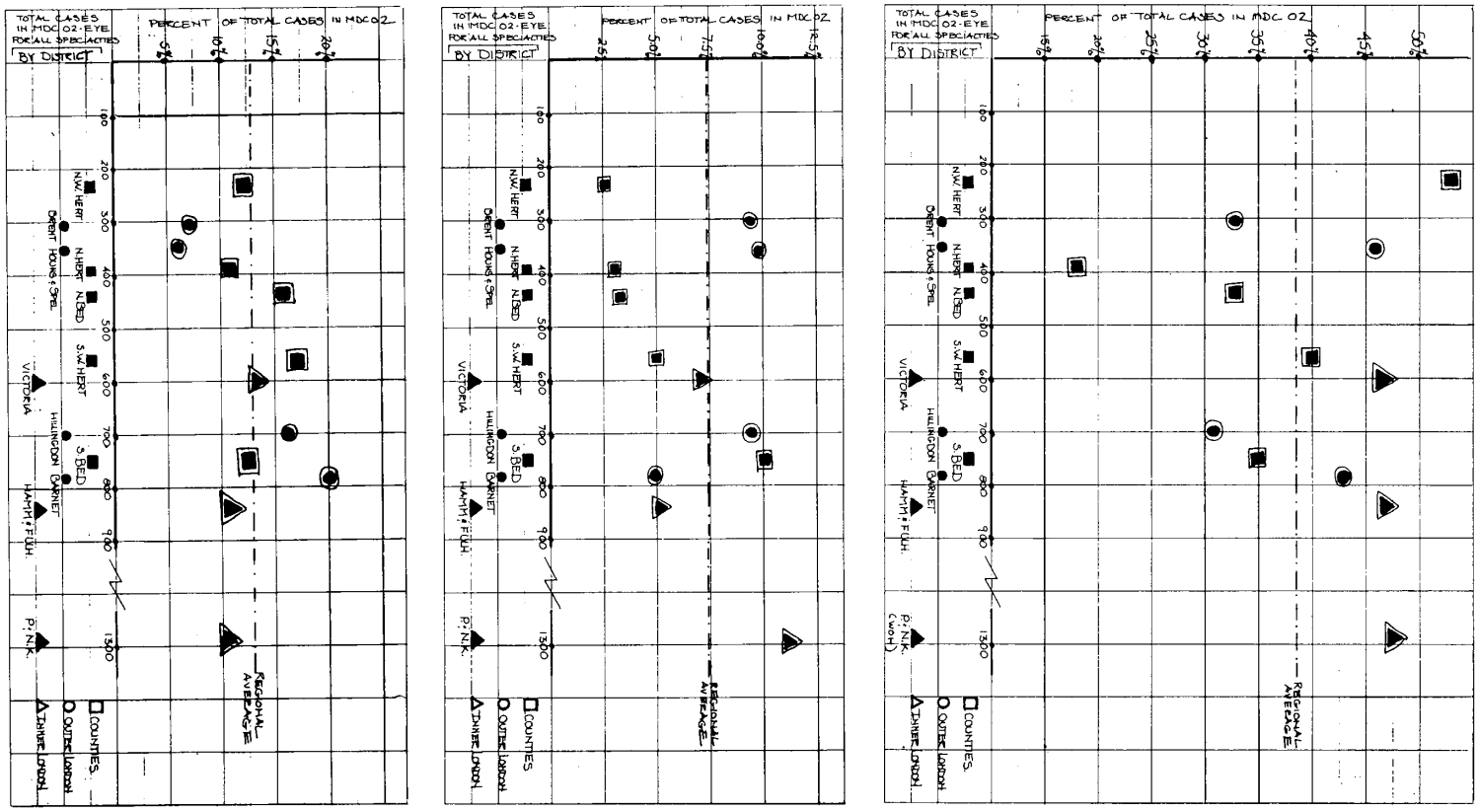
036 Retinal procedures
042 Other intraocular procedures
037 Orbital procedures
039 Lens procedures

Conversely, DRG 040 Other extraocular procedures, 18 years + was composed largely of eyelid procedures and, therefore, thought to represent less specialised work.

Figure 14 displays each district's proportion of the above DRGs in relation to the district's overall MDC 02 caseload and geographical location. Each district's total cases treated in MDC 02 is plotted below the axis; immediately above it is the proportion of cases falling in the DRG(s) under consideration. Other activity-based assessments could include

1 a district's percentage of the region's total cases within a given DRG in relation to its share of all cases in the MDC;

Figure 14 North West Thames RHA district data – 1982: selected DRGs as a percentage of total cases in MDC 02



DRG 040 – other extraocular proc. 18 yrs+

DRGs 036 – retinal proc.
037 – orbital proc.
042 – other intraocular proc.

DRG 039 – lens procedures

2 a ratio for each district of the more specialised DRGs to those considered less specialised; and
3 a ratio of surgical to medical DRGs.

While activity based measures are descriptive and useful, it is when the measures of resource are applied to DRGs that their value to management becomes manifest. Such measures can be expressed as absolute costs or as weighted measures of relative resource use as identified in Chapter 6. The difficulty at present in the UK is identifying and calculating comprehensive and appropriate measures. The only complete sets of costs by DRG now available are those developed in the US for reimbursement by state or federal government. Their applicability here is uncertain as

1 their basis for development is often questioned in the USA and
2 their comparability to UK resource use and cost practices has yet to be determined.

Ultimately the most appropriate cost weights will be those developed specifically for the UK, and considerable work is now being undertaken in this area.

Intra-district description of case mix in different patient care settings

While DRGs provide a meaningful and manageable basis for describing output at the national and regional level, in many respects the range of their application increases at a more local level. An example of a district-based analysis is the description of activity and resource use by 'disease groups' across and within different settings of patient care delivery. The concept of disease groups has been introduced because they provide a categorisation scheme which allows linkage across different settings at a sufficiently detailed level for analysis. Also they use a vocabulary which is meaningful to clinicians from the standpoint of treatment patterns and protocols. DRGs only relate to inpatient and day care while the equivalent ambulatory visit groups (AVGs) are specific to outpatients. Disease groups span the two settings.

A profile of disease groups served by settings provides a basis for a number of clinical and management decision-making activities. Of particular relevance at a time of limited resources is the consideration of alternative practice patterns. In the search for improved cost effectiveness, clinical practice patterns have not been fully investigated for both political and practical reasons. This should now take place. One consideration should be whether the setting in which care is provided for a particular disease group is the most appropriate and economical patient care.

A disease based classification system can indicate where cases are being treated as inpatient, outpatient or day care, and the relative cost of each. This can then serve as a basis for a clinical assessment of the appropriateness of the present distribution, the extent to which alternative methods of care are available and whether they would be more cost effective and/or improve the quality of care.

Alternatively the volume and distribution of case types being treated in an

individual setting can be studied, in order to examine the appropriateness of the pattern of care being delivered and the opportunities for expansion in the range of case types served. Using case type activity and resource use data by care setting, the balance of resources in relation to workload can be assessed. Analysis of relationships may reveal excess provision in one area and under provision in another while the resource implications of planned or projected changes (for example, the effect of early discharge on the need for community support) can be estimated.

Although the diagnostic group costing study focussed primarily on inpatient care it also included an analysis of outpatient care to permit the range of analyses outlined above. This information is of particular importance in ophthalmology where outpatient activity predominates. Table 20 displays some preliminary results from the study's sample data. Workload data by settings of care is presented for particular disease groups, showing the volume of activity and its relative share of all activity in that setting. It is evident that there is considerable variation across disease groups with regard to the setting in which care is provided. Similarly, each setting serves a different mix of patients. The impact of new technology, for example laser treatment, can also be seen and provides a framework to evaluate the impact of such developments on future provision.

The depiction of disease groups across settings reinforces the point made in the previous section – that to rely solely on inpatient and day case data for cross district or regional comparisons of the type of cases served, may be misleading. If practice patterns vary among districts, with some making greater use than others of outpatient settings for certain case types, inpatient DRGs alone will not reflect this variation. Examples in this particular study are people with glaucoma and retinal disorders, the majority of which were being treated by laser as outpatients in the Western Ophthalmic instead of being admitted as inpatients. This does not however detract from using DRG distributions to assess cost variations specifically within the inpatient workload. It is a different issue.

Analyses of activity and costs by care setting similar to those presented in Table 20 depict the relative cost of providing treatment in different settings. They also display the relative balance and relationship of activity and resources between and within settings. Both provide a useful starting point to assist clinicians and managers. However, it must be noted that these analyses are cross-sectional. A longitudinal profile of the number and types of services provided for a particular disease group over a period of time is needed to ensure consistency and to assist other kinds of decisions. When information about a series of treatments or visits is required, cost comparisons of alternative treatment patterns would need to be based on time series data. Similarly, to project the impact of new outpatients in a disease category on bed use it would be necessary to know the expected frequency of outpatient visits over a period of time and the proportion of those eventually admitted. The pilot study collected historical data for the outpatients seen and for selected inpatients which will appear in the final report.

Table 20: Utilisation of patient care settings by select disease groups, Western Ophthalmic Hospital: study data October 1983 – February 1984

Disease group	Case mix group into which disease falls		Selected outpatient services						Inpatient services			
	Ambulatory visit group	Inpatient DRG	Ophthalmology clinics (visits)		Orthoptics (outpatient visits)		Minor operations (outpatient visits)		Laser (outpatient visits)		Day cases (cases)		Inpatient (admissions)	
			No	% of setting	No	% of setting	No	% of setting	No	% of setting	No	% of setting	No	% of setting
Eyelid disorders	External eye disorders	(040,041) Extraocular procedures except orbit	340	2.9	—	—	288	93.2	—	—	31	62.0	17	2.7
Squints	Visual disturbances, strabismus, motility disorders	(040,041) Extraocular procedures except orbit	365	3.1	727	75.3	—	—	—	—	—	—	64	10.2
Cataract, aphakia	Cataract aphakia pseudophakic	(039) Lens procedures	3,150	27.2	—	—	1	0.3	—	—	—	—	325	51.8
Glaucoma	Other anterior segment and optic nerve	(038) Primary iris procedures	1,405	12.1	—	—	—	—	60	13.1	2	4.0	17	2.7
Retinal disorders	Retina choroid and vitreous	(036) Retinal procedures	2,080	18.0	—	—	6	1.9	374	81.8	2	4.0	60	9.6
Total, all causes in patient care setting			11,580		966		309		457		50		627	

Disease groups have been identified by the study.

The case mix groups are those developed by the Health Systems Management Group of the Yale School of Organisation and Management including both inpatient diagnosis related groups (DRGs) and the ambulatory visit groups (AVGs) from the ambulatory patient classification system.

As the two classification schemes do not map into each other precisely, select disease groups have been chosen for the analysis and the respective case mix groups into which they fall indicated. In most instances the disease groups represent only a portion of the cases falling within a case mix groups.

Study data:
Inpatient, day case, laser, minor operations and orthoptics data are from a full five month data collection over the study period.
Outpatient clinic data was obtained from a one week sample in August 1983. The resulting proportion of cases by disease groups was applied to the actual total clinic visits during the study period.

Use of hospital resources by DRG and their associated costs

This final section addresses a principal objective of the study, the development of detailed costs by DRG by service department. The costing of cases by DRG is complete and preliminary results of average costs for all inpatient cost categories by DRG for MDC 02 are shown in Table 21. It must be appreciated, however, that these costs are specific to one district and one site and are not nationally based. As clinical budgeting develops in districts across the country a wider based set of cost estimates should become available.

The study defined three categories of costs and assigned them to inpatient cases, as follows:

Direct costs: those costs associated with and varying with individual patient case types and which are directly attributable to them.

Indirect costs: those costs associated with and varying with the number or volume of patients served, irrespective of case type. In the majority of departments, these costs were assigned to patients on the basis of patient days.

Overhead costs: those costs which do not vary with either the type or the volume of cases in the short term. These costs were apportioned to cases on the basis of patient days.

As evidenced in Table 21, the study found that 62 per cent of inpatient costs were directly attributable to case type; another 22 per cent varied with volume, leaving 16 per cent as overhead.

Average costs by DRG make it possible to estimate the impact of various case mixes on overall costs and resource requirements. Table 21 shows that there is considerable variation in each cost heading between DRGs, and it can be noted that the clinical assessment of more 'complex' DRGs as discussed in the section on regional comparisons is generally borne out in the departmental cost variations by DRG, with these DRGs having higher costs. Also, preliminary analysis indicates that the variation in average cost between DRGs and the average for MDC 02 as a whole, is greater than the corresponding variation in the lengths of stay. And this is true within the surgical DRGs as well as between the surgical and medical DRGs, where a greater variation would be expected due to theatre costs.

These specific cost weights from the study were also compared with those developed by the Health Care Financing Administration for reimbursement of federally funded health care in the USA. Again, there was a greater difference in this study's average costs by DRG. In the study data, the highest cost DRG was nearly six times more expensive than the lowest, whereas the highest HCFA cost weight in MDC 02 was twice that of the lowest. In the MDC 02 surgical DRGs, the difference was less but still marked; study costs varied by a factor of 3 against the HCFA factor of 2. Part of this difference may be explained by differences in practice patterns, particularly with respect to length of stay for certain types of cases. However, the method of costing is likely to be a more significant factor, particularly medical staff costs and the differentiation of nursing staff costs among case types.

The total case cost by DRG was built up from the average cost for each department and it can be disaggregated back. This gives a perspective on each

Table 21: All inpatient costs – average cost by DRG (untrimmed data): study data October 1983 – February 1984

MDC 02: *Eye disorders*	ACTIVITY		DIRECT COSTS (£)						INDIRECT COSTS (£)				
	Number of cases[1]	*Average length of stay (inpatient cases only)*	*Medical staff*[2]	*Nursing staff*	*Med/surg supplies*[3]	*Operative procedures*[4]	*Diagnostic services*[5]	TOTAL DIRECT	*Patient care services*[6]	*General services*[7]	TOTAL INDIRECT	OVERHEAD COSTS[8]	TOTAL COSTS (£)
SURGICAL													
036 Retinal procedures	53	6.6	60.13	209.75	17.22	281.25	10.61	578.96	71.15	95.37	166.52	124.22	869.70
037 Orbital procedures	15	5.2	51.96	140.81	24.10	267.58	15.14	499.59	54.67	76.17	130.84	95.16	725.59
038 Primary iris procedures	16	5.9	51.92	152.86	26.21	143.28	15.73	390.00	65.21	89.02	154.23	114.10	658.33
039 Lens procedures	325	6.0	50.38	150.87	19.68	201.36	15.36	437.65	66.62	90.65	157.27	116.58	711.50
040 Extraocular procedures except orbit, 18 years +	85	4.7	20.18	86.78	16.55	106.45	12.80	242.76	32.25	45.56	77.81	52.77	373.34
041 Extraocular procedures except orbit, 0-17 years	67	2.0	20.46	53.97	2.14	137.17	10.62	224.36	21.28	37.89	59.17	36.67	320.20
042 Intraocular procedures except retina, iris, lens	22	5.9	52.92	183.10	29.42	280.47	14.34	560.25	65.08	88.88	153.96	113.87	828.08
MEDICAL													
043 Hyphaema	5	4.4	36.69	133.40	3.77	—	7.46	181.32	48.83	70.31	119.14	85.45	385.91
044 Acute major infection	23	7.0	52.24	224.61	47.91	—	14.95	339.71	77.69	103.30	180.99	135.94	656.64
045 Neurological disorders	2	4.0	32.85	71.93	10.52	—	46.07	161.37	44.39	65.23	109.62	77.68	348.67
046-48 Other disorders:													
046 18 years + with comorbidity, complications	31	2.9	24.68	66.45	7.88	13.77	14.73	127.51	31.10	49.83	80.93	54.18	262.62
047 18 years + without comorbidity, complications	28	3.0	23.65	54.88	7.73	4.88	14.72	105.86	32.07	50.69	82.76	55.84	244.46
048 0-17 years	5	1.4	15.65	41.85	3.48	6.72	5.26	72.96	15.54	32.24	47.78	27.19	147.93
MDC 02	677	5.2	41.92	132.22	17.59	164.94	14.05	370.72	54.57	75.97	130.54	94.92	596.18

1 includes 627 inpatients and 50 day cases
2 Ward medical staff costs only; other medical staff costs included within respective departments.
3 includes pharmacy, stores, dressings, CSSD
4 includes theatre, laser, minor operations
5 includes x-ray, pathology, ECG, physiotherapy, fluorescein angiography, orthoptics
6 includes medical staff, nursing staff, med/surg supplies
7 includes admitting/medical records and hotel services
8 includes administration and estate management

department's relative share of the overall care cost and permits further analysis of the largest contributions. Departmental costs by DRG also provide a basis for estimating the impact of a particular mix of patients on the resource requirements of specific departments. Total average cost may mask departmental variation in requirements. For example, in Table 21, DRG 044 Acute major infection, has the highest average direct nursing cost although its total cost is below that of other DRGs. Similarly, the specific type of diagnostic services utilised varies among DRGs.

The costs by DRG of selected departments shown in detail in Tables 22 and 23, including ward nursing care and radiology, provide evidence of the variation in costs per case within a DRG. Analysis of this cost variation across patients within a DRG, as well as between DRGs themselves, is a means of assessing the appropriateness and cost-effectiveness of existing patterns of service.

Variation within a DRG as evidenced in the coefficient of variation may be caused by a number of factors. DRGs 036 Retinal procedures, and 040 Other extraocular procedures 18 years +, have consistently high variation across different cost headings. This may be due to the existence of distinct diagnostic sub groups within these two DRGs and is worthy of further examination.

The nursing costs per case were derived from a nursing dependency rating per patient day which largely reflected basic care needs. These ratings included care specific to eye conditions and care for general medical conditions. The existence of secondary medical conditions in this predominantly elderly patient population was thought to be a contributing factor to the variation although, interestingly, the nursing cost distributions within DRGs are comparatively tight. Diagnostic costs per case, as illustrated in the radiology example, are likely to be affected most by clinical policy on pre-operative work-up requirements in relation to the type of anaesthetic used.

The information system developed in this study includes activity, as well as cost data. The type and quantity of workload generated by a mix of DRGs can be translated into the workload requirements for individual departments for planning, budgeting and monitoring purposes. The costing of departmental activity was based as closely as possible on weighted units of resource use per type of activity. The categorisation of departmental activity and the associated resource use weights are useful in themselves as a departmental measure of workload. For example, it is evident from Table 24 that different settings generate a very different 'mix' of workload.

Chest x-rays constitute 81 per cent of the inpatient workload compared with 27 per cent of the examinations in outpatients. Conversely, head area examinations dominate outpatient activity where there is also a greater proportion of higher resource-consuming exam types. The result is an outpatient cost per exam of £5.01 compared with £4.06 for inpatients.

The final output of a case cost is a function of the type and quantity of services received and the cost per item of service. Each factor is the responsibility of different groups. While clinicians are responsible for the control of service orders, individual departments are responsible for the cost per item of service – the so-called 'intermediate' or departmental outputs of the hospital system. Inside a department, the quantity of input resources used and the cost per resource are combined to determine the item cost. By disaggregating the total cost into its component parts, the impact of each

Table 22 Total ward nursing costs-distribution by DRG: study data October 1983 – February 1984

MDC 02 – *Eye disorders*	*Number of cases*	*Average nursing cost per case (£)*	*Standard deviation nursing cost per case (£)*	*Coefficient of variation nursing cost per case (£)*
SURGICAL				
036 Retinal procedures	53	266.64	397.59	1.49
037 Orbital procedures	15	184.56	150.86	0.82
038 Primary iris procedures	16	204.96	80.91	0.39
039 Lens procedures	325	204.10	107.89	0.53
040 Extraocular procedures except orbit 18 years +	85	112.97	256.36	2.27
041 Extraocular procedures except orbit 0–17 years	67	71.05	32.25	0.45
042 Intraocular procedures except retina, iris, lens	22	235.10	130.79	0.56
MEDICAL				
043 Hyphaema	5	172.42	173.36	1.00
044 Acute major disorders	23	286.69	199.83	0.70
045 Neurological disorders	2	107.40	71.32	0.66
046 Other disorders, 18 years + with comorbidity complications	31	91.34	70.84	0.78
047 Other disorders, 18 years + without comorbidity complications	28	80.54	78.50	0.97
048 Other disorders, 0–17 years	5	54.27	18.18	0.33
MDC 02	677	175.89	181.95	1.03

NOTE: Untrimmed data include direct and indirect nursing staff costs

Table 23 Inpatient radiology costs – distribution by DRG: study data October 1983 – February 1984

MDC 02 – *Eye disorders*	*Number of cases*	*Average x-ray cost £*	*Standard deviation x-ray cost £*	*Coefficient of variation x-ray cost £*
SURGICAL				
036 Retinal procedures	53	2.87	4.19	1.46
037 Orbital procedures	15	6.29	7.49	1.19
038 Primary iris procedures	16	2.96	3.85	1.30
039 Lens procedures	325	2.58	2.71	1.05
040 Extraocular procedures except orbit 18 years +	85	1.33	3.00	2.26
041 Extraocular procedures except orbit 0–17 years	67	0.36	2.16	6.00
042 Intraocular procedures except retina, iris, lens	22	5.36	5.25	0.98
MEDICAL				
043 Hyphaema	5	7.04	10.04	1.43
044 Acute major infection	23	4.00	7.89	1.97
045 Neurological disorders	2	14.03	10.22	0.73
046 Other disorders, 18 years + with comorbidity complications	31	2.37	1.88	0.80
047 Other disorders, 18 years + without comorbidity complications	28	1.58	2.27	1.44
048 Other disorders, 0–17 years	5	–	–	–
MDC 02	677	2.44	3.71	1.52

NOTE: Untrimmed data

Table 24 Radiology department – costs for inpatients and outpatients by type of exam: study data October 1983 – February 1984

Type of exam	*Inpatients*		*Outpatients*		*Total*	
	No exams	*Cost* £	*No exams*	*Cost* £	*No exams*	*Cost* £
Chest	356	1296	221	728	577	2024
Head Area (skull, orbits, and so on)	70	403	448	2184	518	2587
Spine	12	51	65	223	77	274
DCG	–	–	34	504	34	504
IVP, cholecyst	–	–	12	179	12	179
All other	1	4	39	274	40	278
	439	£1754	819	£4092	1258	£5846

component can be identified, a determination made as to the appropriateness of its present level and responsibility assigned. There are two obvious applications for this information. The first is to analyse the cost-effectiveness of the present situation and the potential for improving efficiency. The second is to monitor budgets to explain variations between actual and budgeted expenditure.

Conclusions

The value of an output versus input-focused financial information system is as evident now as it was in the early 1970s. Increasing pressure to maximise the use of limited health care resources in the UK and throughout the world has brought about improved measures of resource use in relationship to output. Defining health care outputs in terms of DRGs appears to provide the best classification for controlling variation in resource use. It also provides a vocabulary which is meaningful in describing clinical work.

The effort required to produce such a system is still evident, but perhaps less daunting than 10 years ago. The diagnostic group costing study found the effort to be considerable because of the scope of information required and the number of data systems involved, including activity, resource availability and cost data. Activity data are required for all direct patient care departments and in many instances have first to be collected at the individual patient level and then aggregated. The identification of costs in a format suitable for assignment to activity requires considerable manipulation and adjustment to the present accounting system both during the process of initial coding and subsequently in the way costs are categorised and aggregated.

However, moving towards a complete DRG costing system is becoming more feasible. Various initiatives in the National Health Service will expedite the effort required. There has been a dramatic increase in computerised systems, making comprehensive data collection considerably easier. At the same time there is a wider advocacy of improved information systems. Activity data systems are expanding with the implementation of the Körner[9] recommendations while the Griffiths inquiry[8] has led to management

budgeting requiring the incorporation of meaningful clinical activity measures in the budgetary process. The same initiatives are also leading the way in promoting a proactive management in the NHS rather than a continuation of a largely reactive philosophy. Limited resources, the shift from administration to the broader issues of management, the devolution of responsibility to unit and district level, and greater clinician involvement in resource management, have all led to a need for sensitive output measures as management tools. So perhaps the effort can now be seen to be worthwhile.

References

1 King Edward's Hospital Fund for London. Accounting for health: report on a King's Fund working party on the application of economic principles to health service management (Chairman, Brian Abel-Smith). London, King Edward's Hospital Fund for London, 1973: 13.
2 See 1: 21, 22.
3 See 1: 39.
4 Russell E M. Patient costing study. Scottish Health Services Studies No 31. Edinburgh, Scottish Home and Health Department, 1974.
5 Babson J H. Disease costing. Manchester, Manchester University Press, 1973: 2.
6 Hillman R L and Nix G R. DHSS funded research into specialty costing 1980–1982. London, DHSS, 1983.
7 NHS Financial Information Project. Patient based costing in the hospital. Field trial report. Birmingham, West Midlands RHA, May 1984.
8 Department of Health and Social Security. NHS management inquiry (Leader, Mr Roy Griffiths). London, DHSS, 1983.
9 Department of Health and Social Security. Steering Group on Health Services Information. First report. A report on the collection and use of information about hospital clinical activity in the National Health Service. (Chairman, Mrs Edith Körner) London, HMSO, 1982.

10 PLANNING CLINICAL BUDGETS USING DRGs

Iden Wickings

The ideal tool?

To my mind the human hand comes closest to being the ideal tool. It is always where you left it, which is helpful to the forgetful such as myself. It is almost infinitely adaptable. It can serve as a deadly weapon or be used for the stroke of a lover. Skilled, it can produce glorious music from an inanimate piano or relieve the pain of a suffering patient. Hands can speak for the deaf and allow the blind to read. In combination with other tools, hands can undertake micro-surgery or move mountains.

Turning to the theme of this book, I have to admit that DRGs are not quite such ideal tools, even for the more limited demands of health care managers; nonetheless they are not bad. They do not claim to cover outpatient work and are poor descriptors of long stay care, although these weaknesses are currently being addressed. However, like hands, DRGs can be used with other tools to serve a variety of purposes. Similarly they can be used imaginatively and with skill to form something akin to words. To justify these claims, which some might think exaggerated, I propose to discuss the utility of DRGs as tools for two functions for which they were not originally designed:

as words in a language of health management:
as proxy outputs in budgets for clinical services.

Furthermore I shall be discussing them in the British context of a state funded, nationally controlled health service. We should remember that in such a setting, as in most government controlled services everywhere, there is a considerable emphasis on planning with less attention being paid either to influencing or responding to the choices of consumers in the market. Are DRGs, originally designed for use in the aptly described American health care market, equally useful for the different purposes of those working in a planned system?

DRGs as multiple purpose tools

As will have been gathered from some earlier chapters, health systems across the world differ in many profound ways in their financial systems. There are broad families of financial arrangements, of course such as whether reimbursement is prospective or retrospective, or whether funding is organised in relation to each patient treated or by a global budget for a whole service. I have discussed elsewhere the idea that there may be a 'general theory' of health care budgeting in that these different families have predictable effects, both good and bad, on patterns of health care delivery.[1]

DRGs were first developed to assist in the review of clinical performance, and in chapter 8 Gwyn Bevan discusses how reviews of this nature could be

applied in the NHS. In the US, however, it was not long before DRGs were used to discipline a situation in which health expenditure was widely reported to be virtually out of control[2, 3, 4] and in which most hospital costs were met by retrospective payment for each treated case.

It is worth considering briefly what 'out of control' meant. In the United States there was a multiplicative factor at work: each year more patients were being treated but additionally the average cost of treating similar patients was said to be growing much faster than could be justified on obvious grounds, such as the effects of inflation, new and desirable technology, or demographic changes. I shall call the first factor the 'volume' effect (the numbers of patients treated) and the second the 'unit cost' effect. Both have been discussed in some detail in chapter 6. There can be little doubt that DRGs have proved a powerful tool in controlling unit costs in such a setting[5, 6]. Interestingly they may also have influenced volumes more than was anticipated, by rendering, for example, some treatments less financially viable for the hospitals concerned[7, 8].

