PRODUCTIVITY COMMISSION

Techniques for Measuring Efficiency in Health Services

Staff Working Paper

Stuart Peacock
Chris Chan
Melvino Mangolini
Dale Johansen

July 2001

Inquiries about this staff working paper:

Media and Publications
Productivity Commission
PO Box 80
Belconnen ACT 2616

Tel: (03) 9653 2244
Fax: (03) 9653 2303
Email: dhanek@pc.gov.au

General Inquiries:

Tel: (03) 9653 2100 or
(02) 6240 3200

Citation, with permission from the author(s), should read:

Peacock, S., Chan, C., Mangolini, M. and Johansen, D. 2001, *Techniques for Measuring Efficiency in Health Services*, Productivity Commission Staff Working Paper, July.

Information on the Productivity Commission, its publications and its current work program can be found on the World Wide Web at http://www.pc.gov.au or by contacting Media and Publications on (03) 9653 2244.

Contents

BOXES

FIGURES

TABLES

Preface

Health care expenditure represents a major use of our nation's resources and has been growing rapidly. Factors such as the ageing population, the increased personal use of health care, and medical advances that have opened the way for more treatment options and diagnostics have contributed to a rise in the demand for health care.

Increased pressures on health care resources have led policy makers, administrators, and clinicians to search for more efficient ways to deliver health services. Efficiency improvements in the health sector, even in small amounts, can yield considerable savings of resources or expansion of services for the community.

Efficiency measures can be a useful tool for health planning and policy evaluation. Ross's (1995) survey results suggest that there is a high level of awareness of efficiency evaluation among the sample of Australian government officials who had responsibility for health expenditure decisions or policy advice. However, a lack of expertise in evaluation techniques and dearth of data were found to be major barriers to the use of efficiency measures, hence impeding the management of health service resources with a proper economic perspective.

The paper aims to fill this information gap. It is a précis of techniques and applications for measuring health care efficiency. Given its focus on methodological issues, the paper complements the ongoing work at the Commission and by others in developing and analysing performance data on health services in Australia.

The paper was researched and written by Dr Stuart Peacock, Chris Chan, Melvino Mangolini and Dale Johansen. The Commission contracted Dr Peacock at the Centre for Health Program Evaluation, Monash University. *The views expressed in this paper are those of the authors and do not necessarily reflect those of the Commission.*

The authors wish to thank Mrs Helen Owens for guiding the initial development of this project. Acknowledgments are due to Prof. Knox Lovell, Dr Lynne Williams, Dr Robyn Sheen, Penny Taylor, Ilias Mastoris, Dr Patrick Jomini and Paul Gretton for their helpful comments on a draft of the paper. Wendy Maloney and Janet Savvides were of great assistance with typesetting the paper.

Key points

- Cost-effective health care is an important objective for all governments. Compared to efficiency analysis in other industries, however, measuring efficiency in the provision of health services is complicated by the complex nature of health processes and the production of health services, as well as the characteristics of the health care market.
- A proper economic perspective requires assessing health care efficiency in terms of health outcomes. Further, such an analysis should account for the impact on health of other factors such as education or living standards.
- This study appraises two types of measurement techniques — economic evaluation which compares health programs; and benchmarking analysis which compares service providers. These techniques are complementary in identifying potential efficiency improvements in the health sector.
  - Economic evaluation focuses on measuring health outcomes for a range of health conditions. The last decade has seen the development of composite outcome measures incorporating fatal and nonfatal conditions into the measurement of health outcomes. This has expanded the scope for comparing different health programs for their cost and effectiveness.
  - Data envelopment analysis is the most useful benchmarking technique because of its adaptability to the joint-production nature of health services and to the influence of external factors on productive efficiency. It extends the frequently used simple ratio analysis by producing more informative efficiency measures than partial productivity indicators.
- New methods have been developed to measure the efficiency of service providers in terms of service outcomes. The use of these methods depends on the availability of outcome data, and may require some adjustment in the information systems used in the health sector.
- The World Health Organisation and the National Health Performance Committee in Australia have adopted different approaches to conducting comprehensive comparisons of efficiency in the health sector. This reflects differences in perceptions of various methodological issues and solutions to these issues.

Glossary

The following terms are defined in the context of this report:

Acute care	A health care process in which diagnostic, therapeutic or surgical procedures are performed to relieve symptoms and reduce the severity of illness or injury.
Allocative efficiency	Situation in which health service resources are put to their best possible use so that no further improvements in the health status of the community as a whole are possible.
Ambulatory care	Health services provided by an acute care hospital to non-admitted patients.
Asymmetric information	Situation in which the service provider and the patient have differing amounts of information about the patient's health, treatment options available, and the expected consequences of treatment.
Benchmarking analysis	The comparison of service providers against a benchmark or ideal level of performance chosen on the basis of performance over time or across a sample of comparable service providers, or an externally set standard.
Casemix	The distribution of inpatient cases treated by a hospital as classified by patient illness characteristics and treatment processes; often described with reference to diagnosis related groups.
Case complexity	The intensity or costliness of resource use in the provision of health services.
Comorbidity	The simultaneous occurrence of two or more diseases or health problems that affect the care of a patient.
Cost-effectiveness	Technical efficiency determined from comparisons of a specific set of health programs using economic evaluation techniques. See 'technical efficiency'.

Cost–benefit analysis (CBA)	A method of estimating the net benefit of a health program — that is, total benefit less total cost, with all benefits and costs measured in monetary value.
Cost–effectiveness analysis (CEA)	A method of comparing a health program to its alternative(s) based on the ratio of total incremental cost to total incremental benefit, with all benefits measured in a natural unit such as the number of disease cases treated or the number of years of life saved.
Cost minimisation analysis (CMA)	A method of comparing the total net cost of alternative health programs that yield the same health related benefits.
Cost-of-illness study	A method of estimating the total cost incurred by the community due to a specific disease or health condition.
Cost–utility analysis (CUA)	A method of comparing a health program to its alternative(s) based on the ratio of total incremental cost to total increment benefit, with all benefits measured in a health related wellbeing unit such as the mortality rate or life expectancy adjusted or not for differences in the functional capability or quality of life.
Data envelopment analysis (DEA)	A programming technique for estimating the production frontier and measuring the relative level of efficiency based on a measure of distance to the frontier.
Diagnosis related group (DRG)	A set of patient case types established under a disease and injury classification system that identifies patients with similar illness conditions and processes of care (with each of the DRGs assigned a weighting factor that compares its costliness to the average for all DRGs under a casemix funding scheme).
Direct health care costs	The money value of health care resources consumed in the provision of a health program and in dealing with the side effects or other current and future consequences associated with it, including the costs of diagnostic tests, drugs, supplies, personnel, and medical facilities.
Disability adjusted life expectancy (DALE)	A measure of health related wellbeing, which is the life expectance at birth adjusted for the prevalence of disability using a set of weights that reflect the varying degrees of reduction in functional capacity due to various diseases at each age.

Discounting	The process of converting future costs and benefits to their present values.
Economic evaluation	The comparative analysis of alternative health programs in terms of their costs and benefits.
Effectiveness	The extent to which health programs achieve health improvements in real settings.
Efficacy	The extent to which health programs would achieve health improvements under ideal settings.
Elasticity	Percentage change in a dependent variable resulting from one percent change in an independent variable. Elasticity greater (less) than one in absolute value indicates an elastic (inelastic) relationship.
Frontier analysis	A statistical method of estimating the level of efficiency for service providers relative to the best possible production practice, with inefficiency interpreted as a departure from the best possible production practice or frontier.
Functional ability	An individual's ability to perform roles, tasks or activities, such as the ability to work, enjoy recreational pursuits, and perform normal daily living activities. Functional ability consists in physical, emotional, mental and social aspects.
Health, health condition, health state, health status	The physical and emotional wellbeing of an individual or a defined population at any particular point in time, which may be modified by impairments in functional ability, physical, emotional and mental wellbeing and social functioning, that are influenced by the occurrence of diseases or injuries.
Health care	The provision of goods and services to maintain and improve health or alleviate symptoms of illnesses.
Health care intervention, health program	A set of health services provided for the purposes of health promotion, prevention of diseases, establishing diagnosis, curing diseases, rehabilitation, or palliative care.
Health related wellbeing (or utility)	Satisfaction associated with a particular health condition.

Health sector	A classification of economic activities covering the businesses of hospitals, nursing homes, general and specialist medical services, and other health services in the private and public sectors.
Health technology	The use of drugs, devices and medical and surgical procedures for health care, and the organisations and supportive systems within which such care is provided.
Illness severity	One dimension of the health status that refers to a patient's condition such as the intensity of disease manifestation, the degree of dependency, and the response to therapy.
Incremental cost	The cost of one alternative less the cost of another.
Indirect health costs	A term used in health economics to refer to changes in productivity resulting from illnesses or deaths, including productivity gains or losses resulting from patient time spent in treatment at health care facilities. (This definition should not be confused with accountants' use of indirect costs, which refers to overhead and fixed costs.)
Indirect non-health costs	The money value of resources consumed outside the health sector in the provision of a health program or intervention, and in dealing with the side effects or other current and future consequences associated with it, including welfare services, housing services, police, courts, and education.
Intervention	See 'health care intervention, health program'.
Life expectancy at birth	The average number of life years that a newly born member of the population could expect to live.
Morbidity rate	A measure of the incidence of diseases or illnesses in a particular population.
Mortality rate	A measure of the incidence of death due to diseases and injuries in a particular population.
Net present value	The value of a stream of net benefits to be received in future, discounted to the equivalent of present dollars.
Opportunity cost	The value of the best alternative foregone in order to obtain or produce more of the health services under consideration.

Outcome	An expression of health services in terms of their consequences on the health status of service recipients, such as changes in life expectancy and changes in the rate of disease incidence.
Output	An expression of health services in terms of service units that are often used as the base for calculating service costs or charges, such as the numbers of doctor visits, patient days, patient cases, surgeries performed and hospital discharges.
Perspective	The viewpoint from which an efficiency analysis is undertaken, which determines what consequences and costs of health services are considered and what decision rules are applied in comparing efficiency.
Production frontier (or function)	The conceptual or statistical relationship between the maximum amounts of a defined final product (either output or outcome) that can be obtained corresponding to any combination of inputs.
Productivity cost	A term used in health economics to refer to the cost associated with lost or impaired ability to work or engage in leisure activity due to illness or death.
Quality adjusted life years (QALYs)	A measure of health related wellbeing, which assigns a set of weighting factors between zero and one to reflect the impact of particular health states on the quality of life — a weight of one assigned for full health and a zero weight assigned for a health state equivalent to death.
Quality of care	Aspects of health services that affect the process of care and its consequences on patient health condition.
Quality of life	A concept reflecting subjective or objective judgement concerning all aspects of an individual's existence, including health, economic, political, environmental, cultural, aesthetic, and spiritual aspects.
Randomised controlled trial	A clinical trial in which treatments are randomly assigned to the subject in order to reduce potential bias in the patient response to treatment and thereby establish the statistical basis for analysing trial results.

Regression analysis	A statistical method that posits and estimates a linear relationship between a variable to be explained and a set of explanatory variable(s).
Sensitivity analysis	A mathematical calculation that isolates the factors involved in economic analysis to indicate their individual influence on results of the analysis.
Social perspective	A viewpoint for undertaking an efficiency analysis that incorporates all costs and health related effects regardless of who would incur the costs and obtain the benefits.
Technical efficiency	Situation in which health care interventions for particular health states are each performed with the least amount of inputs.
Time preference	The rate at which one is just willing to trade the present for future consumption of health services. A positive rate of time preference means that one is willing to forego some current consumption in return for a sufficiently large gain in future consumption.
Treatment	A set of health services provided for the diagnosis of an illness or suspected illness together with an attempt to cure that illness or alleviate its symptoms if its presence is confirmed.

Overview

Mounting pressures on health care resources lie behind the need to improve efficiency in the health sector. Australia's ageing population, increased personal use of health care, and increased treatment options associated with medical advances all contribute to the rising demand for health care. Due to the large size and rapid growth of health care expenditure, even small increases in efficiency can lead to considerable savings of resources or expansion of services for the community.

Purpose of the paper

This paper aims to provide some insight into the conceptual and methodological issues in measuring efficiency in health services. Both researchers and health system managers may find the paper useful.

- For researchers engaged in efficiency analysis, it is a précis of measurement techniques and applications.
- For those who have responsibility for health expenditure decisions or policy advice, it points to the appropriate analytical tools and necessary data for measuring efficiency.