Linda Jenkins told us something of the United States' experience in chapter 3. The following paragraphs are more speculative and explore the still largely theoretical potential of DRGs in the UK health care environment which is very different, not only financially but organisationally. There is, for instance, a need to give considerable significance to the British pattern of general practice and its effect upon hospital referrals. But I shall principally be considering the value of DRGs for purposes quite different from those responsible for controlling the cost effects of growths in either volumes or unit costs. I shall be imagining the contribution not they but DRGs could make to the task of planning health care outputs in a health service that is globally funded prospectively. Britain does not have a problem of overall health care cost containment, either NHS or private. Indeed many level headed authorities believe that more, not less, should be spent on the NHS. The contrasts in the organisation of hospital care financing between Britain and the US could therefore not be much greater; but I hope to demonstrate that DRGs (and their associated categorisations AVGs and PDGs which I shall not consider further here) are sufficiently robust as multi-purpose tools to have value not only for planning, but as a language of health care management and, in particular, for budgeting when the intention is to improve service outputs.

Budgeting

For some people, the very word 'budget' implies only cost cutting or at least economising. For others, a budget is simply a finite cash allocation, as if it were a bucket of gold that will suddenly be found to be empty. I wish to use the term more widely to mean planning a resource allocation to achieve a particular purpose. For example, I shall consider how DRGs might fit into the following set of linked definitions which were first used in this form in 1980[9] when I was discussing clinical budgets with a group of cardiac surgeons and cardiologists in Newcastle upon Tyne.

PLAN Objectives attainable within the limits imposed by the resources available.

RESOURCES People, materials, buildings, systems and processes available within a specified period of time.
BUDGET Financial statement of resources to carry out a plan.

The rhetorical question 'How might DRGs fit into a state planned budgetary system?' will be used:

to test the suitability of DRGs as words in a language to be shared by clinicians and system managers;
to assess the potential of DRGs as a description of hospital outputs (regarded as planning objectives); and
to assess their value as descriptors of planned inputs (hospital resources).

Only if they perform well enough in these three tests should DRGs be used routinely in a planned, globally funded health care system using clinical budgets.

Why would clinical budgets provide such powerful tests?

Although I believe the first clinical budgeting experiment in the world took place at the Westminster Hospital, London[10], the longest lasting example is now in Baltimore, US. In Johns Hopkins Hospital clinical budgeting is described as 'decentralisation'[11], which is probably a better description of a managerial system that delegates the responsiblity for operational choices about resource management to those doctors and nurses who are closest to their patients. Delegation cannot, of course, absolve the Director of Johns Hopkins from his own responsibilities, any more than such a system could absolve hospital unit or district general managers in Britain from theirs. Top level managers must still be able to monitor whether the operational decisions that are being made are being undertaken competently and they must be able to impose changes when they are not; but able people given delegated powers usually make sensible choices and demonstrably become more aware of opportunity costs[12]. Decentralisation not only shortens the decision-making chain but frees the general managers to concentrate on their proper concerns with overall strategy and output quality.

If an attempt is made to devolve authority without sufficient clarity about the powers that are to be delegated and how success and failure will be established, the likelihood is that the attempt will prove disappointing to all concerned. With decentralised clinical budgets, questions inevitably arise in clinicians' minds: such as what will be the consequences if unplanned resources are used, perhaps through treating patients who prove unexpectedly expensive, or what will happen if the budgets are not fully expended? In no circumstances can the general manager devolve unfettered freedom to any subordinate, so there is a need for mutual understanding about what can be decided operationally within the delegated authority and what will need reference back. This does not mean that delegation will be restricted to the trivial, but it does mean that there will be limits set on what can be changed. Nonetheless, within these predetermined limits the changes that clinical teams may initiate can still be significant: a recent example in Johns Hopkins Hospital, when I visited there in 1984, had involved a substantial alteration in

the rates of payment for all nurses working in the neurosciences division. To improve patient care, the division had decided that nurses should complete their assigned professional tasks instead of 'clocking off' when their shifts were completed. The particular example that led to the change had been in the operating room, but the principle was to apply throughout the division. Locally this change was described as 'becoming more professional' and basic pay rates were to be slightly enhanced while overtime payments were substantially to be eliminated. This change was being considered for adoption by another division while I was there. In Britain, conditions of service are, of course, negotiated nationally, despite some moves towards limited change[13]; but British experiments in decentralisation have still produced changes in staffing levels, in equipment, in capital facilities and in the planned number of patients to be treated[14].

As these examples show, clinical budgeting can result in changes to planned service outputs – which is the equivalent of altering at least some of the objectives for the clinical service concerned – and in changes to planned resource inputs. It follows that the language used to achieve this level of decentralisation must be adequately descriptive, be well understood and be precise enough to meet the needs of everyone involved.

DRGs as a language of health management

If DRGs are to become a key part of the basic vocabulary in a language of health management (a part that perhaps could even be used internationally) it has to be admitted that there are some differences between American clinical practice and that typical in other countries. These differences need to be built in as adjustments to the DRG data sets being used locally, just as the more significant changes due to local financial and health care organisation systems must be incorporated. For instance, medical fees are currently excluded from the cost components of American DRGs because of the way cost reimbursement is organised. This might be illogical in other locations. If the influence of medical pay on total costs is substantial, and is not a consistent mark-up for all DRGs (as results in Chapter 9 suggest), then different groupings could emerge in different financial settings. This means that although DRGs may form the core words of a widely used vocabulary, there will probably always be the need for some local adaptations. There is the further difficulty that DRGs only cover a part (although a critically expensive part) of all health care.

However, DRGs would seem to meet quite admirably Fowlers'[15] first four rules of vocabulary:

> Prefer the familiar word to the far-fetched.
> Prefer the concrete word to the abstract.
> Prefer the single word to the circumlocution.
> Prefer the short word to the long.

I shall leave on one side their fifth and final rule, which preferred Saxon words to the Romance, as another example of English chauvinism at work. But DRGs are undeniably concrete and short and avoid any dangers of circumlocution. They have advantages, when their familiarity is considered,

in that their terminology when set out in full is medically meaningful and yet the overall number of DRGs does not present too daunting a prospect for learning to be gained by people trained in other disciplines, such as finance or general management. Furthermore, within a single medical or surgical specialty, the number of DRGs that are likely to be discussed when reviewing the service to be provided, the resources to be used and so on, is usually about 20. This simplification of the whole of a specialty's case mix obviously renders it more manageable for planning and budgetary purposes. It also lends itself more readily to statistical modelling and analysis because sufficient numbers are involved. The associated disadvantage (that specificity and explanatory power are lost because of the aggregation of cases into broader DRG categories) is not very serious because the data can always be disaggregated again when necessary and the very method of determining the composition of DRGs was selected to retain meaning for both clinicians and resource managers. An example using general medicine, which is one of the most complicated specialties, is shown in Figure 15. The data are from an English health authority. It will be noted that the display covers all DRGs that make up more than one per cent of the specialty's workload in the year, and yet the number to be considered is reasonable. In our experience, this practical advantage can be contrasted very favourably with any attempt to use the ICD codes, when it becomes apparent all too soon that individual clinicians deal with too few patients in most diagnostic groups for there to be useful discussions based on this categorisation. For the same reason there are also too many of the groups for the data to be manageable.

The foregoing analysis shows that DRGs have most of the attributes of words that describe clinical categorisations in a manner with meaning for clinicians and nurses, planners and epidemiologists, as well as for financial specialists and general managers. The 'words' are familiar, reasonably concrete, short and not prone to ambiguity. For acute, inpatient care, DRGs have now been tested in many countries and, provided that suitable minor amendments have been made to suit local clinical and financial practices, they have consistently been found to be powerful descriptors of medically prescribed treatments, especially bed use, and therefore costs. I believe that they meet satisfactorily the linguistic requirements of the first of my three tests. Next I consider their suitability as descriptors of planned hospital outputs before turning later to their role in explaining resource inputs.

The suitability of DRGs as a tool of service planning – specifying the objectives as outputs

I defined a plan earlier as 'objectives attainable within the limits imposed by the resources available', the objectives being readily distinguishable from vague intentions or woolly ideals. For instance, although the overall intention may be to eliminate cancer from the community, or simply to improve the community's health status, it cannot be said to amount to a fully fledged plan without the necessary resources being made available. The available resources are not just physical assets or money, but systems and processes which include the appropriate skills and knowledge. Therefore, in the sense

Figure 15 Example of data for one year's general medicine in an English health authority

Most common DRGs (over 1% specialty caseload):

			No. Cases	No. Beddays	LOS
14	M	SPECIFIC CEREBROVASCULAR DISORDERS EXCEPT TIA	123	2155	17.52
25	M	SEIZURE & HEADACHE AGE 18-69 W/O C.C.	61	266	3.7
82	M	RESPIRATORY NEOPLASMS	39	559	14.33
88	M	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	64	460	7.19
97	M	BRONCHITIS & ASTHMA AGE 18-69 W/O C.C.	88	404	4.59
121	M	CIRCULATORY DISORDERS WITH AMI & C.V. COMP. DISCH. ALIVE	61	530	8.69
122	M	CIRCULATORY DISORDERS WITH AMI W/O C.V. COMP. DISCH ALIVE	403	2979	7.39
123	M	CIRCULATORY DISORDERS WITH AMI, EXPIRED	91	152	1.67
127	M	HEART FAILURE & SHOCK	104	947	9.11
133	M	ATHEROSCLEROSIS AGE <70 W/O C.C.	34	164	4.82
134	M	HYPERTENSION	50	398	7.96
139	M	CARDIAC ARRHYTHMIA & CONDUCTION DISORDERS AGE <70 W/O C.C.	50	205	4.1
140	M	ANGINA PECTORIS	76	303	3.99
142	M	SYNCOPE & COLLAPSE <70 W/O C.C.	37	122	3.3
175	M	G.I. HAEMORRHAGE AGE <70 W/O C.C.	37	269	7.27
183	M	OESOPHAGITIS, GASTROENT. & MISC. DIGEST DISORDER AGE 18-69 W/O C.C.	48	173	3.6
202	M	CIRRHOSIS & ALCOHOLIC HEPATITIS	32	266	8.31
294	M	DIABETES AGE >36	56	561	10.02
299	M	INBORN ERRORS OR METABOLISM	48	13	0.27
395	M	RED BLOOD CELL DISORDERS AGE >= 18	31	281	9.06
449	M	TOXIC EFFECTS OF DRUGS AGE >= 70 AND/OR C.C.	48	160	3.33
450	M	TOXIC EFFECTS OF DRUGS AGE 18-69 W/O C.C.	451	433	0.96
451	M	TOXIC EFFECTS OF DRUGS AGE 0-17	87	111	1.28
		TOTALS	2119	11871	
			71%	59%	

that I use the term, a plan is a very practical thing. Those agreeing to carry out the plan (the agents through which it will be achieved) are entering into a clear commitment and have to be equally clear that, provided the resources specified are indeed made available as stipulated, they should be capable of delivering the planned outputs. Those authorising the resources, financial or otherwise, must be convinced that the attainable objectives are worth the committed resources since in a lot of cases, these resources could be applied equally well to the attainment of other, competing objectives. To achieve a clear commitment of both parties we have to ask whether specifying the number and type of DRGs to be delivered, either set out as a broad range or more narrowly, is a sufficient statement of the objectives for a clinical plan?

I think that the answer is only broadly satisfactory. Professor Alan Williams, participating in a seminar on output measurement in health care[16], usefully distinguished inputs and throughputs from outputs. In his terms, inputs are the resources; throughputs refer to such activities as the number of patients treated; outputs are achievements as measured by health or longevity improvements for patients. He combines throughputs and outputs in a combined measure, the Quality Adjusted Life Year (QALY). However, ignoring for the present the complications of QALYs but using these valuable distinctions, it can be said that DRGs allow the computation and planning of throughputs to become more sophisticated and precise. Although DRGs have something useful to contribute to discussions about the objectives as throughputs of a clinical service, they say nothing precise about outputs in terms of health improvements as defined by Williams. But, in reality, how severe is this limitation?

Each of us might answer this question differently, depending on our judgment of the usefulness of most medical treatments. Today it seems almost to be conventional wisdom to cast doubts on the scientific objectivity and the reliability, let alone the skills, of the medical profession. Particularly since Illich claimed that the medical establishment posed a danger to our health[17] there has been a mounting tide of criticism based on widely publicised assertions that this or that treatment sold by the practitioners of alternative medicine offers benefits denied to those being treated by 'traditional' methods. No doubt some of these assertions are justified (although anything that I would regard as adequate scientific proof is in lamentably short supply) but it must surely be a nonsense to regard the overwhelming majority of traditional treatments as contraindicated just because of the vehemence of the assertions? There is good evidence, it is true, to support allegations that some hospital treatments could and should be reduced. Health maintenance organisations have sometimes resulted in major reductions in hospital bed use[18]. Studies in Canada have reported a substantial overuse of elective surgery[19]. Hospital referrals are 25 times more frequent from some general practitioners in Britain than from others[20], yet the average rate of referral in the Netherlands greatly exceeds that in the UK. There is indeed abundant evidence of both international and inter-physician variation in clinical practices[21]. But there is not yet any widely agreed template of an acceptable range of treatments. Most criticisms rely on the somewhat unproven view that the outliers in the overall distribution are undoubtedly at fault. This is probably true, but not certainly, and it may be that the apparent offenders at

the high end are still offering marginal benefits to their patients. I am in no sense an unwavering supporter of the medical profession and have often criticised the unsatisfactorily wide range of what is held to be 'acceptable clinical practice[22, 23, 24], but those who are not prepared to regard evidence of throughput as having, more often than not, some reasonable correlation with output (in Williams' terms, health improvements gained) surely need to justify their claims in each case.

I hope I have demonstrated that DRGs have the potential to serve as at least part of the statement of planned outputs for a clinical service. But how valuable would this be in practice?

An example using DRGs as proxy outputs

Let me give a practical example, based on experience, of the valuable aid that DRGs can provide in service planning. Some years ago a British health district was failing to meet the needs of its local community. The problem was a large and long established waiting list for patients needing hip prostheses. It was decided to appoint an additional orthopaedic surgeon in order to reduce the waiting list. When the interviews were held, however, the outstanding applicant by a considerable margin was a surgeon with a special interest in working on hands. He expected to undertake the full range of surgery for his specialty but also to spend a particularly high proportion of his operating hours on his special interest. Indeed he would hope, in time, to build up a pattern of referrals for hand surgery to his department – not just local patients but referrals from far and wide. The interviewing panel agreed that there would be benefits to the orthopaedic department, and to the hospital, from building up such a reputation: for instance better junior staff could be attracted and the cases might be valuable for teaching and research which were other responsibilities of the hospital. In surrounding areas there was also at that time no centre specialising in hand surgery. But the local patients needing hip operations might not be grateful for these wider responsibilities being given priority over their own needs.

What was required to make progress at this stage was a language to express both the local and the wider communities' needs in throughputs in meaningful categories. If the present and planned future throughputs had been expressed in numbers of DRGs then either the applicant could have been asked to commit himself, if necessary contractually, to a specific number of hip operations or, if the interviewing committee simply wished to appoint the best applicant, the other consultants' commitments could have been renegotiated to be complementary with the new surgeon's, so that when aggregated they would still be able to meet local needs.

This shows how DRGs can be especially useful in a state-funded, publicly answerable service whose managers are expected to plan and then deliver health care programmes responsive to all legitimate needs. Of course, many other examples could be given. DRGs will sometimes explain why one hospital's service is considerably more expensive than another apparently similar hospital which is treating the same number of patients. Differences in case mix between hospitals treating similar communities may also reveal omissions from the service being offered. The public and its representatives

have a proper interest in efficiency and coverage but the comparisons must be genuine. DRGs are not quite sufficient to meet Williams' definition of outputs but they, or something remarkably like them, are a necessary component of stated objectives for acute hospital services. Such statements are particularly important if clinical budgeting, or decentralisation, is to be practised.

Local clinical budgets using DRGs

My third test concerned the adequacy with which DRGs describe hospital resource use. If DRGs are going to be used as part of the statement of planned outputs for a given clinical service will, concurrently, the resources likely to be needed to treat that case mix also be specified?

It is interesting to note that the development of clinical budgets at local level, perhaps at the level of the individual consultant, brings out ways in which DRGs may contribute quite outside the reason for their original development. It is another example of their having the potential to serve as multi-purpose tools. The point at issue is that the study of resource use (in the form of planned costs for a planned case mix) raises surprising matters concerned with quality of care.

Of course, costs are always a significant consideration, but at the level of the individual consultant or the clinical team, budgets constructed by aggregating planned costs that are themselves derived from national or regional averages may well be inappropriate. No single figure will ever provide an infallible prospective guide to the resources that will be consumed in the treatment of any one patient, even if allowance is made for the sufferer's diagnosis and age. This still remains true if allowance is made for the relative severity of the patient's condition. It is not only patients and the staff treating them that behave unpredictably; the relative cost of the local resources used will behave in a surprisingly elastic manner as well. Some of the reasons are obvious: nurses on a ward at its busiest period will be able to allocate less time to the average patient than during quieter periods. Some reasons are more complex: for instance, one has to decide whether the cost of the physical accommodation used by a patient should reflect the daily fluctuations in the cost of borrowing capital funds from the money markets. After all, a sizeable capital asset is being employed. It needs maintenance and an allowance for depreciation must be made to permit its ultimate replacement. Both of these financial difficulties are usually dealt with by averaging, but inaccuracies or clumsy methods of cost computation raise great difficulties when the number of patients being treated is very small. This will usually be the case when preparing the budget (a financial statement of the resources to carry out a plan, as defined earlier) for an individual consultant.

To understand why I am coupling the quality and financial issues together, here is a a simple example. In Britain, the number of patients treated annually by each consultant ranges from a mere handful to several thousands. There are many factors at work, including the consultant's specialty, personal reputation, age, location, method of practice, individual availability, special interests and energy, not to mention the resources allocated to the consultant and the skill and scale of the available supporting staff. The number of

patients may be significant in unexpected ways that have nothing to do with costs. To take a surgical example: a consultant who undertakes a particular operation infrequently may well not only complete it more slowly (thus costing more) but even, in some specialties, operate less safely[25]. In general medicine as well as surgery the same multiplicity of factors can be at work. Table 25 illustrates the average length of stay for patients in one DRG admitted to a district. The consultants that treat the condition most frequently are those who keep their patients in hospital for the shortest time. Nearly 80 per cent of the patients were admitted by only three of the 11 consultants but they used only two-thirds of the total bed days. The unit costs of their colleagues, who kept their patients in much longer, will be much greater. We have found the same picture to be broadly true in many other specialties.

Table 25 Consultant activity and average stay for the elderly or complicated cases with peripheral vascular disorders (DRG 130)

Consultant	Cases seen	Average length of stay
A	30	7.8
B	22	7.9
C	21	8.2
D	5	6.8
E	5	15.4
F	4	17.8
G	2	14.0
H	2	3.0
I	1	14.0
J	1	16.0
K	1	29.0
All consultants	94	9.1

Although these varied patterns of treatment raise issues about the comparative costs involved, they also raise questions about the acceptable level of clinical competence and the quality of care provided in clinical terms. Both may often be related. The Scottish surgeon Harper reported many years ago that the patients in his surgical firm who were particularly expensive to treat were those in whom complications had developed:

> The avoidance of inpatient morbidity may be a crucial factor in reducing disease cost and this concept has important implications in regard to the training and supervision of junior staff.....The study of comparative disease costs indicates a relationship between cost and length of stay within disease groups as well as all surgical patients considered. While reduction of the length of stay is more likely to lead to increase in turnover rather than reduce overall service cost, it may indicate improved cost effectiveness. How can the surgeon achieve this? Delayed discharge usually reflects operative morbidity and underlines the need for good assessment and pre-operative preparation, a flawless technique and vigilant post-operative care[26].

Our own research has discovered that patients within a particular DRG who are treated by a 'surprising' consultant (for example in a specialty that does not usually see that type of case) regularly stay in hospital longer. Presumably this means that these patients take longer than necessary to be healed, that is they may have suffered needlessly, thus raising qualitative and even ethical issues as well.

This discussion explains why it would be perhaps unwise to prepare a budget for an individual consultant by aggregating national average DRG costs for the cases that he or she is expected to treat. Averages incorporating poor standards are undesirable models and if pressure is to be imposed to change clinical practices then those applying it need to be very sure-footed and sensitively aware of all relevant considerations. There are simply too many key variables concerning the movement of costs, the numbers of patients treated in different DRGs, idiosyncrasies in clinical practices and the quality of care. The consequences are that either clinical budgets have to be framed at a level where the patient numbers will be large enough for crude, aggregated data to be acceptable, or budgets must be negotiated very carefully with the individual consultants so that the complex details described can be explored sensitively.

Provided these reservations are borne in mind there would seem to be important advantages to be gained from using DRGs in discussions about planned resource use for clinical services. Obviously there would be times when the level of allocated resources would be altered, upwards or downwards, as a result of taking case mix into account when preparing clinical budgets at local level. Of more importance in the long run, however, it can be seen that such discussions could reasonably be expected to play a part in improving standards of care. This could be the result of exposing apparently poor practices in terms of unit costs which in turn should lead either to self-evaluation and improvement by the physician or surgeon concerned, or to policy decisions such as a particular hospital ensuring that all patients in a given DRG are referred to a particular consultant or group of consultants.

Conclusion

I have been considering whether DRGs have a role to play as tools in the planning of clinical budgets in a state funded, nationally controlled health service. To decide this question affirmatively I suggested that DRGs would need to be suitable as 'words' to be shared by clinicians and system managers, be useful descriptors of hospital outputs (regarded as planning objectives), and also be valuable as descriptors of planned inputs (hospital resources).

It may be surprising that my reservations now centre mainly on the third requirement; surprising because it is here that one would have expected DRGs to be strongest. However, I have no doubts that DRGs could prove very powerful tools when clinicians and managers need to discuss service plans and agree both the desirable route forward and what each can expect of the other; that level of agreement is, after all, what clinical budgeting is mainly designed to achieve. My reservations concern the adequacy of DRGs as predictors of resource use at the organisationally small levels where clinical budgeting takes place. Here, an insensitive use of 'average costs' could

actually be dangerous as well as unfair both to the patients and those treating them. However, if DRGs could be employed in pursuance of an understanding and informed discussion about the resources needed for a particular case mix, and the manner in which particular DRGs are to be clinically managed, then qualitative improvements in standards of care could be achieved concurrently with the better management of scarce resources. It would require considerable ability to be developed on the part of the non-clinical managers responsible for discussing these issues with the clinicians, but surely the patients have the right to expect nothing less?

References

1 Wickings H I. Is there a general theory for health care budgeting? Effective Health Care, 2, 2, 1984: 51–55.
2 Committee on Interstate and Foreign Commerce. Issues in the National Financing of Health Care. 1976.
3 Moores B. The buck $tart$ here. Health and Social Service Journal, 17 June 1982: 746–747.
4 Somers A R. Moderating the rise in health-care costs – a pragmatic beginning. New England Journal of Medicine, 307, 1982: 944–947.
5 Gardner B. The impact of DRGs on surgical practice. Surgery, Gynecology and Obstetrics, 159, 1984: 75–76.
6 Dean Witter Reynolds Inc. Hospital supply industry review and outlook. The hospital suppliers in the new health-care environment. 1984.
7 Kinzer D and Warner M. The effect of case-mix adjustment on a admission-based reimbursement. Health Services Review, 18, 2, 1983: 209–232.
8 Spivey B E. The relation between hospital management and medical staff under a prospective-payment system. New England Journal of Medicine, 310, 15, 1984: 984–986.
9 Wickings H I. Clinical budgets. In: The proceedings of the British Heart Foundation symposium, 1980: 28–32.
10 Wickings H I. Putting it together: the patient, the purse, and the practice. The Lancet, 1, 1977: 239–240.
11 Solomon S. How one hospital broke its inflation fever. Fortune, 1979: 148–154.
12 Wickings H I. Opportunity costs and budgets for treatment services. In: France G H (ed). Financial budgets for hospital consultants. Edinburgh, Scottish Health Service Centre, 1981: 4–18.
13 King Edward's Hospital Fund for London and National Association of Health Authorities. NHS pay: a time for change. (Chairman of working party, Lady McCarthy) KEHF/NAHA, 1985.
14 Wickings H I, Coles J M, Flux R B and Howard L. Review of clinical budgeting and costing experiments. British Medical Journal, 286, 1983: 575–578.
15 Fowler H W and Fowler F G. The King's English. Oxford, Oxford University Press, 1958 (3rd edition).
16 Williams A. What is 'output' and why do we need to measure it? In: Output measurement for health services. CASPE Research, 1983.
17 Illich I. Limits to Medicine. Penguin, 1977.
18 Cohen D, Porter D, Dawson N, Lee J, Breslau N, Breslau D, Goldberg H, Hershey C and McLaren C. The cost implications of academic group practice – a randomized controlled trial. New England Journal of Medicine, 314, 24, 1986: 1553–1557.
19 Dyck F J, Murphy F A, et al. Effect of surveillance on the number of hysterectomies in the province of Saskatchewan. New England Journal of Medicine, 296, 1977: 1326–1328.

20 Teeling-Smith G (ed). Health, education and general practice. London, Office of Health Economics, 1986.
21 World Health Organization. Report of the working group on the impact of medical decision-making on hospital costs. Copenhagen, WHO Regional Office for Europe, 1986.
22 Wickings H I. Consultants face the figures. Health and Social Service Journal, 8 December 1983: 1466–1468.
23 See 1.
24 Wickings H I and Coles J M. The ethical imperative of clinical budgeting. Nuffield/ York Portfolio No 10, 1985.
25 English T. The UK cardiac surgical register. In: Output measurement for health services. CASPE Research, 1983.
26 Harper D. Comparative disease costing in surgical patients. In: Hospital costing and the clinician. London, DHSS, 1977: 15–22.

Future opportunities

11 THE IMPORTANCE OF CASE MIX IN HEALTH SERVICE MANAGEMENT

The editors

The current rash of exhortations to improve effectiveness and efficiency in health care are fruitless if it is unclear how these improvements can be made. Better information is necessary in order to manage the expensive resources used for the delivery of health care, and this includes a robust and workable measure of case mix.

This chapter looks at the current state of development of DRGs as the most generally applicable measure of inpatient case mix and considers their further development in the short term. Their initial use by health insurers to reimburse hospitals at fixed prices has not deterred many health care researchers in Europe, Canada and Australia from seeing the value of DRGs in a variety of applications[1]. Over the next 10 years, today's DRG definitions will alter to accommodate new treatments and to reflect the introduction of new technologies. It is also to be expected that case mix groupings will be refined as experience is gained from large data sets and across national boundaries.

DRGs are still relatively new, and in the UK and Europe are only at a developmental stage. However it is clear that they already provide a useful framework in which to address many of the more significant issues now facing health care managers, as discussed in some detail in chapter 7. It also seems certain that managers in 10 years' time will still require case mix measures to assist their decision making. The opportunities for effective case mix management are examined in the second half of this chapter.

DRGs now and in the future

More than a reimbursement system

As Professor Fetter has stated on numerous occasions, and again in his introduction to this book, the motivation for defining case types did not stem from any consideration of their use in hospital reimbursement: rather they were seen as a tool for planning and utilisation review. However the nature of DRGs as groups of patients using similar resources made them a natural candidate for any case-based hospital payment scheme.