The paper offers some guidelines on future research and data collection in the health sector. The discussion of efficiency concepts and evaluation methodology supports the ongoing work of the Commission and others in collecting performance data on health services.

A timely review of new techniques and data series

It is timely to review the state of knowledge about measuring health care efficiency, as this area has made some significant progress recently. Newly developed techniques and data series offer potential solutions for a number of data and estimation problems found in traditional efficiency measures used in the sector. These new techniques and data series open the way for an integrated approach to efficiency analysis, which allows for comparisons of efficiency across different service areas and aggregation levels, and making use of the outcome–output relationship in health care. The paper explains how this approach is implemented by Australia's National Health Performance Committee and the World Health

Organisation in undertaking a comprehensive assessment of efficiency in the health sector.

An economic framework for measuring efficiency

An understanding of relevant conceptual and methodological issues is essential for developing the appropriate data and analytical tools.

By developing an appropriate efficiency analysis framework, the paper illustrates the importance of collecting and using outcome data for measuring efficiency in health services. In Australia and overseas alike, the existing wealth of operational data and dearth of outcome data means that efforts should be directed to the measurement of health outcomes. This may require some adjustment in the information systems used in the health sector.

Further, the evaluation techniques must have a sound theoretical basis and be adapted to features of health and health services.

The basis of efficiency concepts and measures

Efficiency concepts are defined in terms of the objective of production assumed and the scope of activity analysed. In different studies of health care efficiency, the objective of production is perceived to be either providing services or achieving outcomes. Activities compared vary from alternative care procedures (also called health programs), individual service providers, to entire health sectors across countries.

Two components of economic efficiency are examined — technical efficiency and allocative efficiency.

- In the context of achieving health outcomes, technical efficiency is achieved by applying cost-effective care procedures with the least inputs.
- Allocative efficiency is achieved by choosing a set of technically efficient health programs to yield the greatest possible health improvements for the population.

This definition of efficiency is specific to health services and differs from one that is commonly used in other sectors such as manufacturing and agriculture, because health care efficiency is assessed in terms of outcomes achieved instead of outputs produced.

Central to the measurement of efficiency is the notion of the production frontier, which relates inputs to either outputs or outcomes while accounting for the effects

of external factors on productive performance. The production frontier is estimated to reflect technological and behavioural characteristics of the production process. Technical efficiency is measured as a distance to the frontier. Allocative efficiency is measured by comparing different points on the frontier for the extent to which they improve the health status of the population.

Complicating characteristics of health services

Compared with other industries, measuring efficiency in the health sector is complicated by characteristics specific to health and health services. This explains why it is necessary to adapt and modify efficiency concepts and evaluation techniques in the study of health care efficiency.

Market anomalies

Due to market characteristics specific to the health sector, a proper economic perspective requires evaluating health services in terms of health outcomes. There are particular considerations in relying on the market mechanism to guide the use of cost-effective health care procedures. Typically consumers have limited knowledge about health care. The supply of health services is characterised by regulation and market segmentation due to geography, service specialty and reimbursement arrangements.

Providing services versus achieving outcomes

Efficiency measures comparing resources used against the provision of services and, alternatively, against the achievement of health outcomes are not necessarily consistent, as service outputs may not vary directly with the resulting health outcomes. For instance, a costly medical procedure may represent a high level of service output but may offer little health benefits in terms of disease treatment. The paper explains that the adoption of cost-effective care procedures is required for consistency between outcome- and output-based assessments of health care efficiency.

The need to apply complementary measurement techniques

To identify the patterns and sources of inefficiency, efficiency is compared between service providers and between health programs. To support policy and clinical decision making, efficiency analysis evaluates costs and benefits at different hierarchical levels in a decision making system.

The study reviews two complementary types of efficiency measurement techniques:

- benchmarking analysis — which compares service providers, individually or collectively; and
- economic evaluation — which compares alternative health programs.

Benchmarking techniques are used to assess the level of technical efficiency relative to certain benchmark units. Benchmarking results reflect the selection of cost-effective care procedures for particular interventions and the level of operational efficiency in performing the chosen procedures. But the benchmark units may not use cost-effective procedures and can still be assessed as relatively efficient within the sample.

In contrast, economic evaluation is conducted under certain laboratory settings to uncover what health outcomes could be achieved if the available technology and resources were used to full advantage. Economic evaluation is thus useful for establishing the outcome–output relationship — that is, evaluation results provide the basis for converting and aggregating service outputs into an outcome measure. But economic evaluation does not take into account any difference between service providers in their ability to manage resources and administer services.

Economic evaluation

Economic evaluation techniques include:

- cost minimisation analysis;
- cost–effectiveness analysis;
- cost–utility analysis;
- and cost–benefit analysis.

These techniques have been applied in appraising new drugs, diagnostic technologies, surgical procedures and injury prevention activities. In such applications, a set of health programs are compared for their benefits and costs. Benefits are gauged by the consequences of a health program on people's wellbeing or health status. Costs refer to the value of opportunities or benefits foregone from not employing resources elsewhere. Data on costs and benefits are obtained from randomised controlled clinical trials, which use appropriate experimental designs to estimate the health effect of a specific health program.

The various evaluation techniques estimate costs in a similar fashion, but differ in the study perspective and the measurement of health outcomes.

Adaptable study perspectives

Possible study perspectives include those of the society, patients and their family, employers, insurers and service providers. The choice of study perspective affects the range of costs and benefits to consider. As many evaluation studies are funded or commissioned by particular organisations, the study perspective mostly reflects these organisations' corporate objectives or decision making contexts. For instance, a hospital contracting an evaluation study is less likely to be concerned with costs borne outside of its budget than with its own budget. It is suggested that, in such circumstances, a social perspective analysis should be run alongside an analysis adopting a narrowly defined study perspective in order to reconcile the interests of service providers with those of the community.

Different methods of comparing health outcomes

The different ways of measuring benefits lead to a trade-off between the scope for potential use and the practicality of various evaluation techniques.

Cost–benefit analysis converts program benefits in all forms into a dollar value. In principle, it has many potential applications as it can address both technical and allocative efficiency concerns within the health sector and between health and non-health uses. But expressing health outcomes in money terms is problematic and controversial. As a consequence, this technique remains little used in the health field.