It is important to realise that DRGs are not simply a short-lived means of reimbursing hospitals in the US. However long the present federal reimbursement scheme lasts, case mix classification has wider potential in a variety of different health care systems, and for a variety of purposes. It could be argued that DRGs will indeed prove more useful to a health care system such as the NHS than that found in the US. A distinction can be made between a form of national social insurance which places greater emphasis on equity of access, as in the NHS, and the more market-orientated approach in the US where

freedom of choice and consumer sovereignty are of greater importance. In order for the first system to achieve its goals and meet society's expectations, a more sophisticated bureaucracy is required and a greater depth of knowledge of issues – such as need within the population, relative performance and the effectiveness with which resources are used. Questions of planning and resource allocation become critical and the potential of DRGs as a routine description of the way in which available resources are deployed becomes that much greater.

It is clear that DRGs are proving of interest in many countries and it is worth considering their appeal in different health care settings.

What makes DRGs attractive?

The advocacy of DRGs often revolves around two points. The first is that DRGs provide a better description of the end product of a hospital's activities than is currently available on a routine basis, and can therefore provide a much better indication of how money is being spent in providing care. Questions can then be asked about what *is* being done in any given hospital, as well as what *can* be done and, more importantly, what *should* be done.

The second point is more straightforward. DRGs are easy to use, easy to understand and provide information which is relatively inexpensive to obtain, since they do not necessarily demand a massive capital investment. Their descriptions can be interpreted by the medically untrained and therefore offer a vocabulary for dialogue between clinicians and laymen. Of course, this may be thought to be a two-edged sword, but it should ensure that resourcing discussions and others take place using a common information base, if not from a common viewpoint. DRGs in many cases can be used to exploit existing information systems and provide new forms of analysis.

The first point begs a number of questions which are worth exploring in more detail. Specifically, what do we mean by 'better description' and do we need it anyway?

DRGs are better in that they can provide routinely an assessment of the financial implications of treating different types of patient. The description offered by DRGs can be used both as a vocabulary for determining what range and volume of care the hospital should provide and for monitoring performance against desired organisational goals.

If the remit of a health authority is to maximise health within a catchment population, then a description of hospital activities based on DRGs (or other measures of case type) is to be preferred to one based on volume of cases within a specialty or hospital. There are clearly distinct differences in the expected costs of treating different types of patient within a single specialty. In most settings, resources committed in one area are necessarily denied to another, so an economic assessment of the consequences of differences in case mix is required to achieve the greatest possible benefit from the resources available.

More specific descriptions of patients could well be advocated, even beyond the DRG level. However, the value to managers of such a process of increasingly fine differentiations between case types must eventually be traded off against the manageability of the resulting scheme and the costs of

information gathering. Detailed classifications of diagnoses (ICD) have been available for many years, but have been rarely used by management other than on an *adhoc* basis because they have proved too unwieldy.

With the changes now taking place in the management of health care throughout the world, a greater emphasis is being placed on improving management information systems in order to assist decision making. There is also a growing awareness that this new information should facilitate improvements in health care rather than some intermediate aim of the institution. Since appropriate care depends on the condition of the patient it is hardly surprising that case mix measures are coming to the fore, but like all information systems their potential will only be fulfilled if management has the ability and the will to use them effectively.

Problems and potential

As with any novel development which may threaten the status quo, DRGs have received their fair share of criticism. Much has come from the US, and here it is important to separate criticisms of the DRG classification from those levelled at the prospective reimbursement system.

DRGs are not the ultimate case mix classification and they contain some flaws which have been publicised[2,3,4,5]. Many of these criticisms can be considered as second order problems. They do not necessarily mean that DRGs (or a similar form of case mix classification) should not be used, but rather point to ways in which they can be improved, or where care should be taken in their interpretation.

Most criticisms do not compare the value of case mix measures with managing health care without them. For example, US advocates of including severity measures within DRGs do not necessarily feel that Medicare should revert to fee for service funding of hospitals, but rather that the current system introduces avoidable inequities between hospitals.

Similarly the assumption, inherent within the DRGs, of equal quality of care or patient outcomes in a particular group, can be viewed as a weakness. On the positive side, however, the DRG scheme can provide a framework for examining variations in quality and relating them directly to the resources used. In this way comparisons between different balances in the cost/quality equation can be made.

One further criticism concerns the fact that surgical DRGs are defined partially by the chosen treatment. For example, patients who undergo a tonsillectomy will have had tonsillitis to varying degrees, but it has to be assumed that the tonsillectomy was necessary. Once again, though, the DRGs themselves can provide a basis for studying the relative rates per thousand population with which various interventions have been performed. The rate at which certain surgical procedures are undertaken has an important discretionary element and what represents 'unnecessary' interventions is clearly a matter of clinical and managerial judgement; nonetheless, case mix figures can provide background information.

DRGs can be seen as providing operational definitions of a hospital's objectives for inpatient activity; namely, the volume of care to be given across case types and descriptions of how that care is to be provided. With this in

mind their potential applications are numerous. They can be used at all levels in a hospital (planning bed use on a ward or budgeting with clinicians) and in national decision-making on issues such as the expected cost of demographic changes and so on. It is expecting rather too much from a single system that it should replace the variety of classifications currently used for distinct and often very specific purposes. Diagnostic and clinical classifications will still be needed for research and epidemiological purposes and measures of nurse dependency will be used for nurse staffing and manpower planning. Yet it is the balance between the number of case types and the precision of their definitions that makes the DRG classification best suited to a broad spectrum of managerial purposes.

The likelihood that reasonable cost estimates for DRGs can be made available on a routine basis opens up yet more areas for exploiting DRGs – a wider economic analysis of patient care, for example. Even if cost effectiveness/cost benefit analyses can only be performed relatively crudely using DRGs, the ease with which they can be carried out coupled with vociferous demands for the more routine use of such economic appraisals, must make them attractive [6,7,8,9].

Beyond the national level, standardisation of case type by DRG make international comparisons possible. While it may be necessary to make some adjustment to the definitions to reflect national differences in the way care is provided, the majority of groups are broadly comparable and should give a valuable insight into the costs and benefits of care as practised in different countries. Definitional adjustments might also provide opportunities to discuss these alternative care patterns.

Issues for general managers

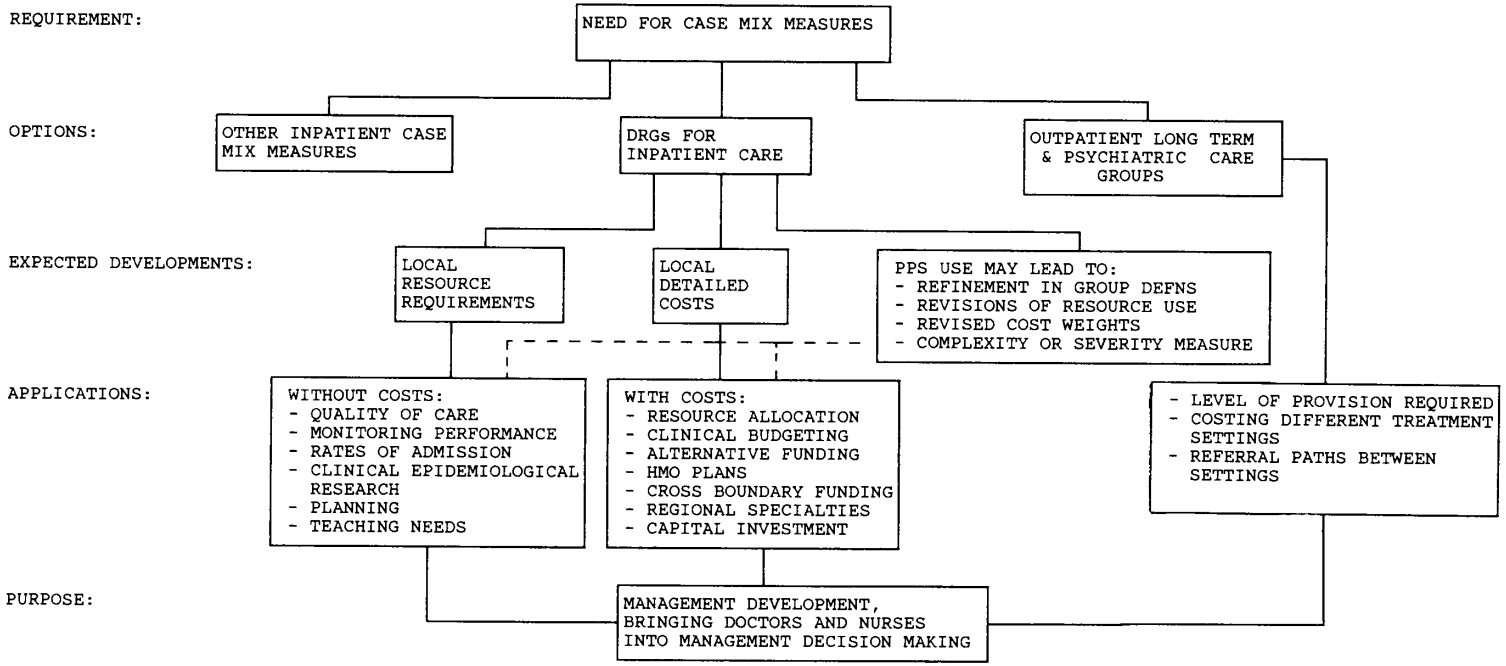
We have already touched on our belief that case mix measures will become indispensable for effective general management within the next ten years, and that DRGs appear to provide the best classification of inpatient care at the present time. This section looks at opportunities for their use in the short to medium term. Whether their potential is fulfilled depends on the ability of managers and on the development of skills to use them effectively.

Figure 16 sets out some possibilities for the application of diagnosis related groups, both as case mix descriptors and, with costed information attributed to case types. Additionally it shows the need for consideration to be given to describing case types in long-term and psychiatric care, as well as outpatients. It also reflects the likely refinements that will be made to the DRG definitions as required by statutory federal regulations in the US, in order to keep them up to date with developments in clinical practice.

Case mix descriptors

It has already been noted that controlling for case types is relevant to a variety of studies. While its relevance in financial planning and budgeting is obvious, it is likely that the next few years will see the issues of quality of care, patient satisfaction and outcome measures come to the fore. Although DRGs are not the only descriptors of case type which can be used as to address these issues

Figure 16 Developments and applications of case mix measures



they do have some significant advantages. They use routinely available information (and hence are less costly to produce); they provide a comprehensive yet manageable number of types; and they relate to other systems, including financial ones.

DRGs are expected to make a useful contribution to establishing standards of care as minimum acceptable treatment profiles, or as ideals against which to compare actual performance. It should be possible to establish expected health indicators at admission and discharge for specific case types. Although DRGs are iso-resource at an aggregate level, it remains to be seen how homogeneous they are in their use of individual items of service. A profile of appropriate care for a case type might include an expected minimum or maximum level of individual services, for example nursing care, diagnostic tests, physiotherapy, and so on. The standards adopted may vary by hospital: for instance a longer stay or more extensive testing might be considered acceptable for a teaching hospital.

Developing standards of care will be a slow and controversial process, but they might be welcomed by nurses and doctors as a means of resisting the more draconian cost-cutting exercises. Standards might also be examined in the light of information concerning patient outcomes (to establish links between resources, process of care and resulting outcomes) and performance at hospital and consultant level. They can also be used in planning at hospital and district level.

Applications of costed DRG information

The biggest potential for DRGs lies in their ability to show the costs of treating different sorts of patient at a level that is more precise than specialty averages. Although there is some consistency in average specialty costs between hospitals, averages clearly mask considerable variation between the cases seen within a specialty.

Figure 16 shows some of the areas where a costed DRG scheme would provide a useful framework in which to work. For internal hospital budgeting, DRGs are seen as a useful tool for both doctors and administrators. They would also help to involve clinicians in decisions on sharing resources among patients, while value for money studies would benefit from controlling for case type, since it would permit a selective review of the resources used by specific patient groups.

Further up the NHS administrative hierarchy at district or regional level there are other obvious applications. Flows of patients between districts could be funded on a case-specific basis rather than average cost, although a choice of local or national standards would have to be made. Similar calculations could estimate the costs of regional specialties and determine the level of provision required at a particular location.

Case mix outside acute hospitals

The use of case mix descriptions will expand – indeed it is already doing so – into settings other than the acute inpatient sector. To keep up with changes in the style of provision – more outpatient treatment, special facilities for the chronically ill, and so on – health service planners and managers will need to

describe patient mix in different settings, identify differences in cost and understand the paths of referral between them. Several schemes currently exist in the US for both outpatients and long-stay cases. Groups for describing psychiatric cases are less well-established and have been developed more subjectively.

The consensus of opinion about both psychiatric cases and long-term care is that the resource demands of patients are *not* primarily diagnosis related. Measures such as the ability to dress, feed and walk are generally accepted as the best discriminating variables for chronically ill patients [10,11], largely because nursing dependency overshadows any differences in the amount of other resources used by patients, such as time with a doctor, tests, investigations or therapy.

Unlike long term care classifications, where there is considerable agreement, the classification of psychiatric patient types has not evolved smoothly [12,13]. Different institutions have developed their own classifications according to their special interests or types of case seen. Although it has been usual to group cases by diagnosis, it has generally been found that there is enormous variability in the resources used, consequent on both the practice pattern of the psychiatrist and the unpredictable remission of most psychiatric disorders.

In the last few years attempts to classify outpatient work have tried to take the best ideas from the acute DRGs approach and other methods of patient clustering. The most advanced scheme is probably that developed at Yale, the ambulatory visit groups (AVGs). These are split between people with known disorders, who may or may not also be treated as inpatients, and those with a range of other reasons for attendance at an outpatient clinic. The first part of the classification mirrors DRGs dealing with major disease categories within which individual case types are identified. The second part uses the reason for visit, for example, innoculations or medical examinations, as a useful predictor of costs and resource consequences.

The extent to which outpatient and other case mix groups are generally applicable across countries will obviously depend on the way health care is organised. It is less likely that these extensions to case mix measures will have as wide an appeal as DRGs. The types of patient receiving hospital inpatient care vary less from one health service to another than care at the primary/outpatient and acute/chronic interfaces, which depend considerably on local policies and provisions.

Developments in DRG definitions

If DRGs are to remain a satisfactory means of paying for inpatient care in the US a number of difficulties will need to be resolved[14]. These lie in the current allocation of cost from functional cost centres to patient types, in the incentives created by attaching particular price tags to patients, and in the need for a constantly evolving and flexible classification scheme.

The hospital funding scheme now used on a national scale in the US excludes some major sources of expenditure from the model that allocates costs to patient types, including doctors' salaries, additional costs associated with teaching hospitals and capital expenditure. These factors can exert a large influence on treatment costs it is argued that they should be brought into

the funding mechanism rather than accounted for independently.

Today's system of US hospital reimbursement involves a delicate balance of financial incentives aimed at cutting overall costs with the least harm. Existing incentives are not necessarily in the interests of patients and can still allow purely profit-orientated health providers to do well. They include incentives to operate (surgical DRGs have higher costs attached to them), or to specialise in profitable lines. This can lead to unnecessary surgical interventions or refusal to admit certain types of case, resulting in uneven provision of services. Considerable thought is needed to create the right incentives. The hastily devised Medicare reimbursement of hospitals provides a good illustration of the problems still to be resolved.

To create incentives that are in the interests of patient, provider and third-party payer will not be easy, especially when the debate involves payers and providers rather than the grateful but powerless patient. Apart from relying on a sense of fairness in hospitals to make provision for all types of cases, and to allow equal access, there are other ways of controlling market forces. The most obvious is to alter the cost weights to manipulate the market, or at least to consider the possibility of doing so. This would make excessive specialisation in profitable lines too risky. Such changes would need a greater understanding of how to attribute costs to patient types.

The scheme must also adapt and evolve in step with changes in hospital methods and costs. Regular revisions will be necessary to reflect changes in treatment practice, the use of new technologies, and inflation – in both wages and the hospital market basket. Changes in costs are in fact assessed by a rate setting commission in an annual review of Medicare's cost weights.

Areas of weakness in the classification process must also be strengthened. In particular the ability of DRGs to describe psychiatric patients is acknowledged to be weak. Across the whole range of case types, cost weights were principally calculated for a sample of older patients and are inappropriate for a more general population or for specialised hospitals.

A further weakness – that of identifying severely ill or complex cases – is difficult to handle when the definition of severity and its relationship to cost remains largely unknown. With DRGs incorporated in a funding process it seems probable that ways will evolve for handling these categories of patients. Possibly it will take the form of identifying a small number of DRGs which are weak because of the heterogeneity of cases falling in them. By subdividing or refining the DRGs as required it would still be easy to roll-up the groups to the original 467, or even to the major diagnostic categories (MDCs).

Concluding remarks

The next ten years will see considerable advances in the measurement and management of case mix. Much pioneering work has established DRGs as a portable, robust and easy to use classification of acute hospital inpatient case types that is generally acknowledged to be the best available.

As projects start to use these new tools in countries around Europe, and in Australia and Canada, the flexibility of DRGs to tackle a range of issues will be demonstrated. Whether with costs attached, as indicators of resource consumption, or as a controlling variable in other studies, they will be seen as

essential tools which can be used at any level of a health service.

It remains to be seen how managers will measure up to the challenging opportunities offered by this additional information.

References

1 DRG Newsletter, June 1986. London, CASPE Research.

2 Neuhauser D and Pine R. DRGs and elective surgery: What's best for the provider? What's best for the patient? Medical Care, 23, 2, 1985: 183–187.

3 Horn S D, Sharkey P D and Bertram D A. Measuring severity of illness: homogeneous case mix groups. Medical Care, 21, 1, 1983: 14–30.

4 Berki S E. The design of case-based hospital payment systems. Medical Care, 21, 1, 1983: 1–13.

5 Birch S and Maynard A. Performance indicators and performance assessment in the UK national health service: implications for management and planning. International Journal of Health Planning and Management, 1, 1986: 143–156.

6 Pineault R. Randomized clinical trial of one-day surgery: patient satisfaction, clinical outcomes, and costs. Medical Care, 23, 2, 1985: 171–182.

7 Jones K R. Predicting hospital charge and stay variation – the role of patient teaching status, controlling for diagnosis-related group, demographic characteristics, and severity of illness. Medical care, 23, 3, 1985: 220–235.

8 Frick A P, Martin S G and Schwartz M. Case-mix and cost differences between teaching and nonteaching hospitals. Medical Care, 23, 4, 1985: 283–295.

9 Garber AM, Fuchs V R and Silverman J F. Case-mix, costs, and outcomes: differences between faculty and community services in a university hospital. New England Journal of Medicine, 310, 19, 1984: 1231–1237.

10 Health Systems Management Group, Yale University. Resource utilisation groups: validation and refinement of a case-mix system for long term care reimbursement. Final report, April 1986.

11 Fries B E and Cooney L M. A patient classification system for long-term care. Yale University School of Medicine. April 1983.

12 Taube C, Lee E S and Forthofer R N. Diagnosis-related groups for mental disorders, alcoholism and drug abuse: evaluation and alternatives. Hospital and Community Psychiatry, 35, 5, 1984: 452–455.

13 Taube C, Lee E S and Forthofer R N. DRGs in psychiatry. Medical Care, 22, 7, 1984: 597–610.

14 Pettengill J and Vertrees J. Reliability and validity in hospital case-mix measurement. Health Care Financing Review, 4, 1982: 101–128.

Appendix I DRG TITLES

The Medicare DRGs, 1985

MDC/DRG	S or M[a]	Abbreviated Title	Relative Cost Weight
MDC 1: Diseases and Disorders of the Nervous System			
DRG			
1	S	Craniotomy age >17 except for trauma	3.3199
2	S	Craniotomy for trauma age >17	3.2488
3*	S	Craniotomy age <18	2.9183
4	S	Spinal procedures	2.2219
5	S	Extracranial vascular procedures	1.6606
6	S	Carpal tunnel release	.3952
7	S	Periph + cranial nerve + other nerv syst proc age >69 and/or C.C.	1.0172
8	S	Periph + cranial nerve + other nerv syst proc age <70 w/o C.C.	.7164
9	M	Spinal disorders + injuries	1.3813
10	M	Nervous system neoplasms age >69 and/or C.C.	1.2951
11	M	Nervous system neoplasms age <70 w/o C.C.	1.2415
12	M	Degenerative nervous system disorders	1.1020
13	M	Multiple sclerosis + cerebellar ataxia	1.0045
14	M	Specific cerebrovascular disorders except TIA	1.3386
15	M	Transient ischemic attacks	.6604
16	M	Nonspecific cerebrovascular disorders with C.C.	.8503
17	M	Nonspecific cerebrovascular disorders w/o C.C.	.8305
18	M	Cranial + peripheral nerve disorders age >69 and/or C.C.	.7833
19	M	Cranial + peripheral nerve disorders age <70 w/o C.C.	.6903
20	M	Nervous system infection except viral meningitis	1.3004
21*	M	Viral meningitis	.6236
22	M	Hypertensive encephalopathy	.7787
23	M	Nontraumatic stupor + coma	1.1448
24	M	Seizure + headache age >69 and/or C.C.	.7203
25	M	Seizure + headache age 18–69 w/o C.C.	.6326
26*	M	Seizure + headache age 0–17	.4304
27*	M	Traumatic stupor + coma, coma >1 hr	1.1250
28	M	Traumatic stupor + coma, coma <1 hr age >69 and/or C.C.	1.0590
29*	M	Traumatic stupor + coma <1 hr age 18–69 w/o C.C.	.7100
30*	M	Traumatic stupor + coma <1 hr age 0–17	.3539
31	M	Concussion age >69 and/or C.C.	.5988
32	M	Concussion age 18–69 w/o C.C.	.4472
33*	M	Concussion age 0–17	.2457
34	M	Other disorders of nervous system age >69 and/or C.C.	.9824
35	M	Other disorders of nervous system age <70 w/o C.C.	.8376
MDC 2: Diseases and Disorders of the Eye			
36	S	Retinal procedures	.7019
37	S	Orbital procedures	.5571
38	S	Primary iris procedures	.4280
39	S	Lens procedures	.4958
40	S	Extraocular procedures except orbit age >17	.3936
41*	S	Extraocular procedures except orbit age 0–17	.3657
42	S	Intraocular procedures except retina, iris + lens	.5845
43*	M	Hyphema	.3788
44	M	Acute major eye infections	.6233
45	M	Neurological eye disorders	.5582
46	M	Other disorders of the eye age >17 with C.C.	.5902
47	M	Other disorders of the eye age >17 w/o C.C.	.5011
48*	M	Other disorders of the eye age 0–17	.4018

MDC 3: Diseases and Disorders of the Ear, Nose and Throat

49	S	Major head + neck procedures	2.5007
50	S	Sialoadenectomy	.7086
51	S	Salivary gland procedures except sialoadenectomy	.6632
52*	S	Cleft lip + palate repair	.6421
53	S	Sinus + mastoid procedures age >17	.5834
54*	S	Sinus + mastoid procedures age 0–17	.6889
55	S	Miscellaneous ear, nose + throat procedures	.4110
56	S	Rhinoplasty	.4101
57*	S	T + A proc except tonsillectomy +/or adenoidectomy only, age >17	.5196
58*	S	T + A proc except tonsillectomy +/or adenoidectomy only, age 0–17	.3097
59*	S	Tonsillectomy and/or adenoidectomy only age >17	.3114
60*	S	Tonsillectomy and/or adenoidectomy only age 0–17	.2616
61*	S	Myringotomy age >17	.4229
62*	S	Myringotomy age 0–17	.3089
63	S	Other ear, nose + throat O.R. procedure	1.0975
64	M	Ear, nose + throat malignancy	1.0700
65	M	Dysequilibrium	.4807
66	M	Epistaxis	.4073
67*	M	Epiglottitis	.6692
68	M	Otitis media + URI age >69 and/or C.C.	.6224
69	M	Otitis media + URI age 18–69 w/o C.C.	.5361
70*	M	Otitis media + URI age 0–17	.3659
71*	M	Laryngotracheitis	.3552
72	M	Nasal trauma + deformity	.4807
73	M	Other ear, nose + throat diagnoses age >17	.5163
74*	M	Other ear, nose + throat diagnoses age 0–17	.3427

MDC 4: Diseases and Disorders of the Respiratory System

75	S	Major chest procedures	2.5773
76	S	O.R. proc on the resp system except major chest with C.C.	1.8539
77	S	O.R. proc on the resp system except major chest w/o C.C.	1.7989
78	M	Pulmonary embolism	1.3949
79	M	Respiratory infections + inflammations age >69 and/or C.C.	1.7795
80	M	Respiratory infections + inflammations age 18–69 w/o C.C.	1.7264
81*	M	Respiratory infections + inflammations age 0–17	.8652
82	M	Respiratory neoplasms	1.1282
83	M	Major chest trauma age >69 and/or C.C.	.9707
84*	M	Major chest trauma age <70 w/o C.C.	.7658
85	M	Pleural effusion age >69 and/or C.C.	1.1342
86	M	Pleural effusion age <70 w/o C.C.	1.1100
87	M	Pulmonary edema + respiratory failure	1.5368
88	M	Chronic obstructive pulmonary disease	1.0304
89	M	Simple pneumonia + pleurisy age >69 and/or C.C.	1.0914
90	M	Simple pneumonia + pleurisy age 18–69 w/o C.C.	.9747
91*	M	Simple pneumonia + pleurisy age 0–17	.5078
92	M	Interstitial lung disease age >69 and/or C.C.	1.0262
93	M	Interstitial lung disease age <70 w/o C.C.	.9623
94	M	Pneumothorax age >69 and/or C.C.	1.4225
95	M	Pneumothorax age <70 w/o C.C.	1.1135
96	M	Bronchitis + asthma age >69 and/or C.C.	.7913
97	M	Bronchitis + asthma age 18–69 w/o C.C.	.7181
98*	M	Bronchitis + asthma age 0–17	.4231
99	M	Respiratory signs + symptoms age >69 and/or C.C.	.7952
100	M	Respiratory signs + symptoms age <70 w/o C.C.	.7650
101	M	Other respiratory diagnoses age >69 and/or C.C.	.8941
102	M	Other respiratory diagnoses age <70	.8930

MDC 5: Diseases and Disorders of the Circulatory System

103*	S	Heart transplant	.0000
104**	S	Cardiac valve procedure with pump + with cardiac cath	6.7815
105**	S	Cardiac valve procedure with pump and w/o cardiac cath	5.1764
106**	S	Coronary bypass with cardiac cath	5.2077
107**	S	Coronary bypass w/o cardiac cath	3.9476
108	S	Cardiothor proc, except valve + coronary bypass, with pump	4.3301
109	S	Cardiothoracic procedures w/o pump	3.6579
110	S	Major reconstructive vascular procedures age >69 and/or C.C.	2.9023
111	S	Major reconstructive vascular procedures age <70 w/o C.C.	2.5582
112	S	Vascular procedures except major reconstruction	2.3256
113	S	Amputation for circ system disorders except upper limb + toe	2.6522
114	S	Upper limb + toe amputation for circ system disorders	2.0848
115	S	Permanent cardiac pacemaker implant with AMI or CHF	3.8743
116	S	Permanent cardiac pacemaker implant w/o AMI or CHF	2.8367
117	S	Cardiac pacemaker replace + revis exc pulse gen repl only	1.8021
118	S	Cardiac pacemaker pulse generator replacement only	1.7624
119	S	Vein ligation + stripping	1.0500
120	S	Other O.R. procedures on the circulatory system	2.4942
121**	M	Circulatory disorders with AMI + c.v. comp. disch. alive	1.8454
122**	M	Circulatory disorders with AMI w/o c.v. comp. disch. alive	1.3509
123	M	Circulatory disorders with AMI, expired	1.1242
124	M	Circulatory disorders exc AMI, with card cath & comp diag	2.1969
125	M	Circulatory disorders exc AMI, with card cath uncomp DX 1	1.6284
126	M	Acute + subacute endocarditis	2.6368
127	M	Heart failure + shock	1.0300
128	M	Deep vein thrombophlebitis	.8549
129	M	Cardiac arrest, unexplained	1.5345
130	M	Peripheral vascular disorders age >69 and/or C.C.	.9545
131	M	Peripheral vascular disorders age <70 w/o C.C.	.9392
132	M	Atherosclerosis age >69 and/or C.C.	.9087
133	M	Atherosclerosis age <70 w/o C.C.	.8510
134	M	Hypertension	.6976
135	M	Cardiac congenital + valvular disorders age >69 and/or C.C.	.9819
136	M	Cardiac congenital + valvular disorders age 18–69 w/o C.C.	.9573
137*	M	Cardiac congenital + valvular disorders age 0–17	.6315
138	M	Cardiac arrhythmia + conduction disorders age >69 and/or C.C.	.9200
139	M	Cardiac arrhythmia + conduction disorders age <70 w/o C.C.	.8217
140	M	Angina pectoris	.7470
141	M	Syncope + collapse age >69 and/or C.C.	.6408
142	M	Syncope + collapse age <70 w/o C.C.	.5621
143	M	Chest pain	.6743
144	M	Other circulatory diagnoses with C.C.	1.1150
145	M	Other circulatory diagnoses w/o C.C.	.9916