Cost–utility analysis applies a generic measure of health status to compare program outcomes. Such an outcome measure combines the effect on mortality (length of life) and morbidity (quality of life). The past decade has seen the development of a variety of composite outcome measures that incorporate fatal and nonfatal conditions into the measurement of health status. This has expanded the scope for comparing different health programs.

In cost–effectiveness analysis, program outcomes are measured in a physical or natural unit of health status, such as the number of lives saved, life years gained, or reductions in disease incidence. Cost minimisation analysis compares program costs in situations where clinical evidence demonstrates the same outcomes for alternative health programs. It requires no explicit measurement of benefits. By comparison, these two latter techniques represent restrictive forms of assessment. But they are popular by virtue of their simplicity in comparing health programs that achieve particular health outcomes.

Benchmarking analysis

Unlike appraising a specific set of health programs under certain laboratory conditions, benchmarking analysis compares the efficiency of service providers in performing a variety of services. This form of analysis accounts for the operational aspects of production, such as resource management and service administration within an organisation.

The paper examines a collection of hospital studies to illustrate the methodological issues involved in benchmarking analysis, particularly the measurement of heterogeneous services.

The benchmarking techniques under review include:

- simple ratio analysis;
- unit cost analysis;
- stochastic frontier analysis; and
- data envelopment analysis.

These techniques are examined with respect to two respects:

- whether the technique is equipped to account for joint production and the influence of external factors on productive performance; and
- whether the technique measures efficiency in relation to treatment outcomes.

Flexible model specification

On the first criterion, data envelopment analysis is considered to be the most useful, chiefly because it allows for flexible specification of a complex production process, as required for modelling hospital services. Being a sophisticated form of ratio analysis, it is superior to simple ratio analysis for its built-in ability to make like-with-like comparisons, taking into account the input structure and service composition of individual hospitals in the sample. It can derive a single indicator of efficiency while allowing for multiple outputs and inputs, without using price or cost data. Nevertheless, due attention should be directed to a few modelling problems identified in previous hospital studies using data envelopment analysis. One problem is the misspecification of input-sharing patterns within a hospital. Model extensions developed in the study of other fields may lend solutions for this and other problems.

Linking efficiency to treatment outcomes

The second criterion is related to the outcome–output relationship. Benchmarking studies measuring hospital outputs in gross terms — like the total numbers of patient cases, patient days, or surgeries performed — risk making biased efficiency comparisons, particularly against hospitals that treat a relatively high proportion of severely ill patients.

A variety of variables have been used to represent hospital services. One widely applied method in measuring acute care services is to classify patients in diagnosis related groups and apply a set of cost weights to aggregate outputs across different patient classes. While this method controls for casemix differences between hospitals, using cost weights to aggregate outputs is problematic. The possible use of non-cost-effective care procedures means that service costs may overstate health benefits, hence leading to a biased assessment of efficiency in achieving health outcomes.

Some studies include crude indicators of case complexity (such as the casemix index and the high-technology index) or hospital-wide quality-of-care indicators (such as the misadventure rate, the unplanned readmission rate and the unexpected death rate). Such variables provide an indirect link between outputs and outcomes.

Using outcome measures in benchmarking analysis

Several recent studies introduce outcome measures to directly gauge the change in health status for individual patients between admission and discharge. Such outcome studies have been applied to intensive care units, obstetric services, and coronary artery bypass graft surgeries. Health status prior to hospital treatment is assessed using an illness severity scoring procedure. Treatment outcomes are assessed with reference to the discharge status or full recovery to normal health status. The assessment captures a systematic form of inefficiency associated with clinical decisions, including the use of ineffective health technology and redundant procedures. However, this approach confines the measurement of efficiency to homogeneous hospital services and requires in-depth clinical knowledge for preparing the required health data.

Lessons from the paper

The paper offers several lessons to improve the measurement of efficiency in the provision of health services.

- By reviewing efficiency concepts, the paper clarifies that studies of health care efficiency can be categorised in two dimensions: the cascade of activity levels; and the use of outcome- and output-based measures of health services. It is important to distinguish the conceptual and methodological issues behind these two linkages, as different techniques are required for addressing different issues.
- By reviewing a variety of efficiency measurement techniques and applications, the paper shows how estimation methodology and data collection is addressing the measurement issues in the study of health care efficiency. Several recently published outcome studies of hospital efficiency, coupled with the latest *World Health Report 2000*, show that outcome-based benchmarking analysis is feasible.
- By exploring how to link the different types of techniques, the paper outlines an integrated efficiency analysis framework providing a guide to the conduct of comprehensive assessment of efficiency in the health sector. The World Health Organisation and the National Health Performance Committee implement this framework in different ways. This reflects the differences in perceptions of various conceptual and methodological issues, as well as the differences in visions for developing analytical tools and data to address these issues.

1 Measuring efficiency

In general, measuring efficiency involves three tasks:

- identify relevant *model variables*;
- formulate an *efficiency measure* incorporating these variables; and
- obtain *data* to represent these variables and calculate the efficiency measure.

The first task depends on a conceptual understanding of the production process, including its technological and behavioural characteristics, as well as factors affecting producers' ability to perform. For the second task, the selection of an appropriate evaluation technique is based on its capability to generate robust and informative efficiency estimates, and to adapt to features of the production process analysed. The third task requires collecting data that are well defined, accurate, and consistent with the conceptual framework underlying the efficiency measure.

Distinctive features of health and health care complicate the various tasks of efficiency measurement in the health sector. Some of these features reflect the intangible nature of service production. Others are related to the complex determinants of health needs, the relationship between service providers and recipients, and factors affecting the choice of care procedures.

1.1 Scope of the paper

This paper reviews a range of techniques that have been applied for measuring efficiency in health services (box 1.1). The focus is on building a conceptual framework of efficiency analysis and identifying appropriate evaluation methods, rather than on the procedure of data collection. The emphasis on methodological issues is in line with the view expressed by Eckermann (1994), who raised concerns about the inappropriate use of data for measuring efficiency in the health sector:

> Both hospital productivity and hospital efficiency require careful conceptual consideration before assessing how current measures of hospital activity can be used to measure such phenomena. There is a temptation to use any available data (even if understandable) to assess efficiency and productivity as efficiency, without due recognition of the need to determine a sound theoretical basis on which to do so (p. 168).