MDC 6: Diseases and Disorders of the Digestive System

146	S	Rectal resection age >69 and/or C.C.	2.6801
147	S	Rectal resection age <70 w/o C.C.	2.4826
148	S	Major small + large bowel procedures age >69 and/or C.C.	2.5228
149	S	Major small + large bowel procedures age <70 w/o C.C.	2.1924
150	S	Peritoneal adhesiolysis age >69 and/or C.C.	2.3499
151	S	Peritoneal adhesiolysis age <70 w/o C.C.	2.0063
152	S	Minor small + large bowel procedures age >69 and/or C.C.	1.4697
153	S	Minor small + large bowel procedures age <70 w/o C.C.	1.2468
154	S	Stomach, esophageal + duodenal procedures age >69 and/or C.C.	2.6621
155	S	Stomach, esophageal + duodenal procedures age 18–69 w/o C.C.	2.3094
156*	S	Stomach, esophageal + duodenal procedures age 0–17	.8382
157	S	Anal procedures age >69 and/or C.C.	.7902
158	S	Anal procedures age <70 w/o C.C.	.6341
159	S	Hernia procedures except inguinal + femoral age >69 and/or C.C.	.9200
160	S	Hernia procedures except inguinal + femoral age 18–69 w/o C.C.	.7596
161	S	Inguinal + femoral hernia procedures age >69 and/or C.C.	.6995
162	S	Inguinal + femoral hernia procedures age 18–69 w/o C.C.	.5793
163*	S	Hernia procedures age 0–17	.4313

164	S	Appendectomy with complicated princ. diag age >69 and/or C.C.	1.8130
165	S	Appendectomy with complicated princ. diag age <70 w/o C.C.	1.5986
166	S	Appendectomy w/o complicated princ. diag age >69 and/or C.C.	1.4179
167	S	Appendectomy w/o complicated princ. diag age <70 w/o C.C.	1.0706
168	S	Procedures on the mouth age >69 and/or C.C.	.8541
169	S	Procedures on the mouth age <70 w/o C.C.	.8899
170	S	Other digestive system procedures age >69 and/or C.C.	2.6326
171	S	Other digestive system procedures age <70 w/o C.C.	2.3727
172	M	Digestive malignancy age >69 and/or C.C.	1.2141
173	M	Digestive malignancy age <70 w/o C.C.	1.0408
174	M	G.I. hemorrhage age >69 and/or C.C.	.9185
175	M	G.I. hemorrhage age <70 w/o C.C.	.8150
176	M	Complicated peptic ulcer	1.2309
177	M	Uncomplicated peptic ulcer >69 and/or C.C.	.7345
178	M	Uncomplicated peptic ulcer <70 w/o C.C.	.6077
179	M	Inflammatory bowel disease	1.0048
180	M	G.I. obstruction age >69 and/or C.C.	.8112
181	M	G.I. obstruction age <70 w/o C.C.	.7763
182	M	Esophagitis, gastroent. + misc. digest. dis age >69 and/or C.C.	.6121
183	M	Esophagitis, gastroent. + misc. digest. dis age 18–69 w/o C.C.	.5593
184*	M	Esophagitis, gastroenteritis + misc. digest. disorders age 0–17	.3782
185	M	Dental + oral dis. exc extractions + restorations, age >17	.6612
186*	M	Dental + oral dis. exc extractions + restorations, age 0–17	.4112
187	M	Dental extractions + restorations	.3949
188	M	Other digestive system diagnoses age >69 and/or C.C.	.7367
189	M	Other digestive system diagnoses age 18–69 w/o C.C.	.6508
190*	M	Other digestive system diagnoses age 0–17	.3344

MDC 7: Diseases and Disorders of the Hepatobiliary System and Pancreas

191	S	Major pancreas, liver + shunt procedures	4.1357
192*	S	Minor pancreas, liver + shunt procedures	3.8790
193	S	Biliary tract proc exc tot cholecystectomy age >69 and/or C.C.	2.4258
194	S	Biliary tract proc exc tot cholecystectomy age <70 w/o C.C.	1.9674
195**	S	Total cholecystectomy w c.d.e. age >69 and/or C.C.	2.1465
196**	S	Total cholecystectomy w c.d.e. age <70 w/o C.C.	2.0380
197**	S	Total cholecystectomy w/o c.d.e. age >69 and/or C.C.	1.4714
198**	S	Total cholecystectomy w/o c.d.e. age <70 w/o C.C.	1.2619
199	S	Hepatobiliary diagnostic procedure for malignancy	2.4319
200	S	Hepatobiliary diagnostic procedure for non-malignancy	2.5550
201	S	Other hepatobiliary or pancreas O.R. procedures	2.7007
202	M	Cirrhosis + alcoholic hepatitis	1.1841
203	M	Malignancy of hepatobiliary system or pancreas	1.0823
204	M	Disorders of pancreas except malignancy	.9581
205	M	Disorders of liver exc malig, cirr, alc hepa age >69 and/or C.C.	1.0710
206	M	Disorders of liver exc malig, cirr, alc hepa age <70 w/o C.C.	.9151
207	M	Disorders of the biliary tract age >69 and/or C.C.	.8404
208	M	Disorders of the biliary tract age <70 w/o C.C.	.7239

MDC 8: Diseases and Disorders of the Musculoskeletal System and Connective Tissue

209	S	Major joint procedures	2.2674
210	S	Hip + femur procedures except major joint age >69 and/or C.C.	2.0617
211	S	Hip + femur procedures except major joint age 18–69 w/o C.C.	1.9327
212*	S	Hip + femur procedures except major joint age 0–17	1.6954
213	S	Amputations for musculoskeletal system + conn. tissue disorders	2.1094
214	S	Back + neck procedures age >69 and/or C.C.	1.8236
215	S	Back + neck procedures age <70 w/o C.C.	1.4765
216	S	Biopsies of musculoskeletal system + connective tissue	1.5434
217	S	Wnd debrid + skn grft exc hand, for muscskseletal + conn. tiss. dis	2.2587
218	S	Lower extrem + humer proc exc hip, foot, femur age >69 and/or C.C.	1.4102
219	S	Lower extrem + humer proc exc hip, foot, femur age 18–69 w/o C.C.	1.0678
220*	S	Lower extrem + humer proc exc hip, foot, femur age 0–17	.9242
221	S	Knee procedures age >69 and/or C.C.	1.2595

222	S	Knee procedures age <70 w/o C.C.	.9794
223	S	Upper extremity proc exc humerus + hand age >69 and/or C.C.	1.0612
224	S	Upper extremity proc exc humerus + hand age <70 w/o C.C.	.8859
225	S	Foot procedures	.6409
226	S	Soft tissue procedures age >69 and/or C.C.	.7901
227	S	Soft tissue procedures age <70 w/o C.C.	.6271
228	S	Ganglion hand procedures	.3588
229	S	Hand procedures except ganglion	.5936
230	S	Local excision + removal of int fix devices of hip + femur	1.3453
231	S	Local excision + removal of int fix devices except hip + femur	.9420
232	S	Arthroscopy	.6000
233	S	Other musculoskelet sys + conn tiss O.R. proc age >69 and/or C.C.	1.7553
234	S	Other musculoskelet sys + conn tiss O.R. proc age <70 w/o C.C.	1.2325
235	M	Fractures of femur	1.7403
236	M	Fractures of hip + pelvis	1.3711
237	M	Sprains, strains, + dislocations of hip, pelvis + thigh	.7847
238	M	Osteomyelitis	1.5350
239	M	Pathological fractures + musculoskeletal + conn. tiss. malignancy	1.0865
240	M	Connective tissue disorders age >69 and/or C.C.	.9608
241	M	Connective tissue disorders age <70 w/o C.C.	.8954
242	M	Septic arthritis	1.5715
243	M	Medical back problems	.7473
244	M	Bone diseases + septic arthropathy age >69 and/or C.C.	.7711
245	M	Bone diseases + septic arthropathy age <70 w/o C.C.	.7102
246	M	Non-specific arthropathies	.7073
247	M	Signs + symptoms of musculoskeletal system + conn tissue	.6491
248	M	Tendonitis, myositis + bursitis	.6072
249	M	Aftercare, musculoskeletal system + connective tissue	1.0097
250	M	Fx, sprns, strns + disl of forearm, hand, foot age >69 and/or C.C.	.7351
251	M	Fx, sprns, strns + disl of forearm, hand, foot age 18–69 w/o C.C.	.5902
252*	M	Fx, sprns, strns + disl of forearm, hand, foot age 0–17	.3496
253	M	Fx, sprns, strns + disl of uparm, lowleg ex foot age >69 and/or C.C.	.7388
254	M	Fx, sprns, strns + disl of uparm, lowleg ex foot age 18–69 w/o C.C.	.6193
255*	M	Fx, sprns, strns + disl of uparm, lowleg ex foot age 0–17	.4638
256	M	Other diagnoses of musculoskeletal system + connective tissue	.8616

MDC 9: Diseases and Disorders of the Skin, Subcutaneous Tissue and Breast

257	S	Total mastectomy for malignancy age >69 and/or C.C.	1.0970
258	S	Total mastectomy for malignancy age <70 w/o C.C.	1.0618
259	S	Subtotal mastectomy for malignancy age >69 and/or C.C.	1.0036
260	S	Subtotal mastectomy for malignancy age <70	.9228
261	S	Breast proc for non-malig except biopsy + loc exc	.7253
262	S	Breast biopsy + local excision for non-malignancy	.4569
263	S	Skin grafts for skin ulcer or cellulitis age >69 and/or C.C.	2.4480
264	S	Skin grafts for skin ulcer or cellulitis age <70 w/o C.C.	2.1802
265*	S	Skin grafts except for skin ulcer or cellulitis with C.C.	1.4804
266	S	Skin grafts except for skin ulcer or cellulitis w/o C.C.	.9386
267	S	Perianal + pilonidal procedures	.6049
268	S	Skin, subcutaneous tissue + breast plastic procedures	.5332
269	S	Other skin, subcut tiss + breast O.R. proc age >69 and/or C.C.	.9844
270	S	Other skin, subcut tiss + breast O.R. proc age <70 w/o C.C.	.8039
271	M	Skin ulcers	1.3659
272	M	Major skin disorders age >69 and/or C.C.	.8530
273	M	Major skin disorders age <70 w/o C.C.	.8200
274	M	Malignant breast disorders age >69 and/or C.C.	1.0003
275	M	Malignant breast disorders age <70 w/o C.C.	.8920
276	M	Non-malignant breast disorders	.6003
277	M	Cellulitis age >69 and/or C.C.	.8771
278	M	Cellulitis age 18–69 w/o C.C.	.8012
279*	M	Cellulitis age 0–17	.4739
280	M	Trauma to the skin, subcut tiss + breast age >69 and/or C.C.	.6137
281	M	Trauma to the skin, subcut tiss + breast age 18–69 w/o C.C.	.5321
282*	M	Trauma to the skin, subcut tiss + breast age 0–17	.3424
283	M	Minor skin disorders age >69 and/or C.C.	.6328
284	M	Minor skin disorders age <70 w/o C.C.	.5909

MDC 10: Endocrine, Nutritional and Metabolic Diseases and Disorders

285	S	Amputations for endocrine, nutritional + metabolic disorders	2.8360
286*	S	Adrenal + pituitary procedures	2.8651
287	S	Skin grafts + wound debride for endoc, nutrit + metab disorders	2.7851
288*	S	O.R. procedures for obesity	1.5532
289	S	Parathyroid procedures	1.3593
290	S	Thyroid procedures	.8460
291*	S	Thyroglossal procedures	.4858
292	S	Other endocrine, nutrit + metab O.R. proc age >69 and/or C.C.	2.0096
293*	S	Other endocrine, nutrit + metab O.R. proc age <70 w/o C.C.	1.4796
294	M	Diabetes age = >36	.8003
295	M	Diabetes age 0–35	.7380
296	M	Nutritional + misc. metabolic disorders age >69 and/or C.C.	.8886
297	M	Nutritional + misc. metabolic disorders age 18–69 w/o C.C.	.7841
298*	M	Nutritional + misc. metabolic disorders age 0–17	.7460
299	M	Inborn errors of metabolism	.9309
300	M	Endocrine disorders age >69 and/or C.C.	.9630
301	M	Endocrine disorders age <70 w/o C.C.	.8058

MDC 11: Diseases and Disorders of the Kidney and Urinary Tract

302	S	Kidney transplant	4.1840
303	S	Kidney, ureter + major bladder procedure for neoplasm	2.5133
304	S	Kidney, ureter + maj bldr proc for non-malig age >69 and/or C.C.	1.7765
305	S	Kidney, ureter + maj bldr proc for non-malig age <70 w/o C.C.	1.6866
306	S	Prostatectomy age >69 and/or C.C.	1.1281
307	S	Prostatectomy age <70 w/o C.C.	.9414
308	S	Minor bladder procedures age >69 and/or C.C.	1.0333
309	S	Minor bladder procedures age <70 w/o C.C.	.9193
310	S	Transurethral procedures age >69 and/or C.C.	.6998
311	S	Transurethral procedures age <70 w/o C.C.	.5810
312	S	Urethral procedures, age >69 and/or C.C.	.7347
313	S	Urethral procedures, age 18–69 w/o C.C.	.6825
314*	S	Urethral procedures, age 0–17	.4323
315	S	Other kidney + urinary tract O.R. procedures	2.4625
316	M	Renal failure	1.3176
317*	M	Admit for renal dialysis	.2360
318	M	Kidney + urinary tract neoplasms age >69 and/or C.C.	.9047
319	M	Kidney + urinary tract neoplasms age <70 w/o C.C.	.7859
320	M	Kidney + urinary tract infections age >69 and/or C.C.	.8039
321	M	Kidney + urinary tract infections age 18–69 w/o C.C.	.6732
322*	M	Kidney + urinary tract infections age 0–17	.4506
323	M	Urinary stones age >69 and/or C.C.	.7057
324	M	Urinary stones age <70 w/o C.C.	.5415
325	M	Kidney + urinary tract signs + symptoms age >69 and/or C.C.	.7172
326	M	Kidney + urinary tract signs + symptoms age 18–69 w/o C.C.	.5814
327*	M	Kidney + urinary tract signs + symptoms age 0–17	.4975
328	M	Urethral stricture age >69 and/or C.C.	.6440
329	M	Urethral stricture age 18–69 w/o DX 2	.5271
330*	M	Urethral stricture age 0–17	.2788
331	M	Other kidney + urinary tract diagnoses age >69 and/or C.C.	.8826
332	M	Other kidney + urinary tract diagnoses age 18–69 w/o C.C.	.7682
333*	M	Other kidney + urinary tract diagnoses age 0–17	.5093

MDC 12: Diseases and Disorders of the Male Reproductive System

334	S	Major male pelvic procedures with C.C.	1.5450
335	S	Major male pelvic procedures w/o C.C.	1.3449
336	S	Transurethral prostatectomy age >69 and/or C.C.	.9974
337	S	Transurethral prostatectomy age <70 w/o C.C.	.8403
338	S	Testes procedures, for malignancy	.8975

339	S	Testes procedures, non-malignant age >17	.6030
340*	S	Testes procedures, non-malignant age 0–17	.4335
341	S	Penis procedures	.9879
342	S	Circumcision age >17	.4184
343*	S	Circumcision age 0–17	.3788
344	S	Other male reproductive system O.R. procedures for malignancy	1.1088
345	S	Other male reproductive system O.R. proc except for malignancy	.8247
346	M	Malignancy, male reproductive system, age >69 and/or C.C.	.9297
347	M	Malignancy, male reproductive system, age <70 w/o C.C.	.8218
348	M	Benign prostatic hypertrophy age >69 and/or C.C.	.8772
349	M	Benign prostatic hypertrophy age <70 w/o C.C.	.6925
350	M	Inflammation of the male reproductive system	.6033
351*	M	Sterilization, male	.2627
352	M	Other male reproductive system diagnoses	.6319

MDC 13: Diseases and Disorders of the Female Reproductive System

353	S	Pelvic evisceration, radical hysterectomy + vulvectomy	1.9175
354	S	Non-radical hysterectomy age >69 and/or C.C.	1.0993
355	S	Non-radical hysterectomy age <70 w/o C.C.	1.0050
356	S	Female reproductive system reconstructive procedures	.8372
357	S	Uterus + adenexa procedures, for malignancy	1.8989
358	S	Uterus + adenexa proc for non-malignancy except tubal interrupt	1.0777
359*	S	Tubal interruption for non-malignancy	.4235
360	S	Vagina, cervix + vulva procedures	.5923
361*	S	Laparoscopy + endoscopy (female) except tubal interruption	.4813
362*	S	Laparoscopic tubal interruption	.3094
363	S	D+C, conization + radio-implant, for malignancy	.6448
364	S	D+C, conization except for malignancy	.3986
365	S	Other female reproductive system O.R. procedures	1.7778
366	M	Malignancy, female reproductive system age >69 and/or C.C.	.8356
367	M	Malignancy, female reproductive system age <70 w/o C.C.	.5726
368	M	Infections, female reproductive system	.7861
369	M	Menstrual + other female reproductive system disorders	.6887

MDC 14: Pregnancy, Childbirth, and the Puerperium

370*	S	Cesarean section with C.C.	.9809
371*	S	Cesarean section w/o C.C.	.7457
372*	M	Vaginal delivery with complicating diagnoses	.5476
373	M	Vaginal delivery w/o complicating diagnoses	.4021
374*	S	Vaginal delivery with sterilization and/or D+C	.5435
375*	S	Vaginal delivery with O.R. proc except steril and/or D+C	.6817
376*	M	Postpartum diagnoses w/o O.R. procedure	.4115
377*	S	Postpartum diagnoses with O.R. procedure	.4712
378*	M	Ectopic pregnancy	.8010
379*	M	Threatened abortion	.3136
380*	M	Abortion w/o D+C	.2677
381*	M	Abortion with D+C	.3565
382*	M	False labor	.1823
383*	M	Other antepartum diagnoses with medical complications	.4272
384*	M	Other antepartum diagnoses w/o medical complications	.3211

MDC 15: Newborns and Other Neonates with Conditions Originating in the Perinatal Period

385*	****	**Neonates, died or transferred**	.6811
386*	****	**Extreme immaturity, neonate**	3.6480
387**	****	**Combined with 388**	1.8267
388**	****	**Prematurity w and w/o major problems**	1.1571
389*	****	**Full term neonate with major problems**	.5425
390*	****	**Neonates with other significant problems**	.3486
391*	****	**Normal newborns**	.2218

MDC 16: Diseases and Disorders of the Blood and Blood-Forming Organs and Immunological Disorders

392	S	Splenectomy age >17	2.7458
393*	S	Splenectomy age 0–17	1.5206
394	S	Other O.R. procedures of the blood + blood forming organs	1.1030
395	M	Red blood cell disorders age >17	.7758
396*	M	Red blood cell disorders age 0–17	.6230
397	M	Coagulation disorders	.9761
398	M	Reticuloendothelial + immunity disorders age >69 and/or C.C.	.8808
399	M	Reticuloendothelial + immunity disorders age <70 w/o C.C.	.8371

MDC 17: Myeloproliferative Diseases and Disorders and Poorly Differentiated Neoplasms

400	S	Lymphoma or leukemia with major O.R. procedure	2.7978
401	S	Lymphoma or leukemia with minor O.R. proc age >69 and/or C.C.	1.2280
402*	S	Lymphoma or leukemia with minor O.R. procedure age <70 w/o C.C.	1.1198
403	M	Lymphoma or leukemia age >69 and/or C.C.	1.1593
404	M	Lymphoma or leukemia age 18–69 w/o C.C.	1.1665
405*	M	Lymphoma or leukemia age 0–17	1.0408
406	S	Myeloprolif disord or poorly diff neoplasm w maj O.R. proc + C.C.	2.2435
407	S	Myeloprolif disord or poorly diff neopl w maj O.R. proc w/o C.C.	2.1144
408	S	Myeloprolif disord or poorly diff neopl with minor O.R. proc	1.1271
409*	M	Radiotherapy	.8049
410	M	Chemotherapy	.3490
411	M	History of malignancy w/o endoscopy	.7146
412	M	History of malignancy with endoscopy	.3365
413	M	Other myeloprolif disord or poorly diff neopl DX age >69 and/or C.C.	1.0861
414	M	Other myeloprolif disord or poorly diff neopl DX age <70 w/o C.C.	1.0251

MDC 18: Infectious and Parasitic Diseases (Systemic or Unspecified Sites)

415	S	O.R. procedure for infectious + parasitic diseases	2.9715
416	M	Septicemia age >17	1.5343
417*	M	Septicemia age 0–17	.7078
418	M	Postoperative + post-traumatic infections	.9864
419	M	Fever of unknown origin age >69 and/or C.C.	.8538
420	M	Fever of unknown origin age 18–69 w/o C.C.	.7939
421	M	Viral illness age >17	.5982
422*	M	Viral illness + fever of unknown origin age 0–17	.4315
423	M	Other infectious + parasitic diseases diagnoses	1.1981

MDC 19: Mental Diseases and Disorders

424	S	O.R. procedures with principal diagnosis of mental illness	2.1710
425	M	Acute adjust react + disturbances of psychosocial dysfunction	.6741
426	M	Oepressive neuroses	.9396
427	M	Neuroses except depressive	.7598
428	M	Disorders of personality + impulse control	.9640
429	M	Organic disturbances + mental retardation	.9424
430	M	Psychoses	1.0820
431*	M	Childhood mental disorders	2.2285
432*	M	Other diagnoses of mental disorders	1.0416

MDC 20: Substance Use and Substance Induced Organic Mental Disorders

433**	****	Substance use + substance induced organic mental disorders, left AMA	.4411

434**	****	Drug dependence	1.0296
435**	****	Drug use except dependence	1.0626
436**	****	Alcohol dependence	.8761
437**	****	Alcohol use except dependence	.6119
438**	****	Alcohol + substance induced organic mental syndrome	.8333
MDC 21: Injury, Poisoning and Toxic Effects of Drugs			
439*	S	Skin grafts for injuries	1.8030
440*	S	Wound debridements for injuries	1.4653
441*	S	Hand procedures for injuries	.7105
442	S	Other O.R. procedures for injuries age >69 and/or C.C.	1.8828
443	S	Other O.R. procedures for injuries age <70 w/o C.C.	1.5053
444	M	Multiple trauma age >69 and/or C.C.	.8738
445	M	Multiple trauma age 18–69 w/o C.C.	.7452
446*	M	Multiple trauma age 0–17	.4796
447	M	Allergic reactions age >17	.4735
448*	M	Allergic reactions age 0–17	.3469
449	M	Toxic effects of drugs age >69 and/or C.C.	.7255
450	M	Toxic effects of drugs age 18–69 w/o C.C.	.5895
451*	M	Toxic effects of drugs age 0–17	.2882
452	M	Complications of treatment age >69 and/or C.C.	.8404
453	M	Complications of treatment age <70 w/o C.C.	.8926
454	M	Other injuries, poisonings + toxic eff drugs age >69 and/or C.C.	.8139
455*	M	Other injuries, poisonings + toxic eff drugs age <70 w/o C.C.	.6121
MDC 22: Burns			
456**	****	Burns, transferred to another acute care facility	2.0685
457**	****	Extensive burns	6.7918
458**	S	Non-extensive burns with skin grafts	2.8275
459**	S	Non-extensive burns with wound debridement + O.R. proc	2.7282
460**	M	Non-extensive burns w/o O.R. procedure	1.4077
MDC 23: Factors Influencing Health Status and Other Contacts with Health Services			
461	S	O.R. proc with diagnoses of other contact with health services	1.6335
462*	M	Rehabilitation	1.8078
463	M	Signs + symptoms with C.C.	.7622
464	M	Signs + symptoms w/o C.C.	.7246
465**	M	Aftercare with history of malignancy as secondary DX	.2049
466**	M	Aftercare w/o history of malignancy as secondary DX	.6311
467	M	Other factors influencing health status	.9697
468		Unrelated O.R. procedure	2.0818
469***		PDX invalid as discharge diagnosis	.0000
470***		Ungroupable	.0000

a = Surgical (S) or Medical (M) DRG.

* MEDPAR data have been supplemented by data from Maryland and Michigan for low volume DRGs.

** DRG categories combined (in pairs) in the calculation of the case mix index.

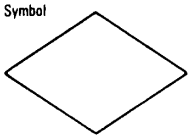
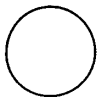
*** DRGs 469 and 470 contain cases which could not be assigned to valid DRGs.

**** According to HCFA staff these DRGs could be Surgical or Medical.

Source: *Federal Register*, vol. 49 (Aug. 31, 1984), pp. 34780–34790.

Appendix II DRG DIAGRAMS

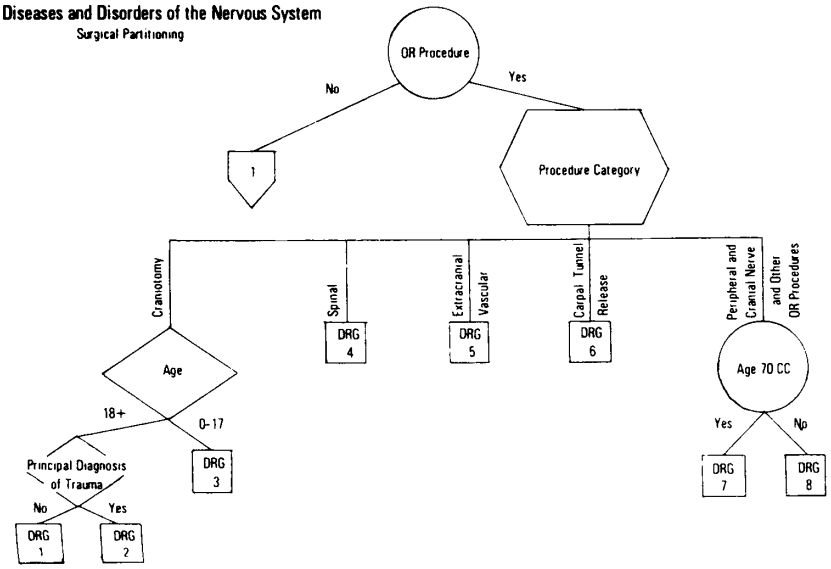
A-1
Decision Trees for the ICD-9-CM DRGs
Definitions of Symbols Used in Decision Trees

Symbol	Definition
	Decision operation
	Looping variable
	Hierarchy of operating room procedures
	Connector
	Terminal

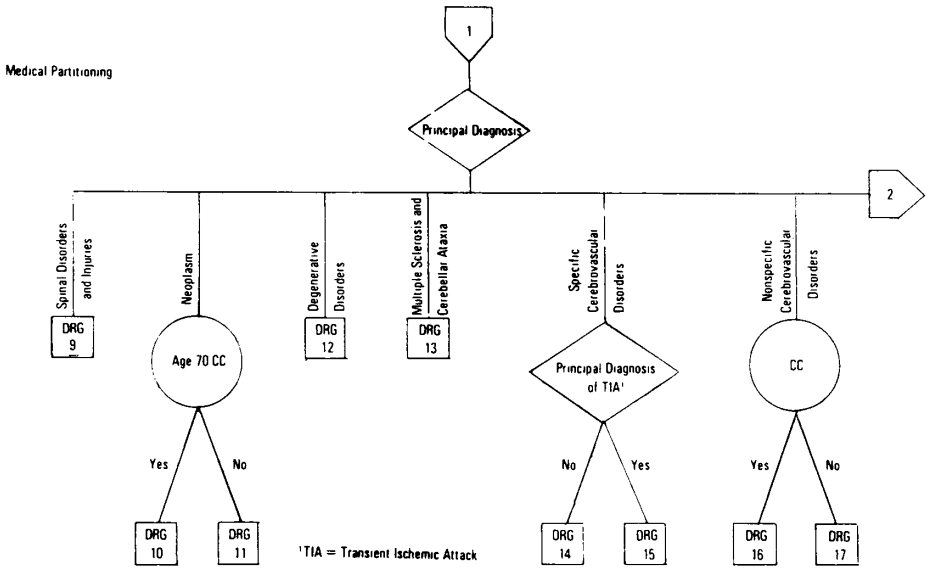
Source: *The Revised ICD-9-CM Diagnosis Related Groups: Grouper User Manual* (New Haven, CT: Health Systems International).

MDC 1: Diseases and Disorders of the Nervous System

Surgical Partitioning



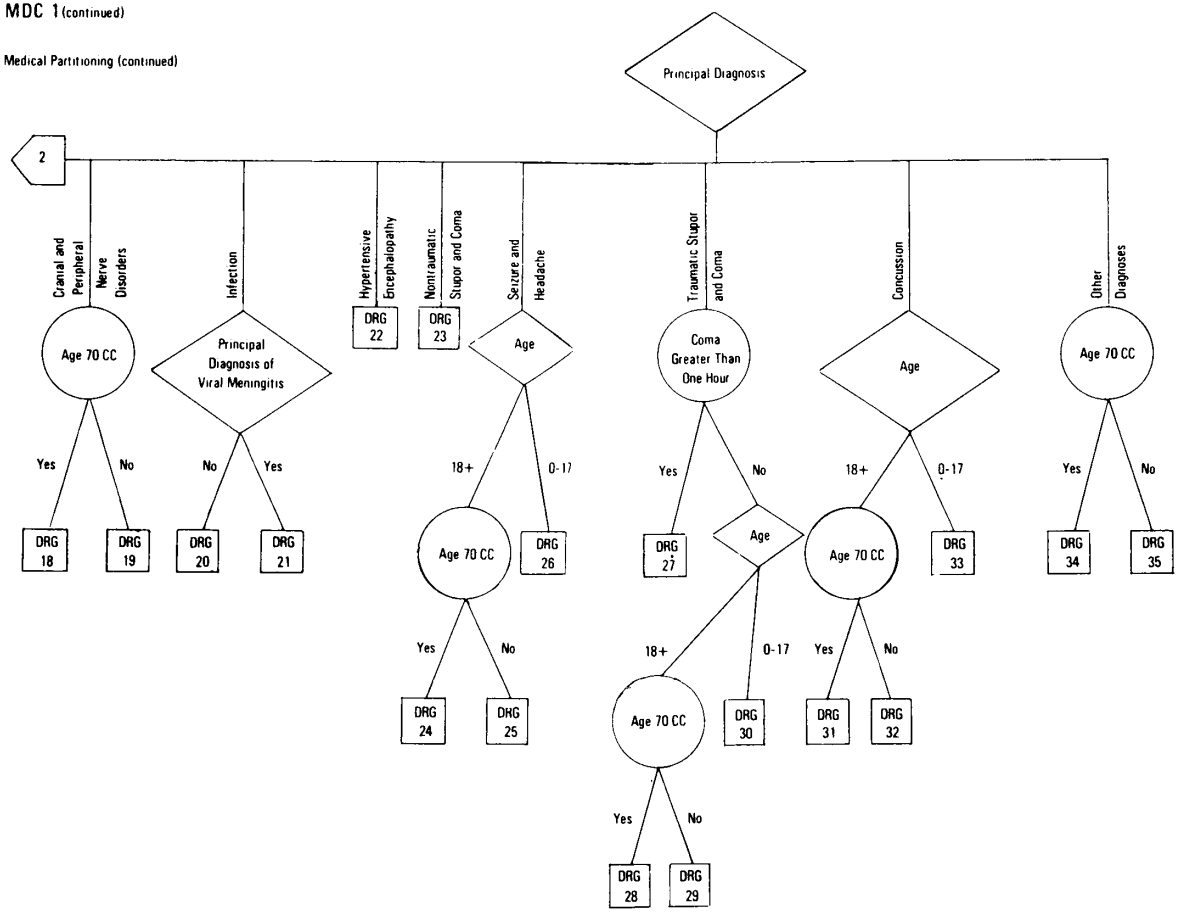
Medical Partitioning

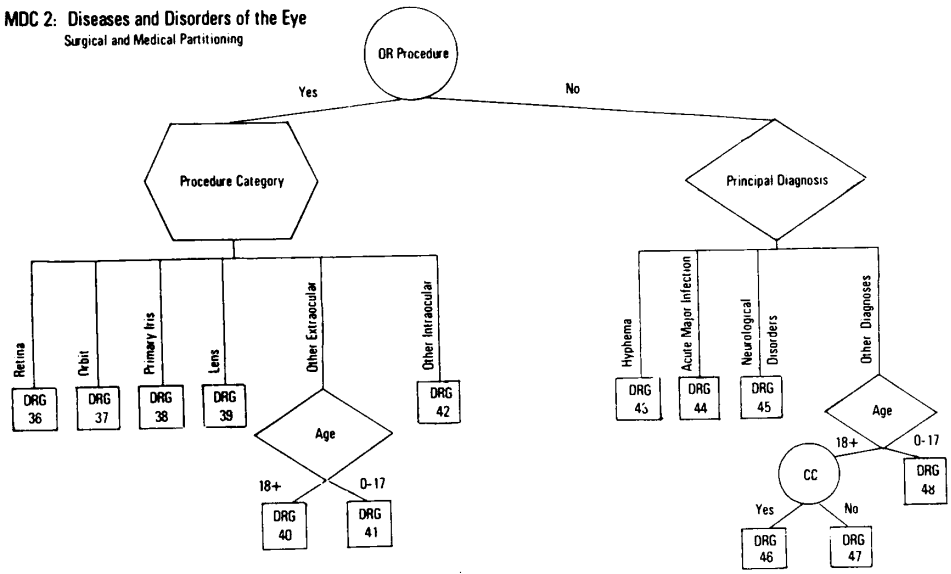


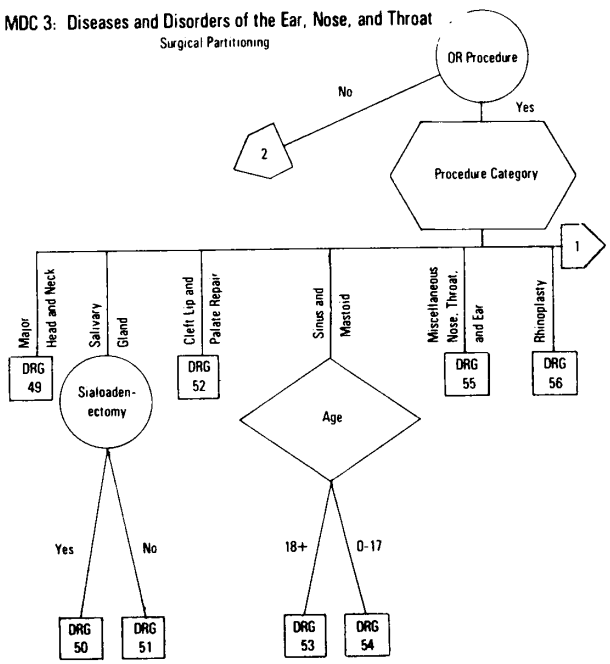
*TIA = Transient Ischemic Attack

MDC 1 (continued)

Medical Partitioning (continued)

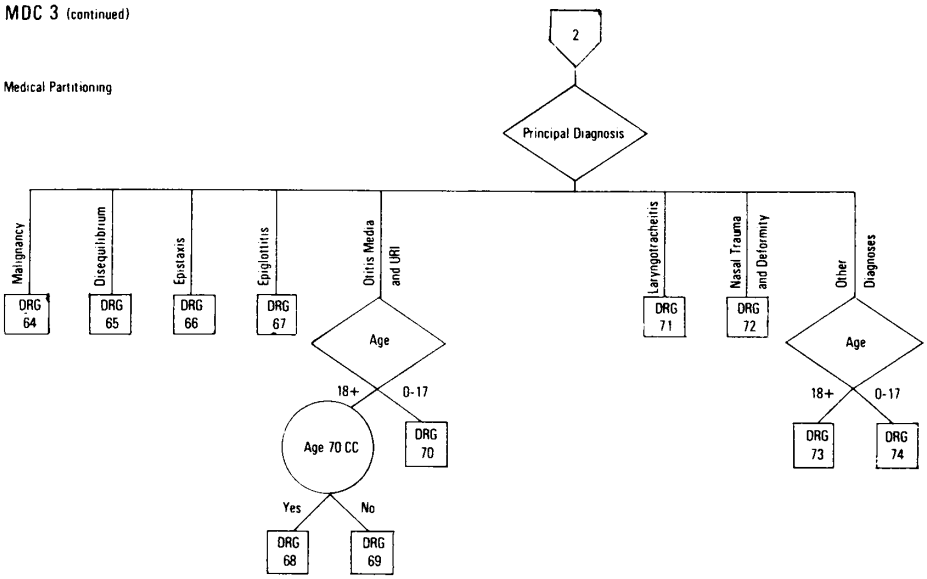


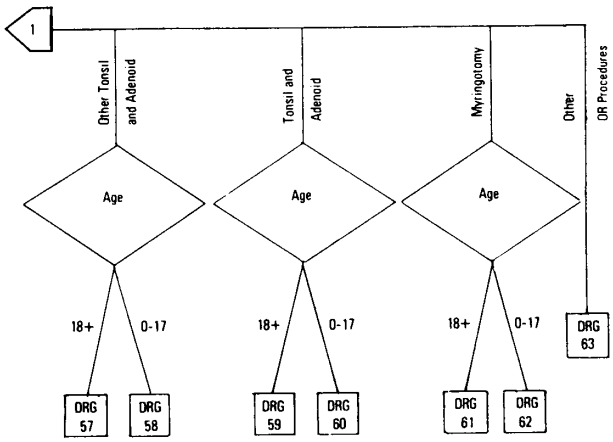

MDC 2: Diseases and Disorders of the Eye
Surgical and Medical Partitioning
OR Procedure
Yes
No
Procedure Category
Principal Diagnosis
Retina
Orbit
Primary Iris
Lens
Other Extraocular
Other Intraocular
Hyphema
Acute Major Infection
Neurological Disorders
Other Diagnoses
DRG 36
DRG 37
DRG 38
DRG 39
DRG 42
DRG 43
DRG 44
DRG 45
Age
18+
0-17
DRG 40
DRG 41
CC
Yes
No
DRG 46
DRG 47
DRG 48


MDC 3: Diseases and Disorders of the Ear, Nose, and Throat
Surgical Partitioning
OR Procedure
No
Yes
2
Procedure Category
1
Major Head and Neck
Salivary Gland
Cleft Lip and Palate Repair
Sinus and Mastoid
Miscellaneous Nose, Throat, and Ear
Rhinoplasty
DRG 49
Sialoadenectomy
DRG 52
Age
DRG 55
DRG 56
Yes
No
18+
0-17
DRG 50
DRG 51
DRG 53
DRG 54

MDC 3 (continued)

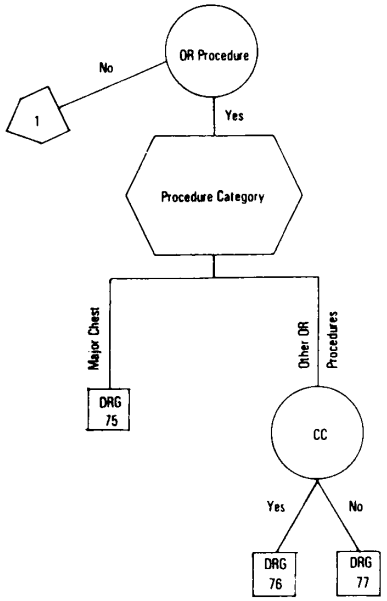
Medical Partitioning



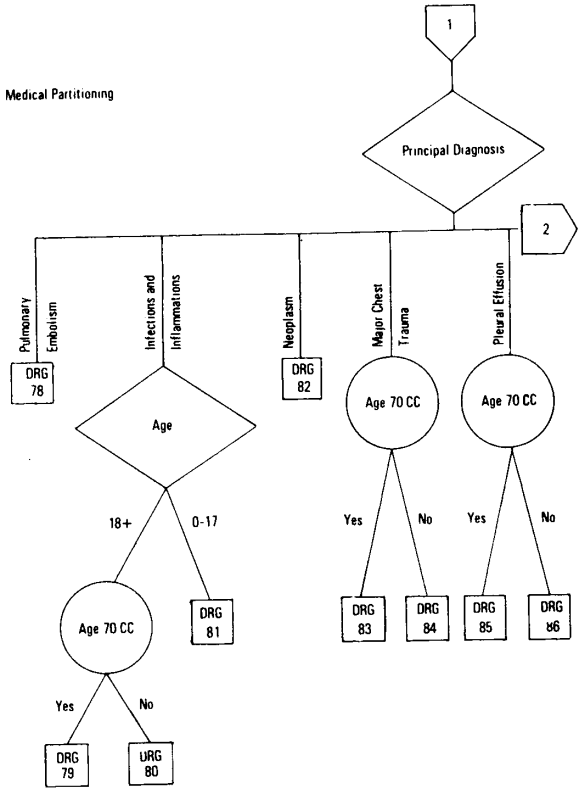


MDC 4: Diseases and Disorders of the Respiratory System

Surgical Partitioning

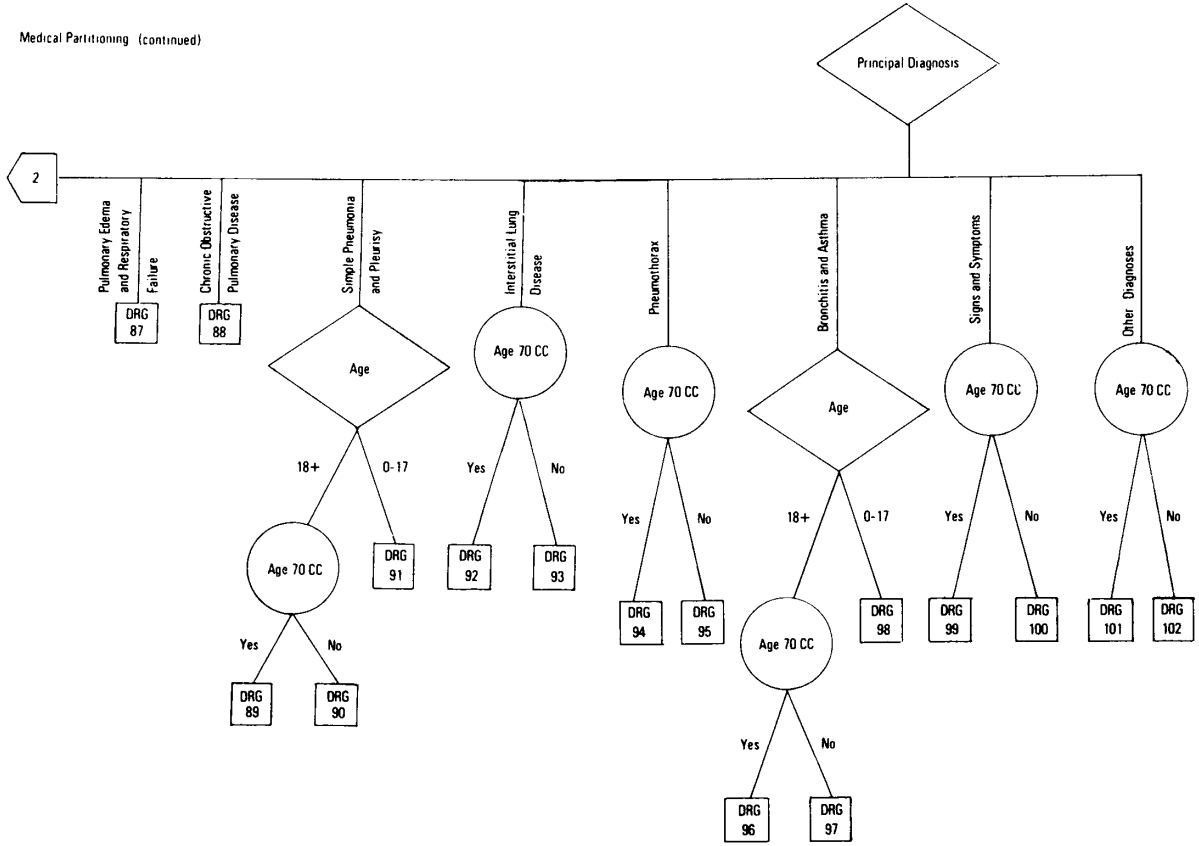


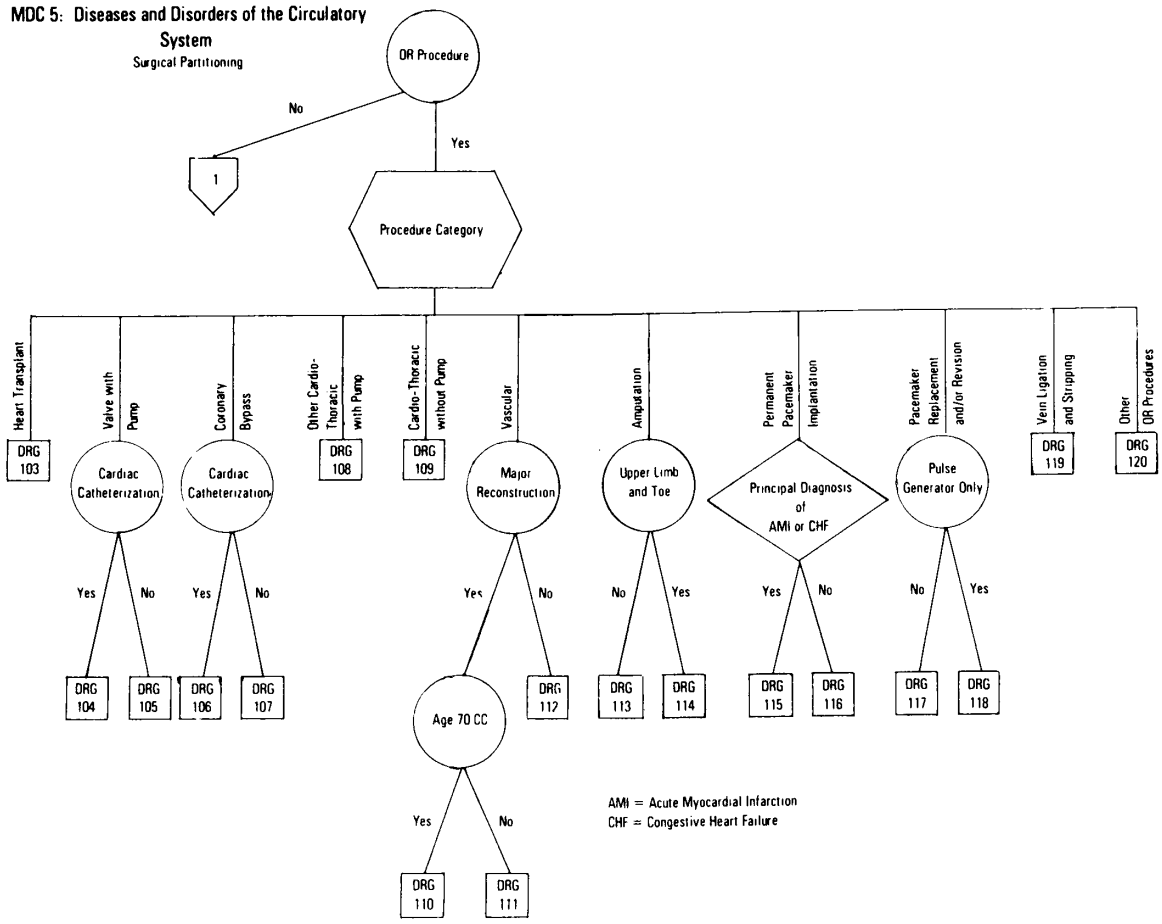
Medical Partitioning



MDC 4 (continued)

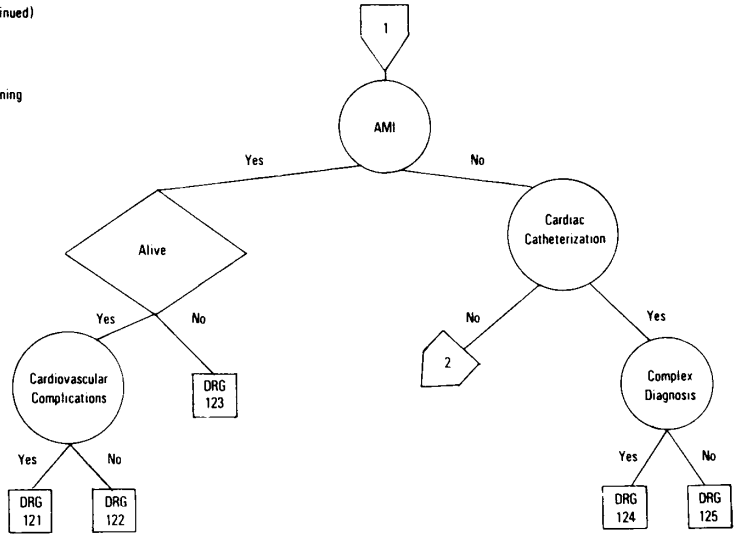
Medical Partitioning (continued)

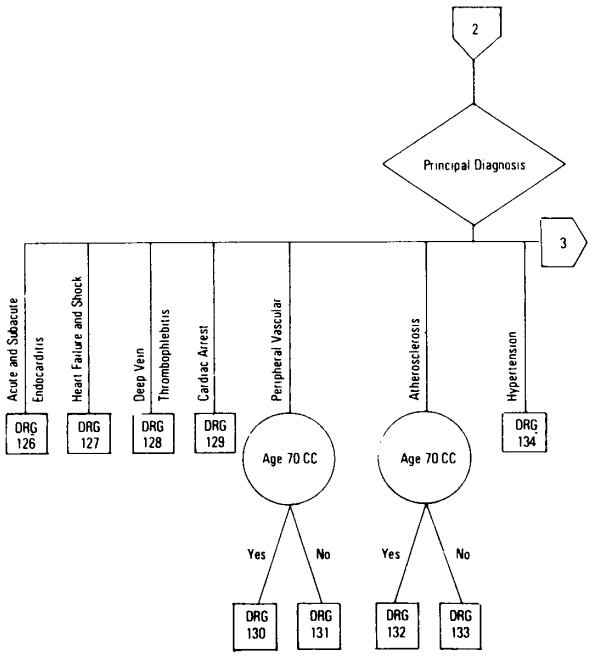



MDC 5: Diseases and Disorders of the Circulatory System
Surgical Partitioning
OR Procedure
No
Yes
1
Procedure Category
Heart Transplant
DRG 103
Valve with Pump
Cardiac Catheterization
Yes
No
DRG 104
DRG 105
Coronary Bypass
Cardiac Catheterization
Yes
No
DRG 106
DRG 107
Other Cardio-Thoracic with Pump
DRG 108
Cardio-Thoracic without Pump
DRG 109
Vascular
Major Reconstruction
Yes
No
Age 70 CC
Yes
No
DRG 110
DRG 111
DRG 112
Amputation
Upper Limb and Toe
No
Yes
DRG 113
DRG 114
Permanent Pacemaker Implantation
Principal Diagnosis of AMI or CHF
Yes
No
DRG 115
DRG 116
Pacemaker Replacement and/or Revision
Pulse Generator Only
No
Yes
DRG 117
DRG 118
Vein Ligation and Stripping
DRG 119
Other OR Procedures
DRG 120
AMI = Acute Myocardial Infarction
CHF = Congestive Heart Failure

MDC 5 (continued)

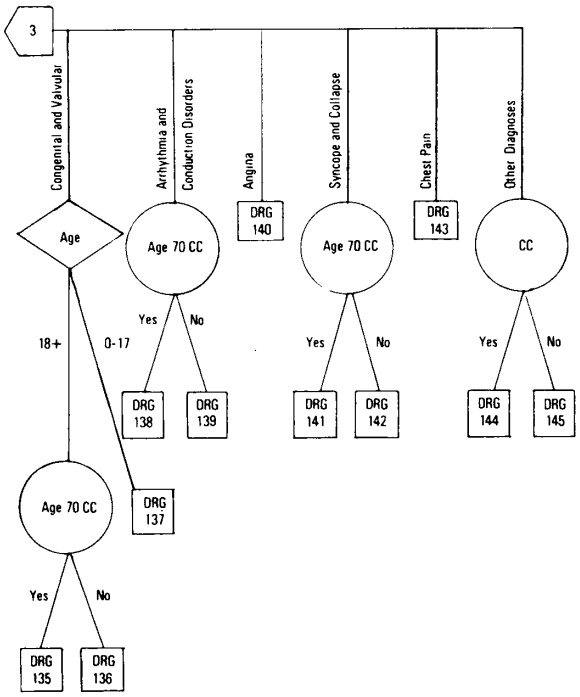
Medical Partitioning





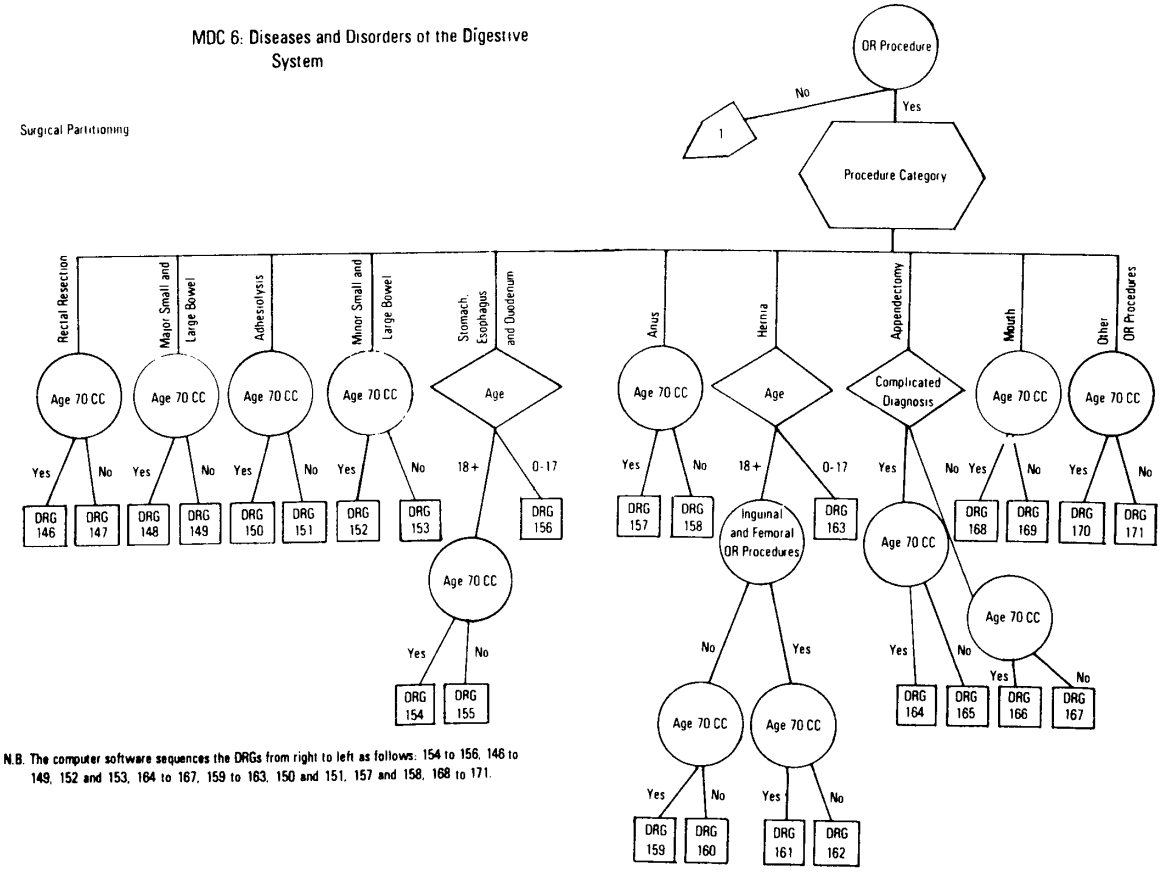
MDC 5 (continued)