Further, the Commission and others have been undertaking much work in developing and analysing performance data on health services in Australia (PC 1999; SCRCSSP 2001; AIHW 2000b, 2000c; ABS 1997, 2000; DHAC 2000a, 2000b; NHPC 2000). This paper aims to complement those research efforts.

Specifically, this paper aims to:

- investigate the state of knowledge about efficiency measurement in the health sector, both in Australia and abroad;
- show how features specific to health and health services affect the suitable methodology for measuring efficiency; and
- subject to the use of appropriate evaluation methods, point to areas where better data and modelling are needed to enhance efficiency analysis.

Data requirements are discussed in relation to the theoretical basis of efficiency analysis. By reviewing numerous efficiency studies in Australia and overseas, the paper shows the use of different available data series to represent particular model variables and highlights limitations of some of the data series used. In practice, however, data availability may restrict the selection of model variables and evaluation techniques.

Box 1.1 **Overall performance of health services**

Efficiency is one facet of the overall performance of health services. The WHO (2000) equates the performance of national health sectors to how well the countries pursue the service goals of achieving efficiency, responding to health needs, and averting severe personal financial losses due to ill health. In an Australian context, SCRCSSP (2001) considers efficiency, appropriateness, accessibility and equity, and quality of care to be the different parts of an evolving framework for reporting performance of public hospitals. The NHPC (forthcoming) similarly includes efficiency as one of the many service indicators in a national health performance framework.

Efficiency evaluation is one task under the broader, multidimensional approach to performance assessment and is undertaken subject to the achievement of other service goals.

1.2 Organisation of the paper

Chapter 2 defines efficiency concepts in the context of using resources to improve population health. To characterise the use and supply of health services, a conceptual model of health production is developed. This model provides the theoretical basis of efficiency measures.

Two types of techniques are needed to assess the various efficiency components in the health sector:

- *economic evaluation* (chapter 3) — which compares a number of health programs or health care interventions in terms of their costs and benefits; and
- *benchmarking analysis* (chapter 4) — which compares individual service providers against best-practice standards of the production process.

The subject of analysis distinguishes these two types of efficiency analysis. Economic evaluation is about choosing between alternative ways to provide health care, such as the use of alternative medical technologies or clinical procedures to cure particular illnesses. The following techniques are commonly used for economic evaluation purposes and will be discussed in the review:

- cost minimisation analysis;
- cost–effectiveness analysis;
- cost–utility analysis; and
- cost–benefit analysis.

For benchmarking analysis, the subject of analysis is a single or collective service provider — such as a surgeon or hospital, which in a sense organise and execute a range of health programs or health care interventions for delivering health services. The benchmarking techniques under review are:

- simple ratio analysis;
- average cost analysis;
- data envelopment analysis; and
- stochastic frontier analysis.

Chapter 5 examines the complementary role of these two types of techniques. They can be applied to address various methodological issues within an efficiency analysis framework. The World Health Organisation (WHO) and the National Health Performance Committee (NHPC) in Australia have proposed different ways to implement such a framework, which are compared in this chapter. The discussion sheds light on the necessary model extensions and data developments for enhancing efficiency analysis in the health sector.

2 Production of health and health services

Measuring efficiency requires a conceptual framework with which to specify the production process, identify the determinants of performance, and derive efficiency measures in terms of well defined variables.

This chapter explains how the economic concept of efficiency can be adapted to health services (section 2.1). It uses a model of health production to highlight the necessary attributes for analysing efficiency in health services (section 2.2). Two recent studies that apply this framework are examined for their methods and results in comparing relative efficiency in health services across countries. The latest empirical work of the WHO represents a macro approach to efficiency measurement (section 2.3). In contrast, the study by Baily and Garber (1997) adopts a micro approach in measuring efficiency at the disease level (section 2.4). These different approaches are shown to be complementary in producing informative measures of health care efficiency.

2.1 Concept of efficiency

The efficiency of a production process — that is, *productive efficiency* — refers to how well inputs are converted into final products. A production process may be defined to cover activities at different organisational levels — such as a whole industry, a firm, a production line, or a work procedure. Subject to the scope of production activities, the level of productive efficiency is measured by comparing actual and optimal amounts of inputs and products. The optimum is defined in terms of production possibilities (*frontier*) (box 2.1). A relative measure of efficiency is a function of distance to the frontier. In such a comparison, efficiency is *technical* since the frontier reflects the state of technology and the operating environment in which production takes place.

It is also possible to define the optimum in terms of the behavioural goal of the production process. Subject to the objective of production, efficiency is measured by comparing actual and optimal attainment of whatever the production process is supposed to pursue within the constraints on production possibilities. For instance,

the production process may be analysed as minimising cost, maximising income, or maximising social wellbeing.

Box 2.1 **Productivity and efficiency**

The terms — *productivity* and *efficiency* — are often used interchangeably to describe the comparative performance of production units. The concepts that underlie these two terms are related, but not identical.

Productivity is a gross concept that is measured by the ratio of products to inputs. It is gross in the sense that a productivity ratio may vary with differences in production technology, differences in the efficiency of the production process, and differences in the environment in which production occurs.

Efficiency is a component of productivity and refers to the comparison between actual and optimal amounts of inputs and products. The comparison can take the form of the ratio of actual to maximum potential products obtainable from the given inputs, or the ratio of minimum potential to actual inputs required for producing the given products. In both cases, the optimum is defined in terms of production possibilities and accounts for the impact of differences in the operating environment and production technology.

In general, efficiency is productivity adjusted for the impact of environmental factors on performance. In practice, however, few measures of efficiency account for all significant environmental factors.

Source: Lovell (1993).

The scope of activities defined and the objective assumed for a production process determine the appropriate efficiency concept and applicable evaluation techniques. In the health economics literature, efficiency comparisons have been conducted at different activity levels — the health sectors across countries, individual service providers, and alternative health programs or care procedures. Further, the objective of production is perceived to be either providing services or achieving outcomes. Because of such differences in the study context, there are considerable variations in the efficiency terms used. To maintain consistent terminology, this paper defines health care efficiency in a specific context as follows.

Health services are interventions provided to improve health for people in different health states. This reflects the primary objective of health care from a social perspective. In this context, *health care efficiency* refers to how well health care resources are used to obtain health improvements and comprises two components:

- *Technical efficiency* — whether health care interventions for particular health states (such as the treatment of illnesses) are each performed with the least amount of inputs.

- *Allocative efficiency* — whether a set of technically efficient interventions (such as the mix of interventions for treating different illnesses) is chosen to yield the greatest possible amount of health improvements.

The above definition of efficiency is specific to health services and differs in several ways from one that is commonly used in such sectors as manufacturing and agriculture. First, the final product of health care interventions is conceptualised to be the health consequences of service provision, rather than the amount of goods or services as measured for other production activities. Second, the definition of technical efficiency in health care is based on two types of comparisons: (i) comparisons of alternative diagnostic or treatment procedures applied to particular health states; and (ii) comparisons of service providers who choose and implement these care procedures. In non-health sectors, technical efficiency is defined mostly in terms of the latter type of comparisons only. Third, the definition of allocative efficiency in health care compares different forms of health care interventions for their impact on peoples' health, such as preventive care and acute care. This is different from the restrictive definition used in production economics (as in Lovell 1993), which considers only the substitution of inputs or outputs within a single production process in light of prevailing market prices. Further, allocative efficiency compares resource uses within the health sector, but not between health and non-health sectors.

In comparisons of health programs or care procedures, technical efficiency is referred to as *cost-effectiveness*. In comparisons of service providers, technical efficiency comprises the use of cost-effective care procedures for particular health states and the implementation of such procedures at least cost. The latter requirement reflects the *operational efficiency* of service providers — that is, the ability to manage resources and administer services within an organisation. This is not taken into account in the evaluation of cost–effectiveness for the care procedures used.

To achieve allocative efficiency, each intervention needs to be technically efficient. In addition, health resources need to be allocated across an optimal mix of interventions to yield the greatest amount of health improvements. The concept of allocative efficiency translates into choosing interventions according to their costs and health benefits to the population. Measuring allocative efficiency is meaningful only if the compared interventions are each practised in a technically efficient way. Otherwise, resources can be reallocated to increase health improvements for some people without making any others worse off.

To determine the frontier of health production — that is, the attainable level of health improvements, either a 'bottom-up' approach or a 'top-down' approach can

be used. This is illustrated in box 2.2, where various efficiency components are integrated into the overall efficiency of the health sector.

The 'bottom-up' approach involves measuring technical efficiency for individual service providers in benchmarking analysis, and that for alternative care procedures in economic evaluation. Benchmarking analysis compares the level of technical efficiency for a group of service providers against certain benchmark units. This assessment reflects the selection of care procedures for particular interventions and the level of operational efficiency in performing the chosen procedures. However, benchmarking techniques do not determine whether the benchmark units would use cost-effective procedures. Depending on the sample coverage, service providers not fully adopting cost-effective procedures may still be assessed as relatively efficient and chosen to be the benchmarks. As a result, benchmarking analysis may omit inefficiency arising from the use of non-cost-effective care procedures. In contrast, economic evaluation focuses on variations in the diagnostic and treatment approaches to caring for specific health states or diseases. This assessment uncovers what health improvements could be achieved if the available technology and resources were used to full advantage. But economic evaluation techniques cannot assess operational efficiency of service providers.

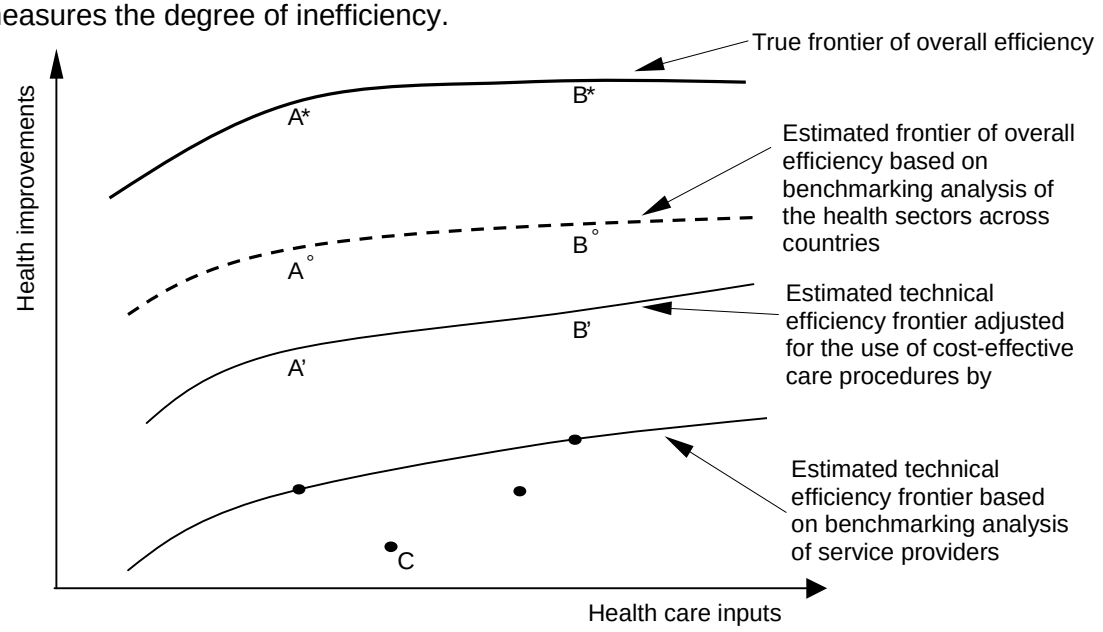
In the 'top-down' approach, the estimated frontier of overall efficiency is based on comparisons of aggregated health care inputs and health improvements achieved for the entire health sector. This way of measuring overall efficiency abstracts from comparisons of the relative worthiness between disparate interventions (for example, treating cancer versus heart disease) and from interpersonal comparisons of health gains. As with the 'bottom-up' approach, the assessment is made relative to observed benchmarks and, therefore, the estimated frontier may understate inefficiency.

There are several necessary conditions for these two approaches to lead to consistent assessments of the overall efficiency:

- The various efficiency components are measured with respect to the same objective of improving population health.
- Benchmarking studies are conducted, respectively, for all types of service providers — hospitals, general practitioners, community health centres, nursing homes, and so on.
- Economic evaluations are conducted to determine the scope for adopting cost-effective interventions over a complete range of health states.
- Economic evaluations are conducted to assess the priority of all health programs and thereby to determine allocative efficiency within the health sector.