Medical Partitioning (continued)



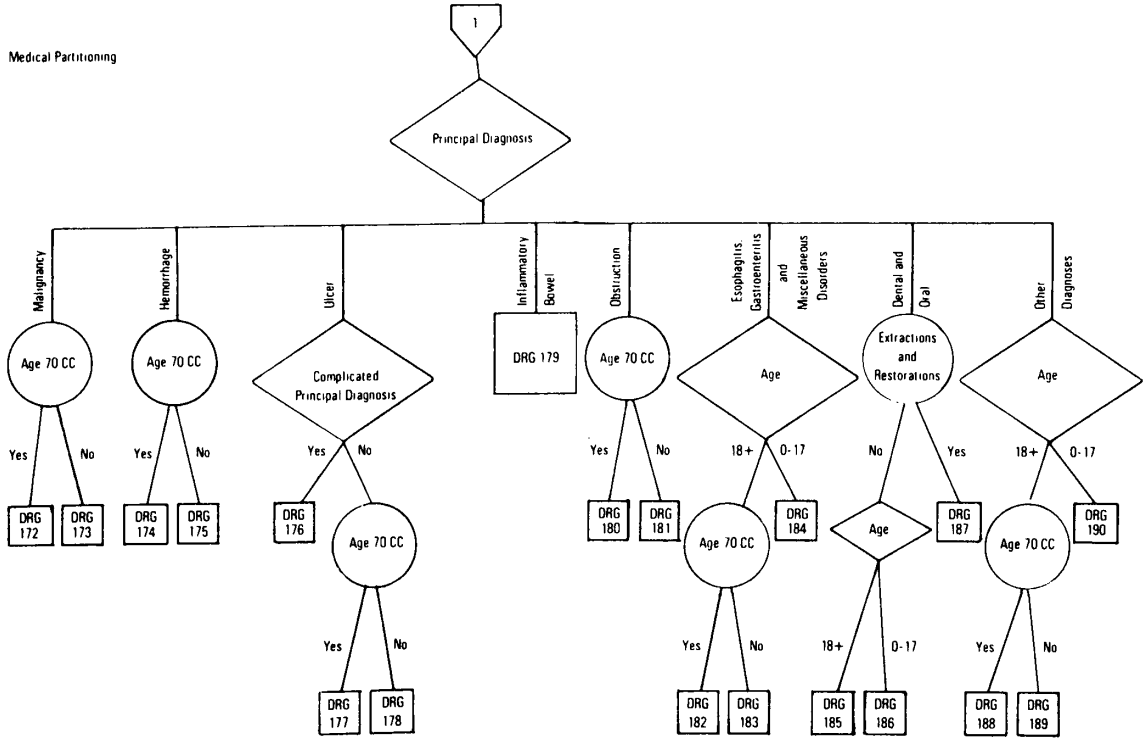
MDC 6: Diseases and Disorders of the Digestive System

Surgical Partitioning



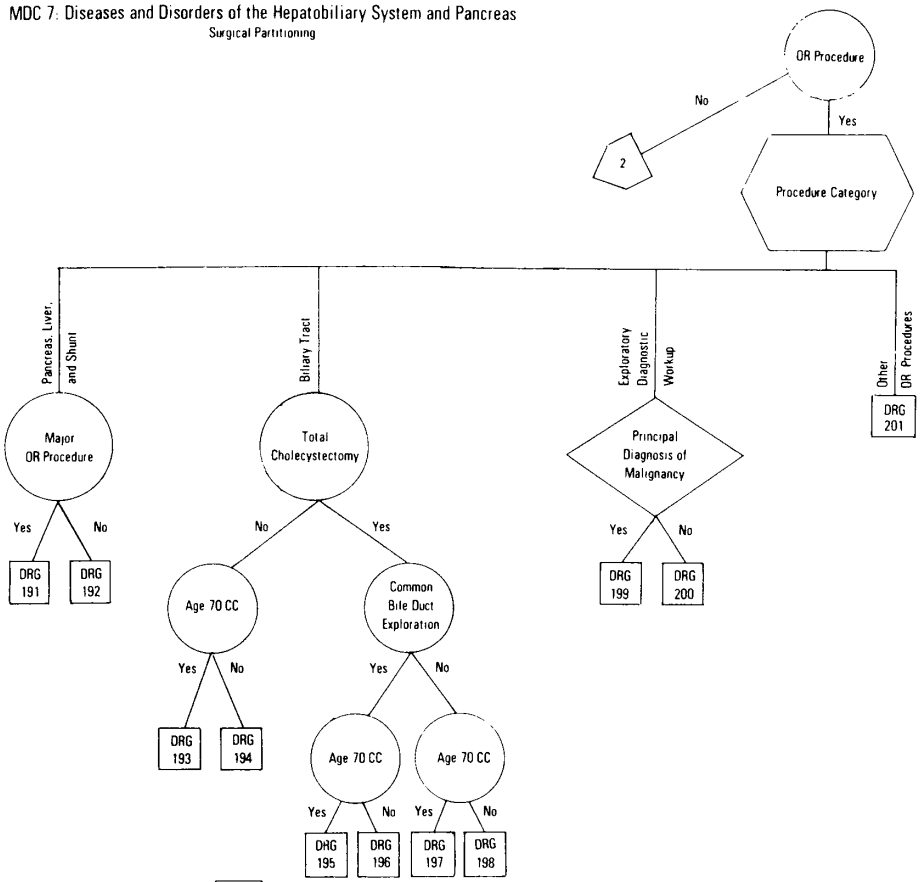
N.B. The computer software sequences the DRGs from right to left as follows: 154 to 156, 146 to 149, 152 and 153, 164 to 167, 159 to 163, 150 and 151, 157 and 158, 168 to 171.

Medical Partitioning

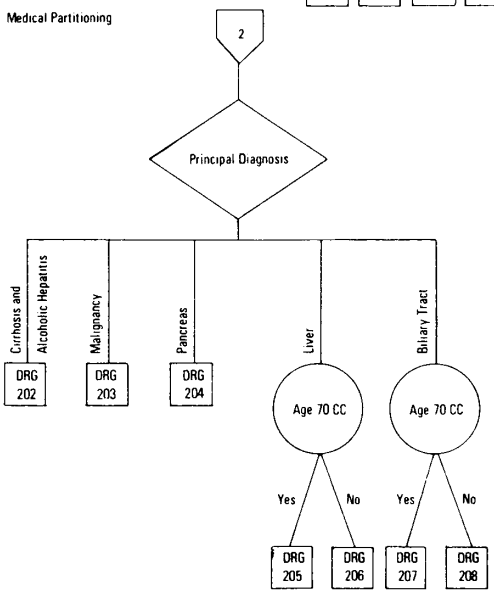


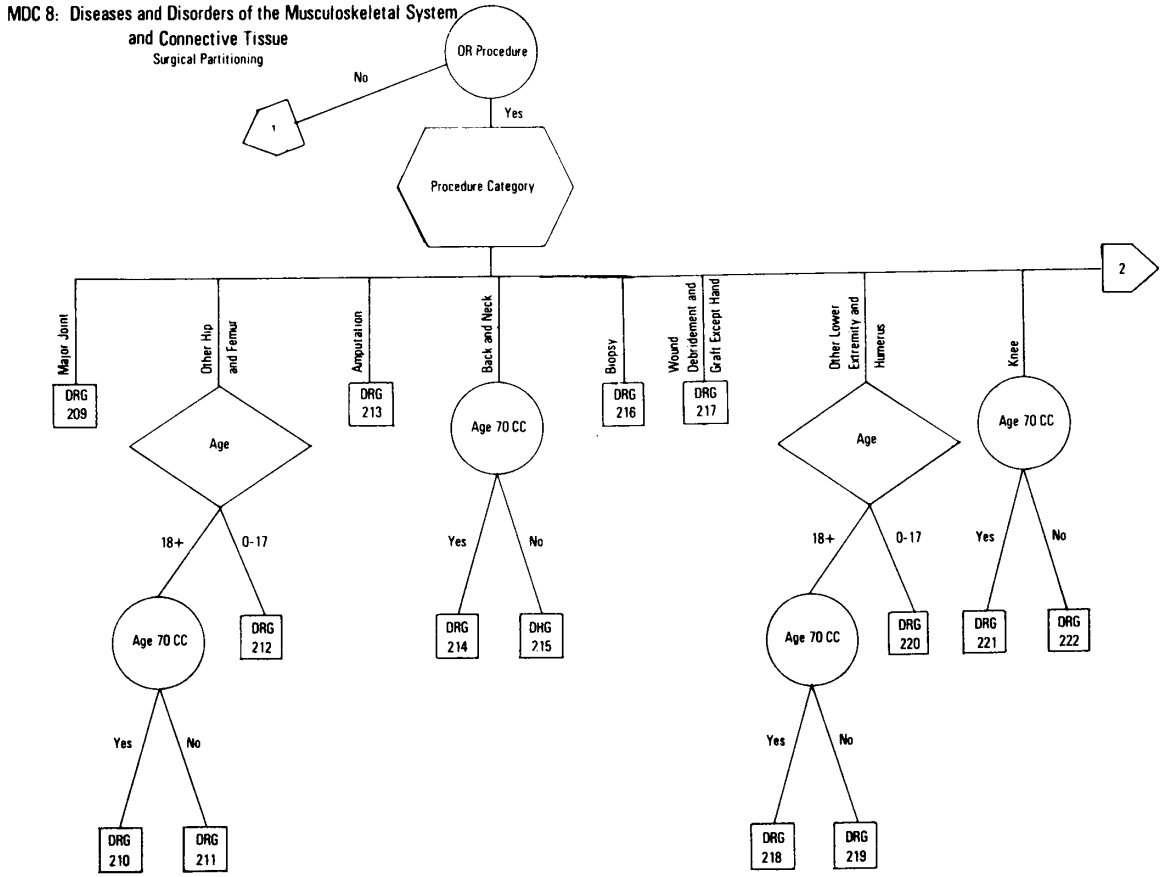
MDC 7: Diseases and Disorders of the Hepatobiliary System and Pancreas

Surgical Partitioning



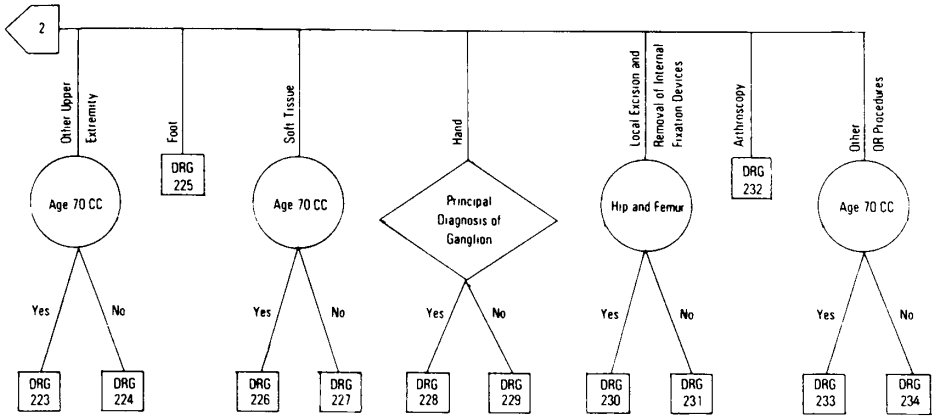
Medical Partitioning




MDC 8: Diseases and Disorders of the Musculoskeletal System and Connective Tissue
Surgical Partitioning
OR Procedure
No
Yes
1
Procedure Category
2
Major Joint
Other Hip and Femur
Amputation
Back and Neck
Biopsy
Wound Debridement and Graft Except Hand
Other Lower Extremity and Humerus
Knee
DRG 209
Age
DRG 213
Age 70 CC
DRG 216
DRG 217
Age
Age 70 CC
18+
0-17
Yes
No
Age 70 CC
DRG 212
DRG 214
DRG 215
DRG 220
DRG 221
DRG 222
Yes
No
Yes
No
DRG 210
DRG 211
DRG 218
DRG 219

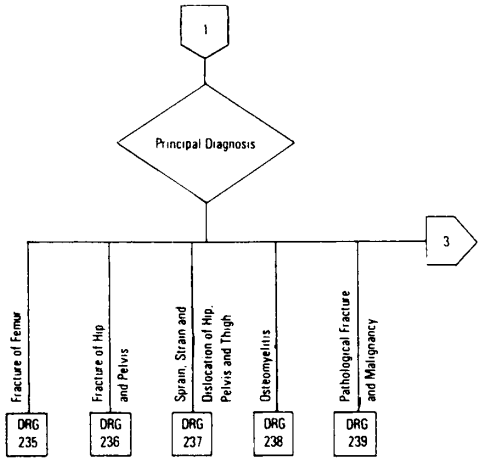
MDC 8

Surgical Partitioning (continued)



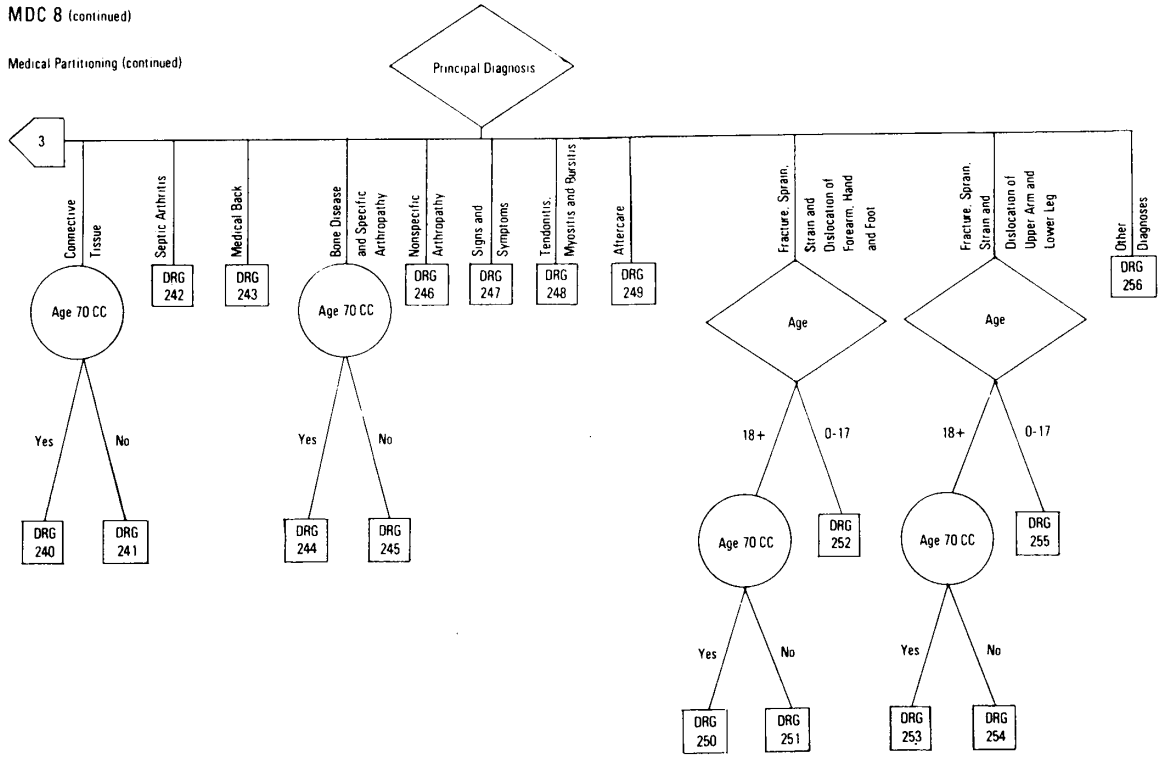
Note: DRGs 230 and 231 immediately precede DRGs 223 and 224 in the official DRG grouper.

Medical Partitioning



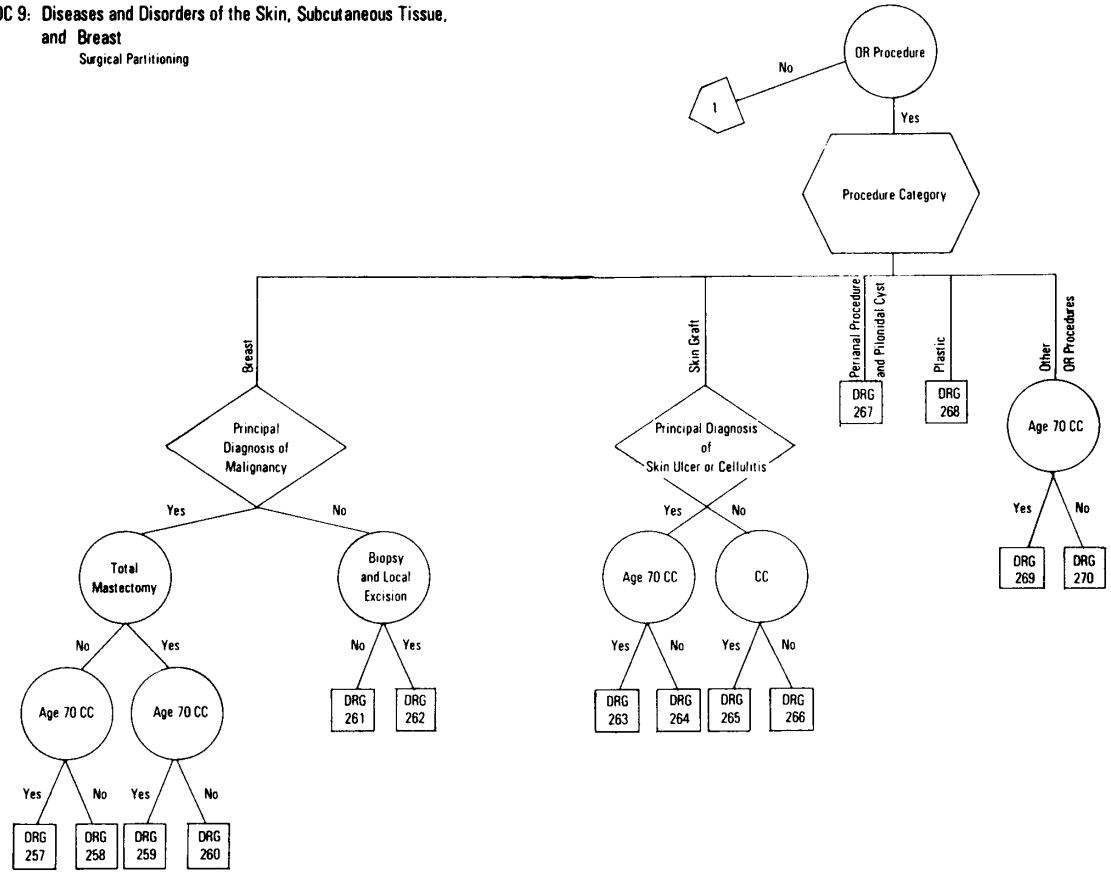
MDC 8 (continued)

Medical Partitioning (continued)



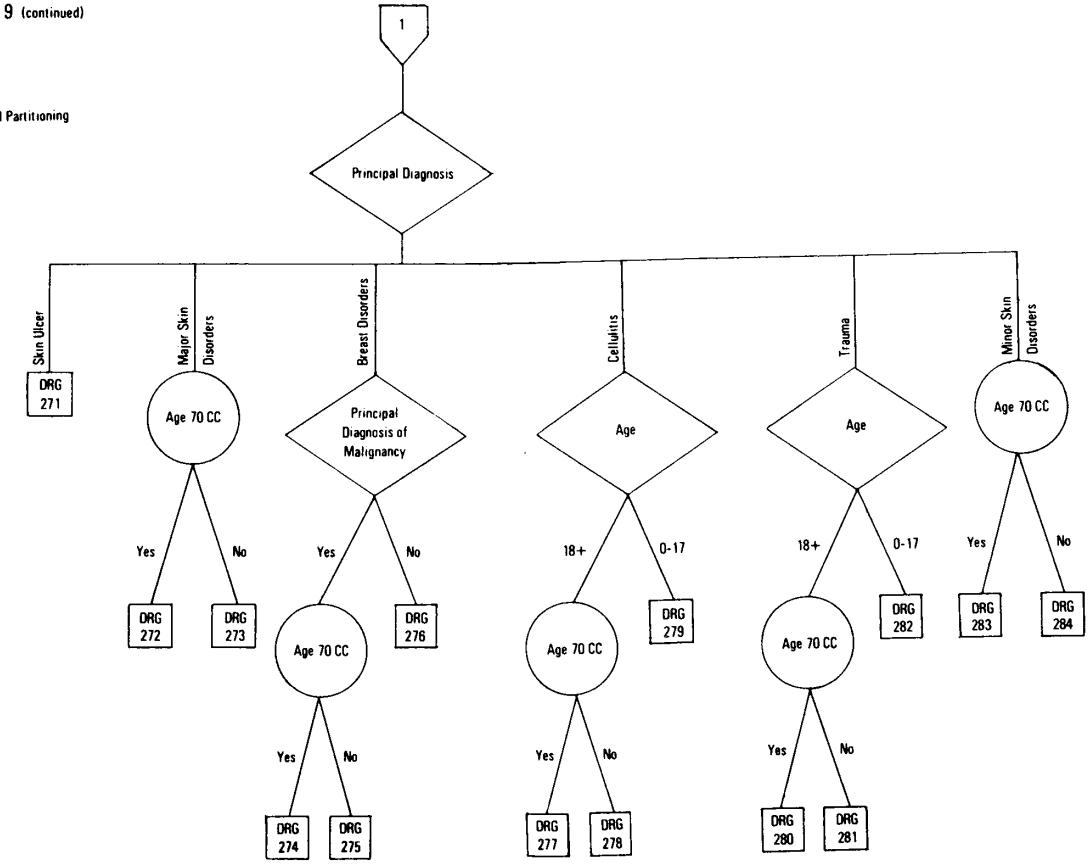
MDC 9: Diseases and Disorders of the Skin, Subcutaneous Tissue, and Breast

Surgical Partitioning

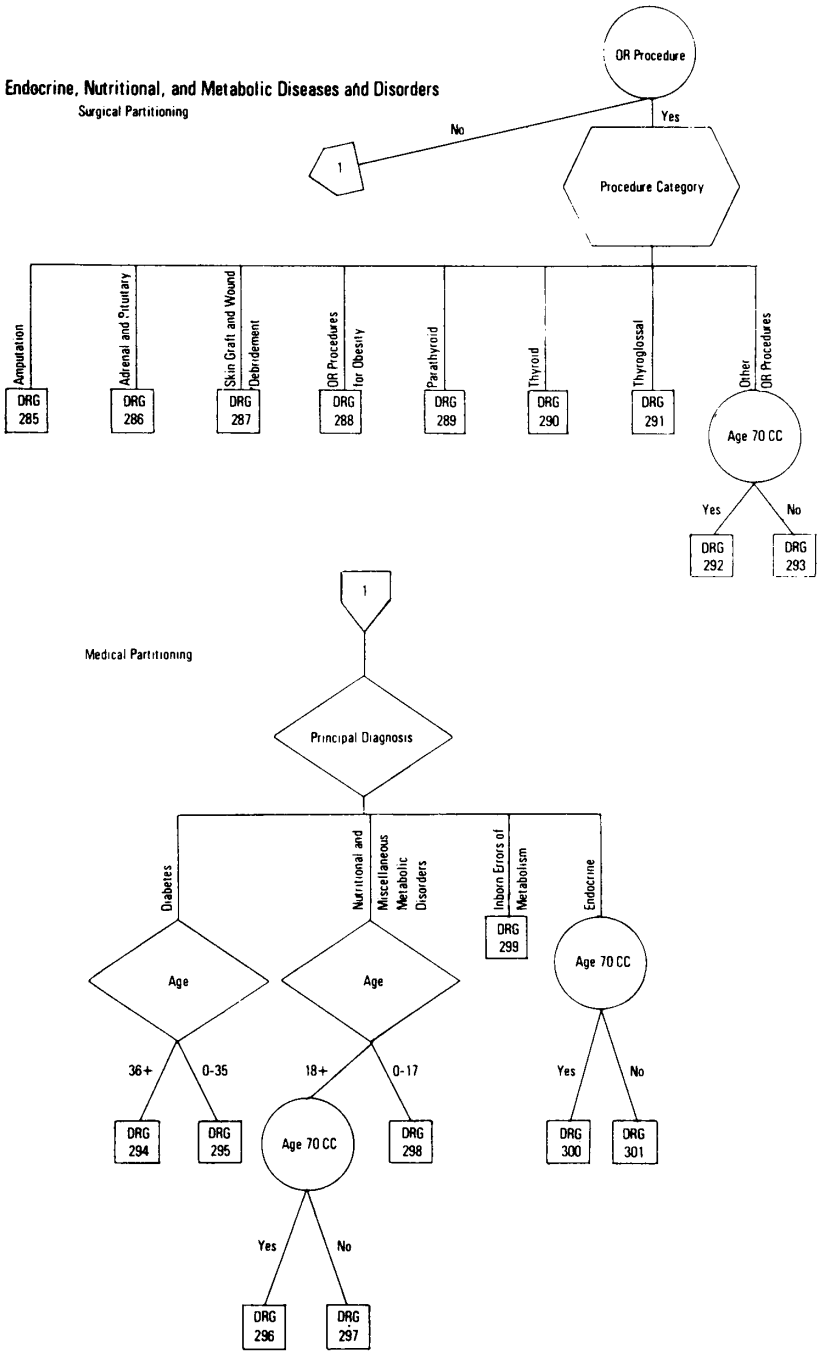


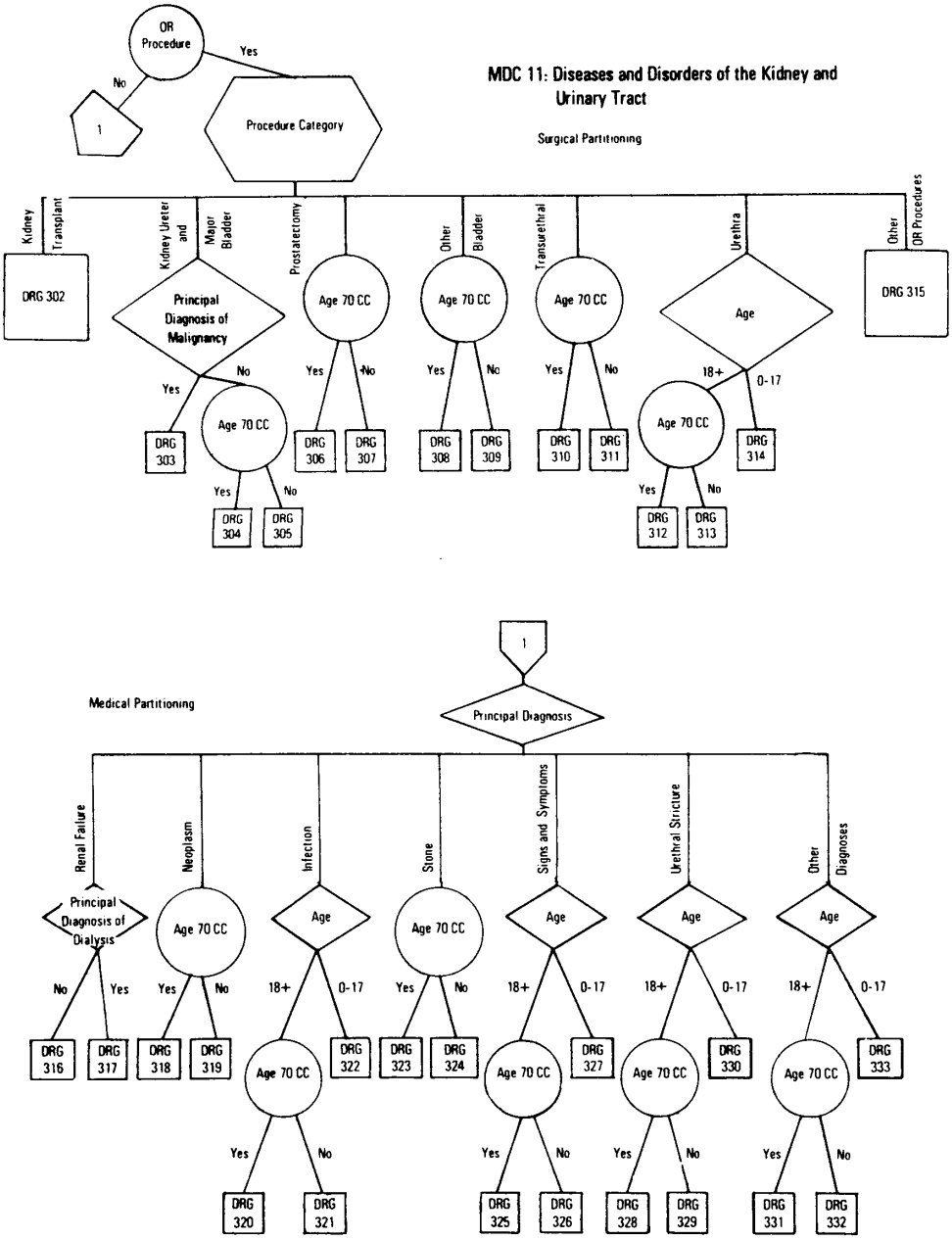
MDC 9 (continued)