Box 2.2 **Estimating the frontier of health production**

Conceptually the 'bottom-up' and 'top-down' approaches can be consistent in estimating the production frontier. To illustrate, in the figure below health improvements are measured on the vertical axis while health care inputs are on the horizontal axis. The various frontier lines represent the maximum possible levels of health improvements for each level of inputs as different components of efficiency are accounted for. For any data point, its vertical distance from one of the frontiers measures the degree of inefficiency.



The 'bottom-up' approach begins with the measurement of technical efficiency for a sample of service providers in four countries (*A* to *D*). The frontier line labelled as *AB* represents the averaged results for countries *A* and *B*, which are assessed as relatively efficient. Lying below the frontier, service providers in countries *C* and *D* are technically inefficient on average.

Suppose a series of evaluations of clinical procedures reveal the scope for introducing cost-effective treatment methods in countries *A* and *B*. Accounting for the potential health gains identified with the benchmark countries, the frontier is adjusted upwards to line *A'B'*.

Health improvements may also come from reallocating resources within respective countries. The impact is represented by a further upward shift of the frontier from *A'B'* to *A*B**. Line *A*B** represents the level of overall efficiency across the countries.

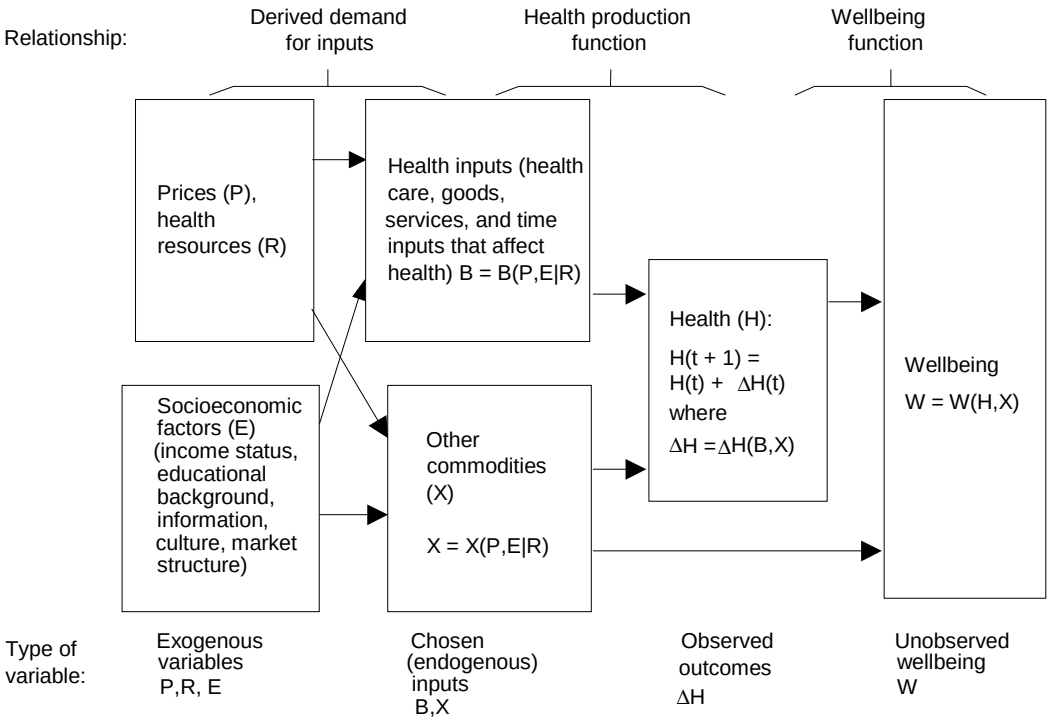
The 'top-down' approach uses aggregated country data to estimate the frontier of overall efficiency. The estimated frontier is the dotted line *A°B°*. It sits above line *AB* as it partly captures allocative efficiency. Line *A°B°*, however, understates the 'true' level of overall efficiency as even the best performing countries may have scope for efficiency improvement. Based on results of economic evaluation, *A°B°* can be adjusted upwards to match *A*B**.

In summary, economic evaluation and benchmarking techniques complement each other in revealing the diverse sources of inefficiency. No one technique can determine the full potential for improving efficiency in the health sector.

2.2 A model of health production

Figure 2.1 is a schematic representation of the process of health production. The model is adapted from DaVanzo and Gertler (1990) and draws on the theory of health investment pioneered by Grossman (1972a, 1972b). SCRCSSP (2001), AIHW (2000a) and NHPC (forthcoming) have all used a similar model to formulate performance indicators. Here, the model is used to identify the key variables in the frontier of health production.

Figure 2.1 **Model of health production**



Source: DaVanzo and Gertler (1990).

Production of health

Individuals inherit a certain health status at a point of time (denoted by H in figure 2.1), which changes over time (by the amount ΔH). Health status degenerates as a result of ageing. Some behaviour such as smoking leads to a decline in health.

Diseases and injuries cause reductions in health too. But health status can be augmented through health improvements.

The use of health inputs is derived from the demand for health. The demand for health is, in turn, determined as people seek to maximise their wellbeing (W) through an optimal mix of health and other commodities (X). The latter decision is affected by resource prices (P) and subject to the constraint of available resources (R) and the influence of socioeconomic factors (E) on preferences.

Gross improvements in health result from the use of health inputs (B), which include health care as well as the consumption of market goods and services for diet, exercise, recreation, housing and so on. Appendix A discusses the significance of health care relative to other health-enhancing activities. For instance, empirical studies suggest that human capital investment, education in particular, is probably the most important determinant of people's health and life expectancy.

The health production function represents the process of converting health inputs into health improvements. In the context of health production, the efficiency of health care is gauged by comparing the input of health care against the resulting health outcomes for a given amount of other health inputs and commodities.

The wellbeing function represents the satisfaction associated with a particular health condition. Other things being equal, wellbeing would increase with a longer and healthier life. To reflect health related wellbeing changes, health outcomes can be expressed in terms of *changes* in mortality, life expectancy, or functional ability. Pliskin et al. (1980), Mehrez and Gafni (1989), and Garber and Phelps (1997) elaborate the theoretical basis of linking wellbeing to life expectancy and health conditions. Their research supports the use of life-year and quality-of-life measures as proxy indicators for health related wellbeing. Nevertheless, these outcome indicators cannot be used to compare health and non-health programs, or to determine the optimal level of health care input. For such purposes, a broader metric of wellbeing is required to measure full potential benefits, whether health or non-health related (see the discussion of cost–benefit analysis in section 3.3).