Medical Partitioning



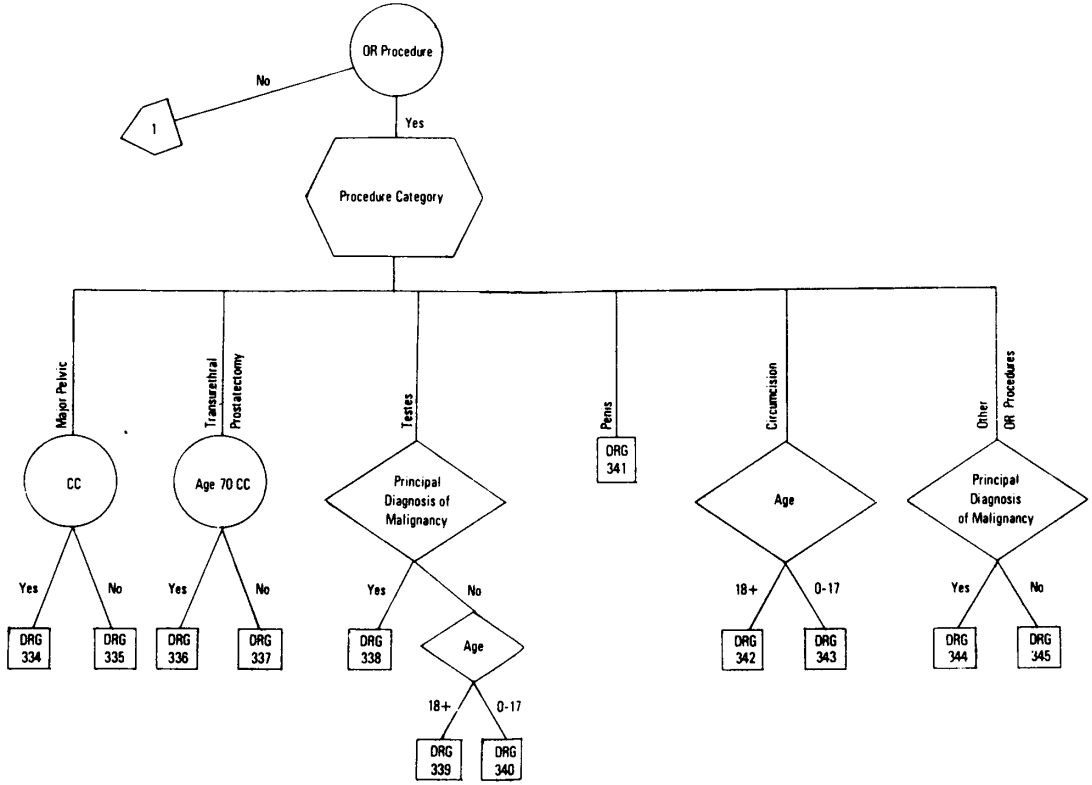
MDC 10: Endocrine, Nutritional, and Metabolic Diseases and Disorders




MDC 11: Diseases and Disorders of the Kidney and Urinary Tract
Surgical Partitioning
OR Procedure
Yes
No
1
Procedure Category
Kidney Transplant
DRG 302
Kidney Ureter and Major Bladder
Principal Diagnosis of Malignancy
Yes
No
DRG 303
Age 70 CC
Yes
No
DRG 304
DRG 305
Prostatectomy
Age 70 CC
Yes
No
DRG 306
DRG 307
Other Bladder
Age 70 CC
Yes
No
DRG 308
DRG 309
Transurethral
Age 70 CC
Yes
No
DRG 310
DRG 311
Urethra
Age
18+
0-17
Age 70 CC
Yes
No
DRG 312
DRG 313
DRG 314
Other OR Procedures
DRG 315
1
Principal Diagnosis
Medical Partitioning
Renal Failure
Principal Diagnosis of Dialysis
No
Yes
DRG 316
DRG 317
Neoplasm
Age 70 CC
Yes
No
DRG 318
DRG 319
Infection
Age
18+
0-17
Age 70 CC
Yes
No
DRG 320
DRG 321
DRG 322
Stone
Age 70 CC
Yes
No
DRG 323
DRG 324
Signs and Symptoms
Age
18+
0-17
Age 70 CC
Yes
No
DRG 325
DRG 326
DRG 327
Urethral Stricture
Age
18+
0-17
Age 70 CC
Yes
No
DRG 328
DRG 329
DRG 330
Other Diagnoses
Age
18+
0-17
Age 70 CC
Yes
No
DRG 331
DRG 332
DRG 333

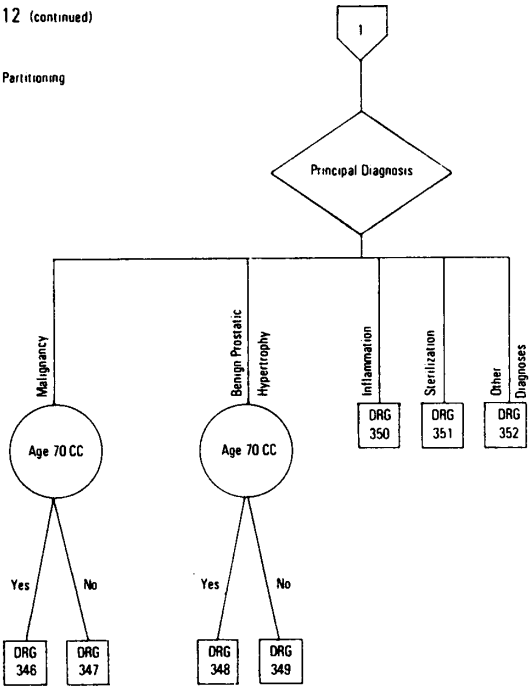
MDC 12: Diseases and Disorders of the Male Reproductive System

Surgical Partitioning



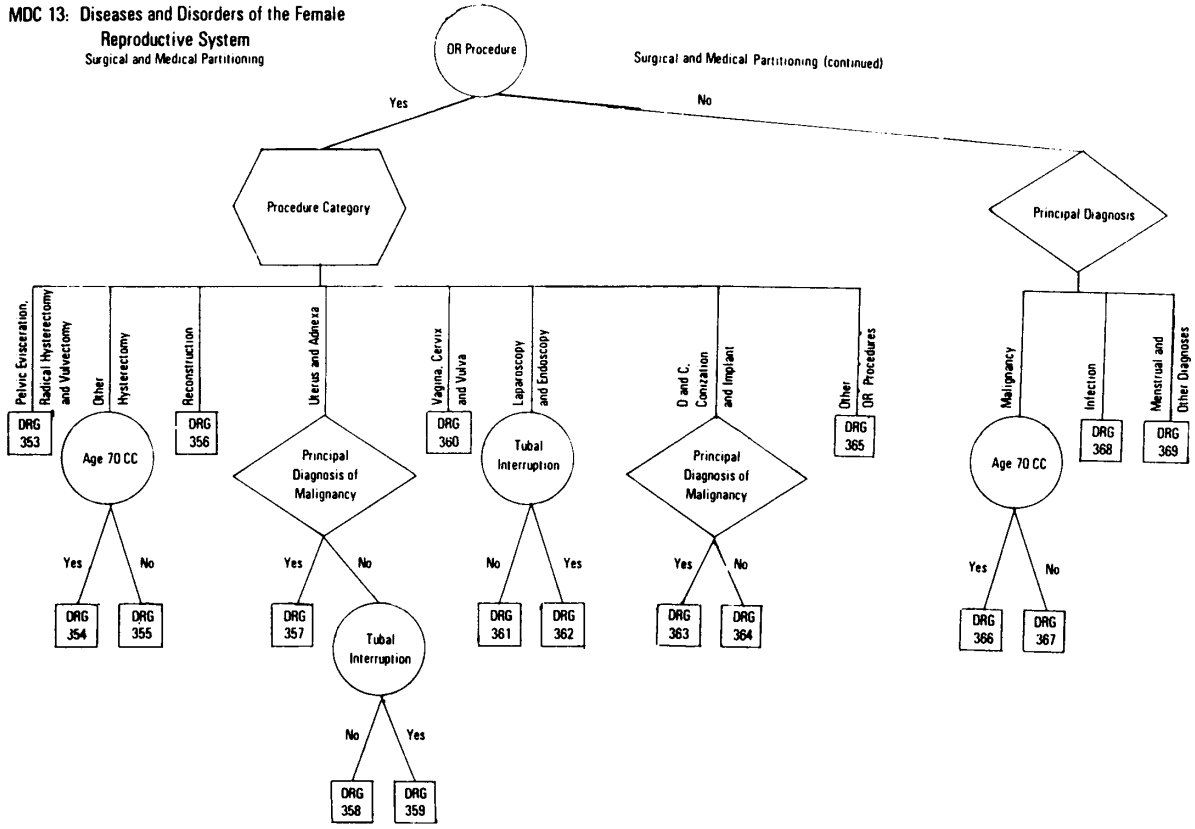
MDC 12 (continued)

Medical Partitioning

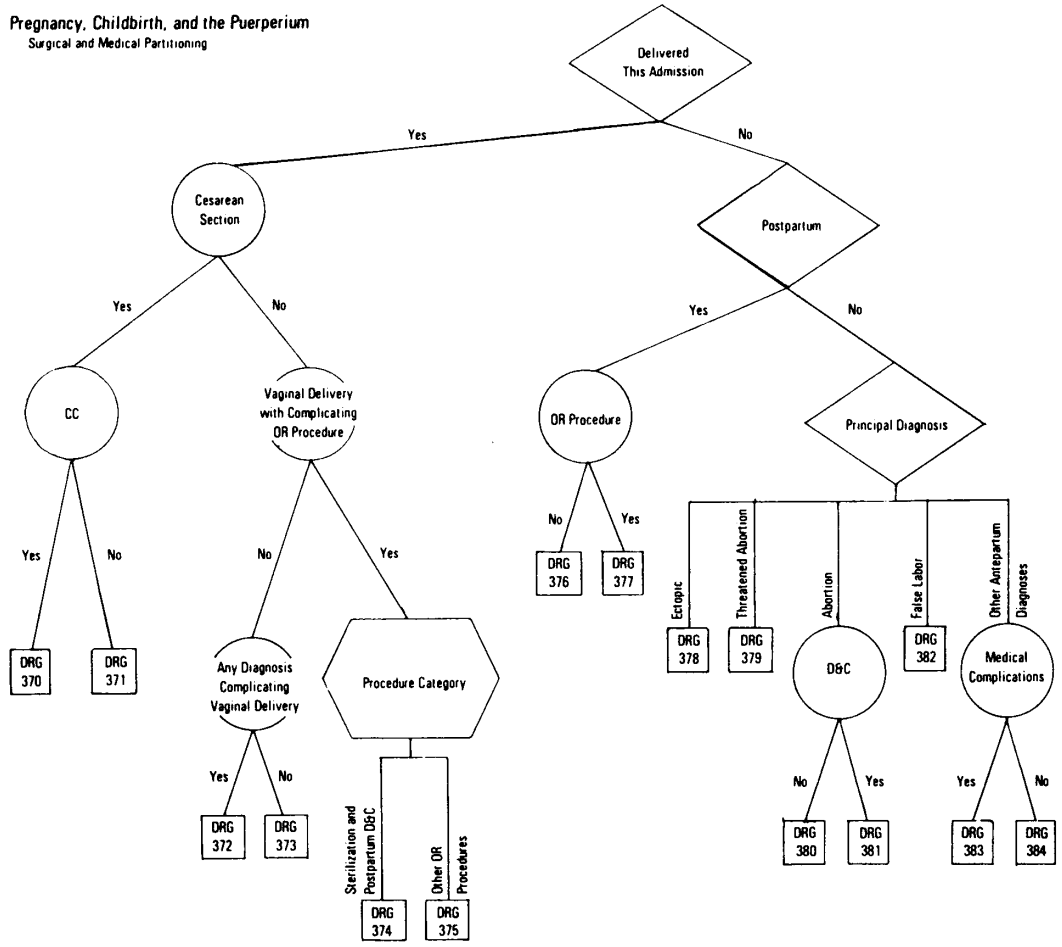


MDC 13: Diseases and Disorders of the Female Reproductive System

Surgical and Medical Partitioning

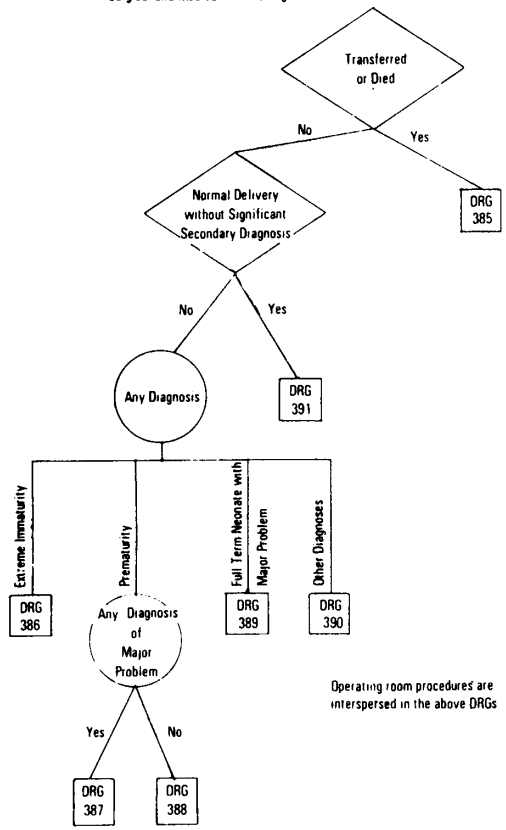


MDC 14: Pregnancy, Childbirth, and the Puerperium
Surgical and Medical Partitioning



MDC 15: Newborns and Other Neonates with Conditions Originating in the Perinatal Period

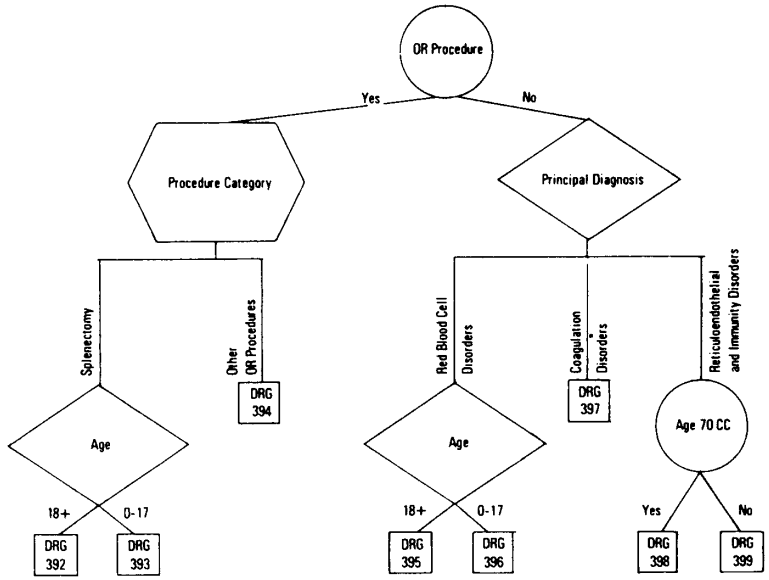
Surgical and Medical Partitioning



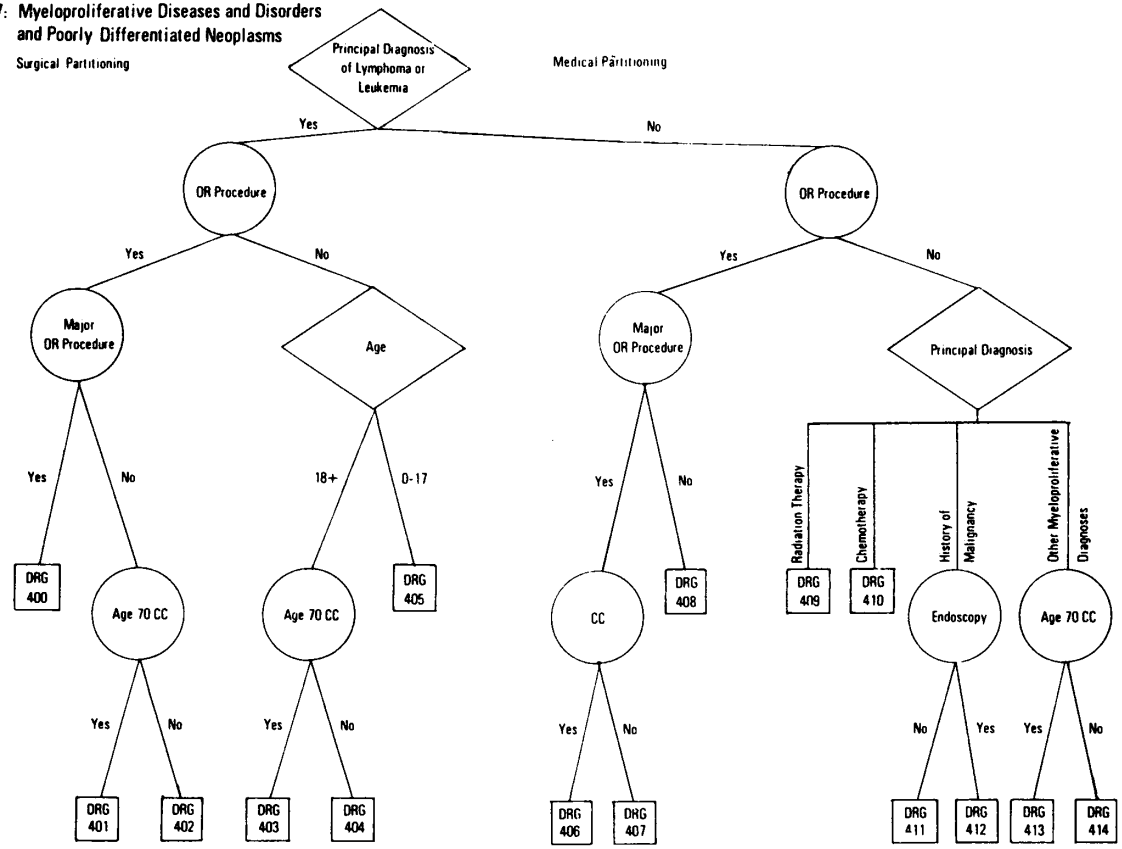
Operating room procedures are interspersed in the above DRGs

MDC 16: Diseases and Disorders of the Blood and Blood-Forming Organs and Immunological Disorders

Surgical and Medical Partitioning

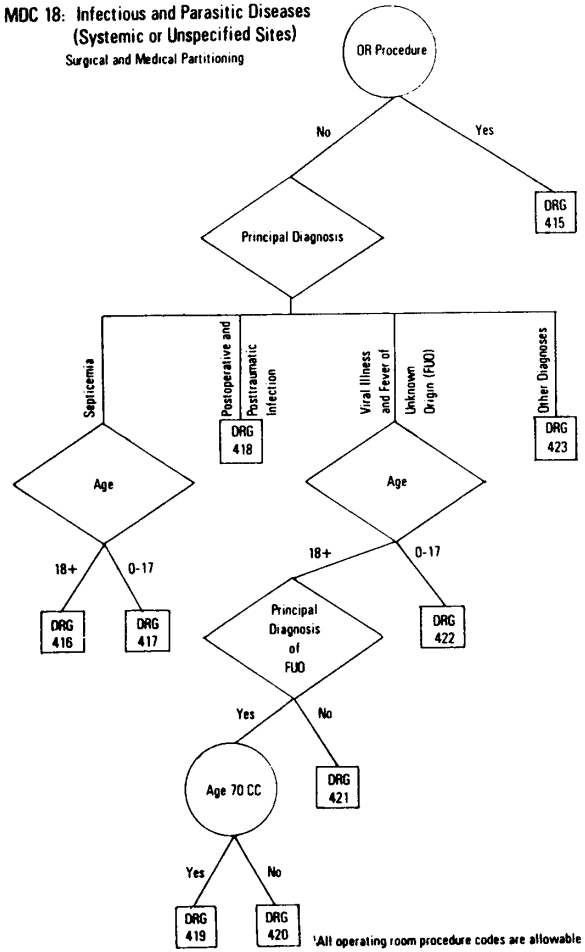


MDC 17: Myeloproliferative Diseases and Disorders and Poorly Differentiated Neoplasms



MDC 18: Infectious and Parasitic Diseases (Systemic or Unspecified Sites)

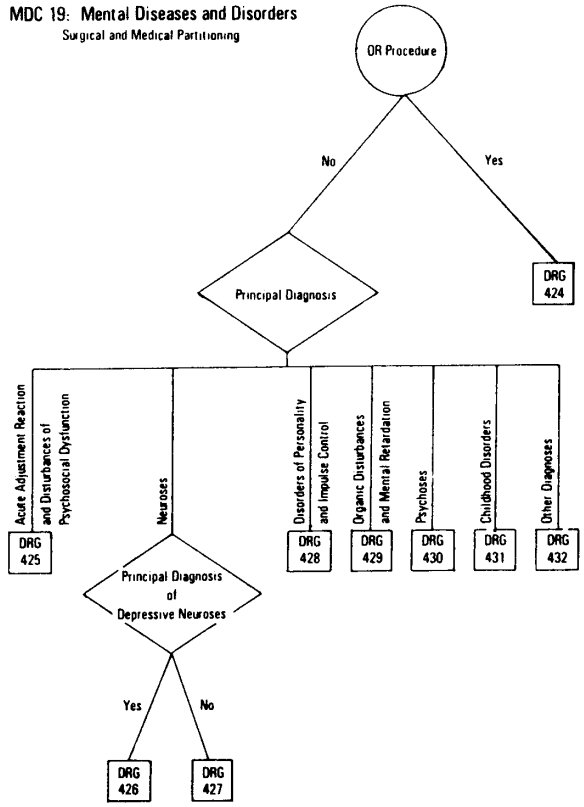
Surgical and Medical Partitioning



[1]All operating room procedure codes are allowable

MDC 19: Mental Diseases and Disorders

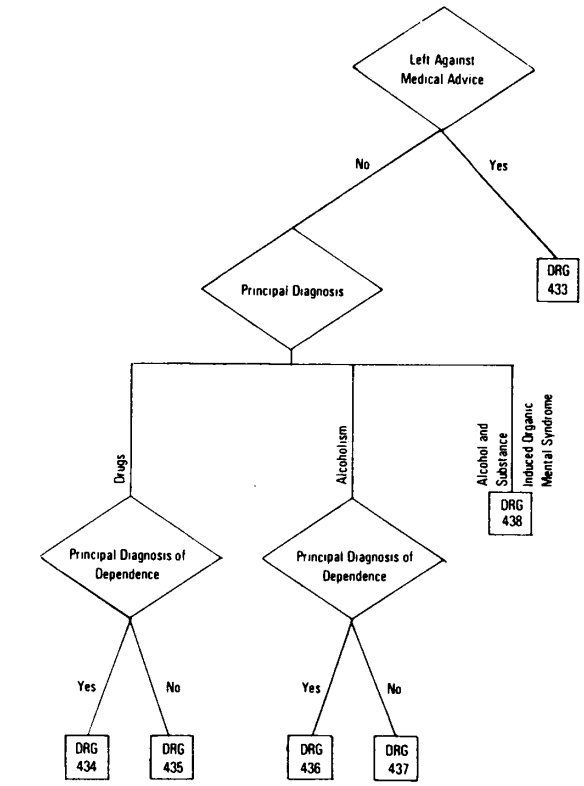
Surgical and Medical Partitioning



[1]All operating room procedure codes are allowable.

MDC 20: Substance Use Disorders and Substance Induced Organic Mental Disorders

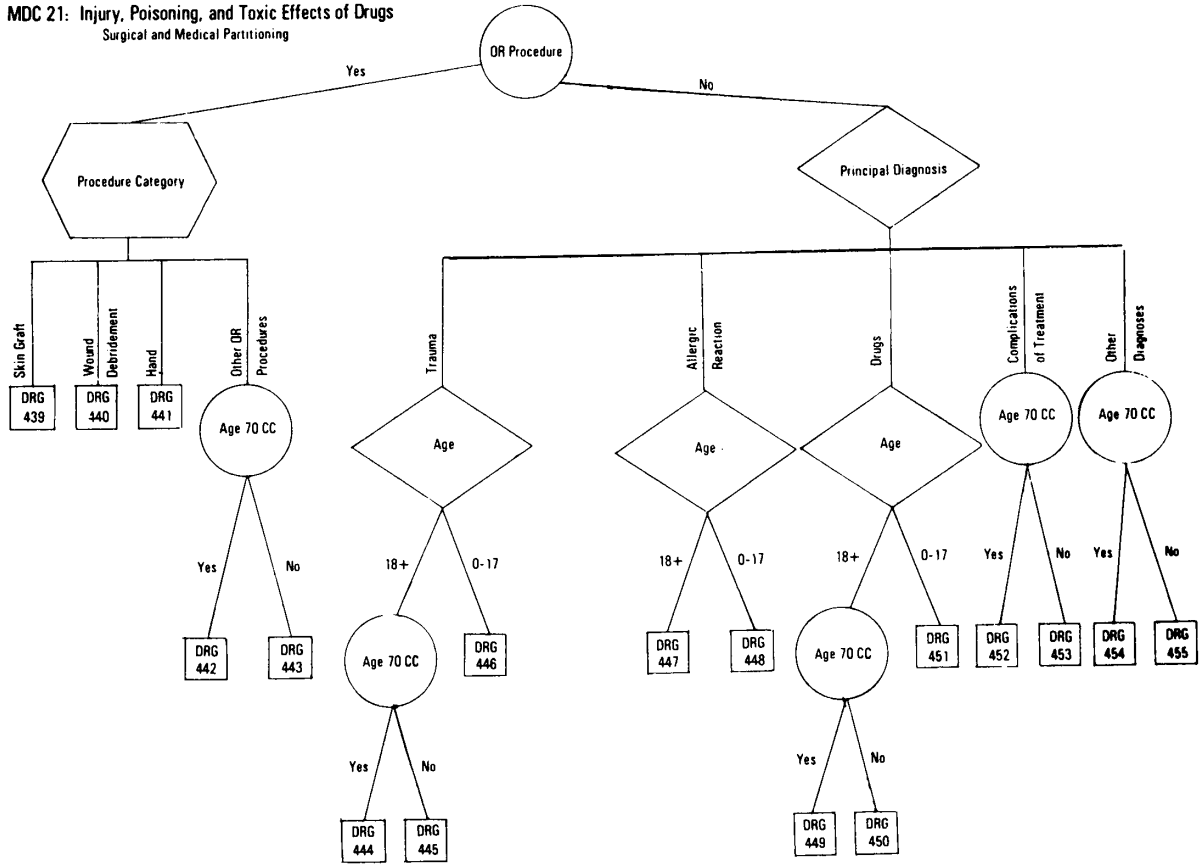
Surgical and Medical Partitioning



Operating room procedures are interspersed in the above DRGs.

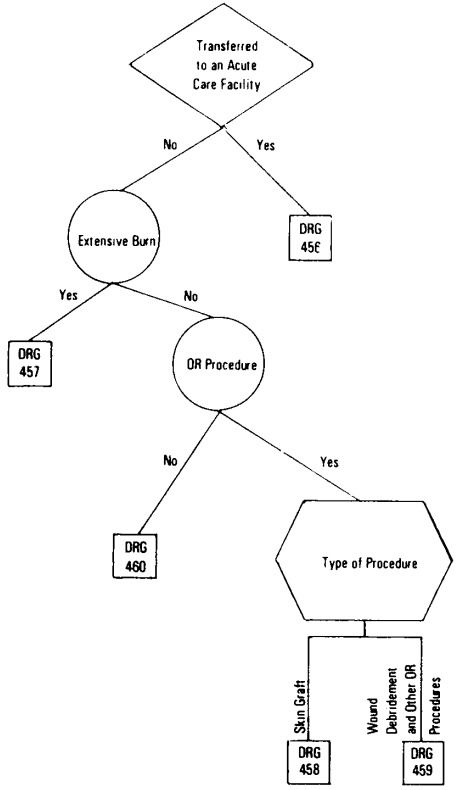
MDC 21: Injury, Poisoning, and Toxic Effects of Drugs

Surgical and Medical Partitioning



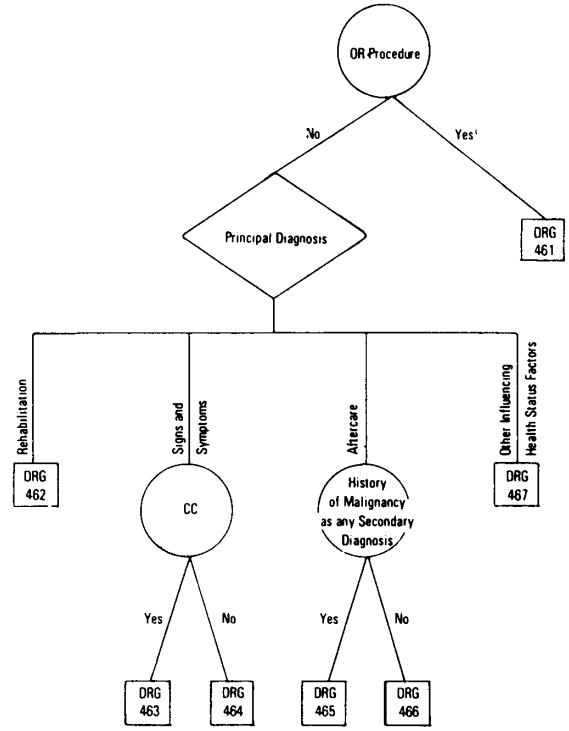
MDC 22: Burns

Surgical and Medical Partitioning



MDC 23: Factors Influencing Health Status and Other Contacts with Health Services

Surgical and Medical Partitioning



All operating room procedure codes are allowable.

Appendix III TECHNICAL GLOSSARY

Activities of daily living (ADL) A disability scale used to measure patient dependency, in particular for nursing long term or chronic cases.

Ambulatory patient group (APG) A classification scheme for outpatients developed by Fetter and Thompson at Yale School of Management in 1984, based on work by Schneeweiss in 1983. It has now been superseded by the more comprehensive ambulatory visit groups (AVGs). The 154 APGs, grouped into major ambulatory categories (MACs), were formed by using an approach similar to the DRG derivation. The dependent variable was contact time with a physician and the groups contained patients with similar disorders who spent a similar time with the doctor. 'Similar' was defined statistically and modified by expert opinion, in the same way as for DRGs.

Ambulatory visit groups (AVG) A classification for outpatients developed by the Health Systems Management Group at Yale in 1986. It is a re-worked version of the APG, designed for the management and the payment of outpatient services. Based on more variables (up to 14 are examined to decide assignment of a case) than an AVG, it uses more measures of resource use in addition to physician contact time. The scheme has two distinct sections: one based on diagnostic categories that can be linked to the DRG scheme, the other identifying administrative reasons for outpatient visits, such as preventative and screening tests, and so on.

APACHE An acute physiology and chronic health evaluation system developed to predict resource use and outcomes of patients in intensive care units. The second version, APACHE II, calculates a score from a combination of twelve physiological variables, age and an evaluation of the health of the patient six months before admission. The method was developed using a panel of clinicians to select and weight the variables used. The scheme has proved successful in predicting the resource use and outcomes of patients admitted to intensive care units.

Autogrp The name given to the interactive software that was used to find homogeneous groups of inpatients (DRGs) and which may be used to look for different groups or subdivisions of DRGs. It incorporates a mutivariable statistical method (automatic interaction detector) to select the variable that provides the most distinct groups, subject to any constraints supplied by the user. It is a quite separate computer program from the grouper software which merely assigns a patient to the already defined DRG.

Automatic interaction detector (AID) A statistical algorithm built into a software package to perform multivariate statistical analysis. It proceeds by diasggregating the data into groups which maximise the between group Euclidean distance (and minimise the within group variability). The user provides the dependent variables, constraints on splits and stopping rules. It is the reverse of cluster analysis; instead of grouping the most similar it separates the most different.

Blue Cross/Blue Shield The 'Blues' run insurance plans covering most of the

employer-provided schemes in the US. The rate of hospital reimbursement is based on a pre-negotiated percentage of the hospital charges.

Budget neutrality A device used by the federal insurance scheme, Medicare, to fix the spending on health care. The relative cost weights remained the same across DRGs, but absolute values are calculated so that the total cost of treating an expected workload equalled the budgeted amount.

*Capitation*Payment per head, usually fixed in advance for a population and independent of services consumed.

Case mix Frequency of patients falling into types according to some predetermined characteristics. These may be social, demographic or severity measures, but are more normally diagnosis, age and treatment. The number of cases in each group can be used to calculate a case mix index.

Case mix 3000 A classification of case types used in a computer package of performance indicators for hospitals in England. It assigns cases to about 3000 groups defined on diagnosis (3-digit ICD-9), sex and seven age-bands.

Central limit theorem A mathematical theorem stating that if random samples are repeatedly taken from a population, the distribution of the mean of the samples approaches the normal distribution. The approximation to normality is sufficiently good for sample sizes of over 30, whatever shape the underlying parent population.

Certificate of need (CON) Before a hospital in the US can be reimbursed for particular services it must be granted a certificate of need from the local health service agency which approves the opening of the service or facility.

Clinical budget A plan of objectives for clinical activity that incorporates detailed resources required to complete the specified level of activity and puts the associated costs into a financial statement. The plan should be agreed by clinicians in conjunction with service providers and finance officers.

Coefficient of variation (CV) The ratio of standard deviation to mean, used as a measure of dispersion. It is sometimes considered that a coefficient of variation should be less than 1 for a homogeneous distribution.

Day case Those cases who attend hospital for investigation, treatment or operation under clinical supervision on a planned non-resident basis and who occupy a bed in a ward, recovery room, or day unit. They can be assigned to a DRG and were included in the data producing US cost weights.

Diagnosis related group (DRG) Groupings of patients that are clinically and resource homogeneous as defined by developers of the scheme at Yale School of Management. Each case belongs to one and only one of the 467 groups. It can be assigned after inspecting the principal diagnosis, main operation, secondary diagnosis, age, sex and disposal from a patient computer abstract.

DRG creep A term applied to the anticipated trend in the US towards changing the way a patient's abstract is recorded in order to maximise the hospital's reimbursement. By re-ordering or adding diagnoses it may appear that patients are suffering from more serious injuries or more complications. A 15 per cent increase has been noted in the recording of secondary diagnoses, but it is not clear whether this reflects more complete

recording of patient characteristics, a more complex case mix or merely an attempt to 'game' the system.

DRG payment Payment for hospital treatment on a fixed scale of DRG payments according to the type of case.

Disease groups A categorisation of five groups of patients introduced in the study of ophthalmology work at the Western Ophthalmic Hospital described in chapter 9. They were used to describe activity and resource use across and within different settings of patient care delivery and cover ophthalmology patients only.

Disease staging A method of describing the extent to which a disease has progressed and so assessing the severity of a patient's condition. The four major stages are:

no complications
problems limited to organ system
multiple site involvement
catastrophic/death

Within these stages a series of substages has been defined on specific clinical criteria developed by a panel of physicians. Stages of a disease have been interpreted in terms of diagnostic codes which enables a computer program to identify disease stages from information on discharge abstracts. Being a clinical measure, staging does not adequately explain differences in resource use.

Global budget Method of setting an overall expenditure limit below which a hospital, or group of hospitals must remain. Used for containing costs, as seen in the fixed amount of resources made available to NHS hospitals which then requires methods for allocation. From October 1983 the Medicare budget was also fixed under prospective payment by DRG.

GROUPER Computer software developed at Yale which assigns inpatients to DRGs. Runs on IBM mainframe and PC machines and is suitable for records coded in ICD-9-CM diagnoses and operations.

Groupes homogenes de malades (GHM). The name given to DRGs by the French project for the medicalisation of information systems (PMSI).

Health Care Financing Association (HCFA) The agency of the federal government responsible for all federally supported health programmes.

Health maintenance organisation (HMO) A health care organisation that acts both as insurer and provider of services by contracting with groups of physicians to provide a range of cover for the population enrolled with the HMO. HMOs expanded sharply in the US in the mid 1980s. Employers have found them attractive as they can provide extensive cover for their staff at fixed cost. Doctors or hospitals with whom the HMOs contract are prepared to negotiate competitive prices to obtain a steady income. Premiums paid by the employer are fixed, within narrow limits to account for age difference, by the government. This suggests that HMOs may have difficulty offering cover for less healthy populations. It is too early to tell whether competitive tendering will lead to selective enrolment in HMOs or, as has been suggested, to poorer outcomes and lower levels of care.

Homogeneity Degree of similarity. Used in case mix classifications to indicate how well the definitions of case types explain differences between patients. A homogeneous group or DRG is one with cases clustered around the

mean and with few extreme cases or outliers. An acceptable level of homogeneity may be defined by a ratio of standard deviation to mean less than 1.

Hospital activity analysis (HAA) A computer database used in the NHS for recording social, demographic and medical data on hospital inpatients. Required nationally and organised regionally, it is an abstract from patient notes containing sufficient variables to assign cases to DRGs. In 1987 it will be slightly modified to provide management information at consultant level more suited to clinical budgeting.

Hospital Inpatient Enquiry (HIPE) A 10 per cent sample of HAA data used to provide a national UK database of hospital inpatient statistics. Published annually up to 1985.

International classification of diseases – 9th revision (ICD-9) The current World Health Organization system of coding diseases, adopted in 1975 and due for revision in the late 1990s. The 4-character code is sometimes extended with a fifth digit or clinical modification (ICD-9-CM). There is also a classification of surgical and other procedures produced as a supplement to the disease classification. Both coding systems were used in DRG definitions.

Length of stay (LOS) Number of nights during inpatient stay, often used as a proxy measure for resource use and used as the dependent variable in DRG derivations.

Longterm care (LTC) Care given to the chronically ill in nursing homes, geriatric wards in hospitals and other longterm care insitutions. Patients falling in this category should not be described by using DRGs. Other classification schemes have been and are being developed to describe case mix in longterm care institutions. These include PDGs and RUGs and have used the activities of daily living as a measure of patient dependency.

Major ambulatory diagnostic category (MADC) The 21 broad categories of disease used in the classification of ambulatory patients or outpatients. MADCs (the forerunners of which were MACs) are closely related to the major disease categories used in DRGs.

Major diagnostic categories (MDC) The 23 broad categories of disease into which the DRGs are grouped. The MDCs cover the complete range of ICD codes and are arranged into categories such as nervous system, respiratory, circulatory, and digestive system disorders.

*Management budgeting*The name of an initiative in the NHS derived from recommendations in the Griffiths enquiry (1983), which proposed improvements in managerial style. Management budgeting is an essentially similar approach to that of clinical budgeting and focuses on the devolution of financial responsibility to smaller administrative units supported by better information. In 1986 DHSS Health Notice (86)34 gave fresh impetus to the initiative, renaming it *resource management*. The new approach is to be piloted in selected acute hospitals, aiming for greater medical and nursing involvement.

Medicaid The US federally funded and administered health insurance scheme which covers hospital and other forms of care for those with low incomes.

Medicare Similar health insurance scheme to Medicaid, but covers people over 65 years old and the disabled, irrespective of income levels. The Part

A insurance covers hospital treatment, while the optional Part B (at extra cost to the insured) covers outpatient and primary care. Those not eligible for Part A may enrol and pay monthly premiums. Medicare patients have been funded by pre-set DRG rates since October 1983 in an attempt to avoid the fast-approaching bankruptcy of the Medicare trust funds. Medicare funding accounted for 40 per cent of all hospital expenditure in the US in 1984.

MEDISGRPS A severity scoring system based on a number of key clinical findings, comparing the admission score to one after ten days can highlight differences in the quality of care.

*New Jersey Department of Health*The state department of health which first used DRGs in a large-scale experiment to set hospital funding by case type treated. Partly as a response to the use of the early version of 383 DRGs from 1980 in New Jersey, the DRG system was reworked to produce the 1982 version containing 467 DRGs.

*Outlier*Extreme cases in DRGs are usually termed outliers. These are the cases which fall beyond pre-defined limits of length of stay or total costs. They may be reimbursed at a lower rate to discourage hospitals keeping cases in a long time and under Medicare reimbursement are subject to reviews before reimbursement is authorised. Each DRG contains a range of patients with a distribution of lengths of stay. The patients falling in the upper and lower tails of the distribution however may be problem cases which should not be considered as belonging to the group. Depending on the focus of interest the outliers may need special analysis or may be unwanted because of their strong influence on characteristics of the DRG's distribution. Trim points need to be fixed to identify outliers as required by the study in hand.

Outpatient An ambulatory patient attending a hospital, for treatment, tests, procedures and so on, but not staying overnight. In the US ambulatory patients include those described in the NHS as day cases who use a bed but do not stay overnight.

Patient dependency group (PDG) A classification of longterm care patients, mainly in nursing homes, developed at Yale School of Management in 1986. The five groups are based on the amount of help patients require for performing the basic activities of daily living. No account is taken of disease, disorders or special treatments. Similar to DRGs, the scheme used samples of data for defining and validating groups which were homogeneous from the point of view of nursing time required per day.

Patient management category, or generalised patient management paths. A classification of hospital inpatients developed by Young and sponsored by Blue Cross of Western Pennsylvania with the object of grouping patients with similar disorders and similar resource consumption. The classification was driven by panels of physicians who considered reason for admission, disease-specific treatment and patient management plans. The approach, unlike DRGs, allowed a separation between the investigative work leading to a diagnosis and subsequent decisions on treatment. Patient management categories have been defined for about 750 treatment patterns, which cover over 90 per cent of the patient mix in an acute hospital.

Patient severity of illness index A four point severity scale, developed by Horn

in 1981 in Johns Hopkins Hospital, which based a patient's scores on seven indicators. The original indicators were:

stage of principal diagnosis
development of complications
interactions between comorbidities
rate of response to therapy
residual response rate
dependency
non operating room procedures

The system was criticised for the subjectivity introduced by the rating scales.

A computerised version is being developed which takes specific clinical findings, to define the levels of severity for about 1,000 disease categories. This version is to be validated against the original index and the resulting severity rating will be expressed as a 6th digit on the disease code.

*Pass-through*Lump sum payment made by Medicare to hospitals to cover expenditure not included in DRG cost weights. These were principally for teaching costs and capital expenditure.

Peer review organisation (PRO) Groups of physicians organised regionally in the US after the introduction of prospective payment rates to perform a mandatory utilisation review of all cases funded by Medicare. The PROs review both quality of care and appropriateness of admissions, including a detailed check of re-admissions and outliers. In 1984 PROs found 3.55 of admissions to be inappropriate and disallowed payment. PROs replaced the professional standards review organisations.

Per case payment Fixed method of payment based on average cost which does not allow for volume effects or severity of case.

Per diem payment Fixed daily payment per case. Otherwise, the same as per case payment.

Performance indicator (PI) A measure used by health service managers to study levels of activity and costs. It may be used to compare performance. Indicators include percentage of occupied beds, throughput per available bed, staff per available bed, cost per case, and so on.

Preferred provider organisation (PPO) A provider-sponsored organisation (as distinct from the payer and provider role of HMOs) in which doctors or hospitals make an agreement with patients and their insurers to provide health care at a discounted price. The patients must use the 'preferred providers'.

Professional standards review organisation (PSRO) Locally organised groups of doctors required under 1972 Social Security Amendments in the US to monitor the quality and appropriateness of health care provided under Medicare and Medicaid. PSROs have been superceded by PROs.

Prospective payment assessment commission (ProPAC) An independent review body which makes annual recommendations to the federal government on aspects of prospective payment. These include modifications to DRG definitions in the light of changing treatment styles, or technology, in order to maintain homogeneity within groups and adjustments to cost weights. It may also be recommended that relative cost weights be changed

in response to new treatments, and the total cost of Medicare is uplifted by an inflation factor appropriate to the health care sector. A reduction in this factor has usually been suggested for efficiency savings.

Prospective payment scheme (PPS) A scheme, such as that used for reimbursing hospitals for treating Medicare patients, which pays fixed rates for each case type treated. The hospital is at financial risk as costs over the payment rate are not usually reimbursed and have to be met from the cases whose treatment costs less than the rate received.

Quality adjusted life year (QALY) A measurement of life expectancy which is adjusted according to quality of life. In this way a treatment which offers a year of excellent health may be equivalent to ten years of poor health or impairment. The concept has been developed by health economists and relies on the general acceptance and reproducibility of scoring different degrees of well-being.

Reduction in variance (RIV) A statistical measure used to indicate the usefulness of grouping items to explain the underlying variations between them. The reduction in variance compares the variance within groups with the variance of all cases as a single group. In the context of hospital admissions, about 40 per cent of the variability in length of stay for a large sample of cases is explained by assigning them to DRGs, that is an improvement on treating all cases as belonging to the same group which is highly significant.

Resource utilisation group (RUG) An earlier version of the patient dependence groups used to describe types of case in longterm care. Based on activities of daily living such as dressing, feeding and mobility, the five RUGs were improved into the current PDGs. Apart from the RUGs defined at Yale, there are other slightly different resource utilisation groups in use for longterm care.

Standard hospital accounting and rate evaluation (SHARE) The system of reimbursing hospitals in New Jersey used immediately before prospective payment by DRG was introduced. Per diem rates were set which allowed for different types of hospitals and different types of patients, but did not adequately account for patient mix.

Tax equity and fiscal responsibility act of 1982 (TEFRA) The public law enacted in the US which required case mix to be incorporated in Medicare reimbursement rates. More detailed legislation in the following year (the revised section 223 limits) laid down the way DRGs would be used to set rates, the definition of outliers, means of payment for teaching costs and capital expenditure, and the review bodies to monitor the scheme.

Trimming A means of removing extreme cases or outliers from a group. This may be desirable to identify extreme cases for review, remove bad data, or to provide a better description of group characteristics.

In cost containment schemes, trim points have been established beyond which a different reimbursement rate may be applied. This is usually a reduced per diem rate or is restricted to a limited percentage of admissions.

It is possible that some patients appearing as outliers have bad data and should not belong to that DRG. These might be cases brought in for surgery which did not take place, or patients with errors in admission or discharge dates. Such cases can be removed by trimming.

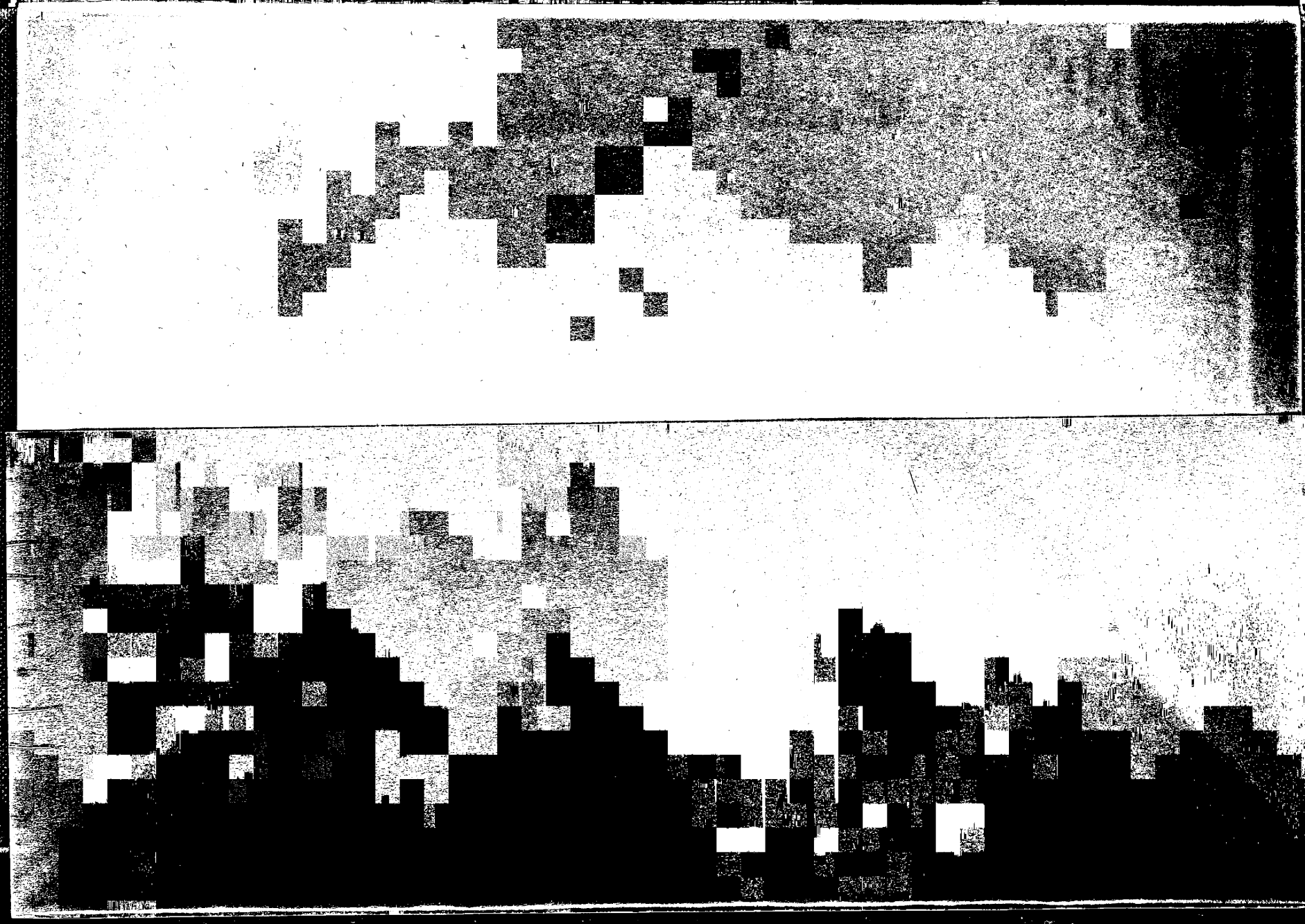
If the underlying characteristics of a group are required extreme cases *must* be trimmed. Often the removal of a small percentage (two–four per cent) of cases has the effect of halving both mean and standard deviation of the length of stay distribution. The object of trimming is to remove the smallest number of cases to achieve the greatest refinement of the mean and standard deviation. Although the median or the mode of the distribution might be used, the disadvantages are that the median cannot be used in calculations of resource use, and the mode might be an extreme point, such as zero day's stay.

Trim points The selection of trim points to identify outliers in DRGs can be done in a number of ways. They may be expressed in terms of cost, but more normally in days of hospital stay. The simplest method is to pick a number, such as 100 days, but this is somewhat crude. A statistical approach often used is to transform the length of stay distribution to a symmetric or approximately normal distribution (by taking logarithms of stay length) and define trim points, or cut-off points, as two standard deviations either side of the mean.

There are variations to this, such as calling day cases a ½ day stay, or widening the trim points to three standard deviations, or only calculating a high trim-point. Yale researchers proposed a two stage trim which first used a non-parametic measure to find the range into which most cases fell and then applied the parametic approach described above to the cases falling within that range.

Utilisation review (UR) One of the activities which DRGs were formed to help. Utilisation review is formalised into the work of professional standards review organisations in the US, who check that the care given and the associated treatment costs are reasonable and necessary. DRGs allow comparisons to be made between individual doctors or hospitals, and between treatment styles and costs for the same type of case.

INDEX



THE CONTRIBUTORS

Martin Bardsley Research analyst, CASPE Research

Gwyn Bevan Senior Lecturer in Health Economics, Department of Community Medicine, UMDS, St Thomas' Hospital

James Coles Associate Director, CASPE Research, and Fellow in Case Mix Accounting, King's Fund College

Robert B Fetter Professor of Health Care Management, Yale School of Organisation and Management

Linda Jenkins DRG Project Leader, CASPE Research

Laurence F McMahon Assistant Professor, Department of Internal Medicine, University of Michigan Medical Center

Antoinette B Newman Financial planner, Paddington and North Kensington HA

Jean-Marie Rodrigues Professor of Public Health, University of St Etienne

Hugh Sanderson Specialist in community medicine, Wessex RHA

Iden Wickings Director of CASPE Research and Deputy Secretary, King Edward's Hospital Fund for London