Production of health services

Health care is an intermediate input in the context of health production, but an output in the context of service provision. Some measures of efficiency in health services are formulated comparing resource use against service provision (such as number of patient cases or consultations with service providers), instead of the resulting health benefits to service recipients. A fundamental issue for adopting the

context of service provision in efficiency analysis is whether efficiency is measured reflecting the objective to improve health of the population.

Service providers may seek to maximise revenue or minimise cost, or pursue other motives. Their behaviour can be in harmony with the objective to maximise social wellbeing through health improvements, provided that market prices provide effective signals to guide consumer and producer behaviour. The health sector, however, is not characterised by such model features of a competitive market.

In the case of health services, consumers are typically poorly informed — they have limited knowledge of their health care needs, available treatment options, and treatment outcomes. On the supply side, significant market imperfections exist — health services are often characterised by regulated doctor supply, market segmentation, and asymmetric information available for doctors and patients respectively.

Moreover, the agency relationship between doctors and patients means that service providers predominantly decide on the use of health services. Clinical decisions primarily reflect service providers' judgement and assessment of patients' circumstances, and involve a degree of uncertainty with respect to the outcomes and effectiveness of any chosen intervention. Service providers may have varying practice styles that are considered to be within the bounds of appropriate care. The discretion in clinical decisions leads to variation in resource use for obtaining particular outcomes. Service costs may vary among patients receiving the same treatment due to differences in their health conditions, complications in treatment, and the skill and operational efficiency of service providers.

Due to market anomalies in the health sector, the link between service provision and health improvements is potentially a weak one. The objective to improve health is not necessarily addressed by measuring efficiency in the context of service provision. Inefficiencies in caring for patients' health, including the use of non-cost-effective health care procedures, are undesirable both from the economic perspective and from the medical and ethical perspectives.

Determinants of economic efficiency in the health sector

Health services are mostly regulated and often directly supplied by government. Efficiency is measured reflecting the inseparable effects of clinical, operational, managerial, and policy decisions on service performance. This is the case not only for measuring overall efficiency of the health sector as a whole, but also for benchmarking technical efficiency among individual service providers. As Palmer (1991a) observes, a hospital is not like a conventional firm in the market sector and

cannot be appraised as if it runs a set of well defined production activities and makes managerial decisions under the control of a single transactor.[1]

As Drummond (1987) points out, health care policy can affect the diffusion and use of health care technology[2] and facilities in two ways — by *directive* and by *incentive*. Regulation by directive is enacted through the planning and funding of facilities, specialist departments or specific technologies. This mechanism is most relevant to expensive interventions and 'big ticket' technologies, which usually require significant research support to develop and a sufficient number of service recipients to maintain viable operation. On the other hand, regulation by incentive is enacted through controls of health service budget and reimbursement schemes (including private insurance) for health service providers.

The health sector may operate with a mixture of the two mechanisms, depending on institutional arrangements and policy settings. Regulation by directive tends to prevail in publicly financed health care systems like the British and Australian systems, whereas regulation by incentive is prevalent in a market-driven health care system like the one in the United States. Appendix B contains some examples showing the link between planning and clinical decisions in the health sector.

2.3 An application of the 'top-down' approach

The World Health Organisation (WHO 2000) adopts a 'top-down' approach in comparing the level of economic efficiency attained by the health sectors in 191 countries. The assessment of overall efficiency accounts for the combined influence of health policy, clinical decision making, and operating performance of the health sector.

The study breaks new ground in devising a measure of health outcomes and a statistical method to separate the effects of health care from those of other factors

---

1 Charnes et al. (1978) has generalised this point to the measurement of technical efficiency for other public sector programs. It is hard to distinguish between inefficiency in public services originating from planning decisions and service operations, as 'the meaning of a "technology" is likely to be more ambiguous than in the case of manufacturing in the private sector (p.430)'.

2 Health technology can be defined broadly, as by the US Office of Health Technology Assessment, to include "the drugs, devices, and medical and surgical procedures used in medical care, and the organisations and supportive systems within which such care is provided". A restricted definition of health technology would cover mainly the relatively expensive medical or surgical equipment or procedures — for example, magnetic resonance imaging, liver transportation and cervical screening for cancer. The discussion here adopts a general definition of health technology, although assessment of expensive facilities tends to have the major influence on health care policy and clinical practices (Drummond et al. 1991).

(AIHW 2001). Box 2.3 details the estimation procedure. In many respects, the WHO's approach conforms to the theoretical framework of health production:

- The production frontier is estimated relating changes in health status to the provision of health services and other health-enhancing inputs.
- Input variables, including health care, are expressed in physical or real terms.
- Relative efficiency is measured indicating what percentage of the greatest possible health improvements each country has achieved.
- Benchmark units are identified from a large pool of panel data for estimating the production frontier and constructing performance targets on the frontier.
- Income is excluded from the production frontier, as it is not a direct input to health production — it affects health status indirectly by driving the level of inputs used for improving health.

However, data and methodological problems remain and need to be acknowledged when interpreting results. Major limitations with the WHO's approach to measuring efficiency are as follows:

- By using a single variable — education — to represent all non-health inputs to health production, the production frontier may be misspecified. The resulting efficiency scores may be biased, depending on whether the omitted input variables vary directly with education and health care expenditure. Empirical evidence in appendix A suggests that diet, sanitation, housing and lifestyle are among the key determinants of health. However, it is difficult to obtain consistent data on these variables from all countries.
- The production frontier does not include any variable to capture certain nontrivial effects of country conditions on service performance. Further research is being undertaken by the WHO seeking to explain any estimated differences in health care efficiency across countries by means of relevant country-specific factors (Lauer et al. 2001).
- The corrected-ordinary-least-squares method is applied to estimate the minimum level of attainable health assuming the absence of modern health care. This technique is defective as the related equation is estimated to demonstrate average properties of the production relationship across countries rather than frontier properties pertinent to observations located near the data boundary.
- Shifts in the minimum level of attainable health since the early 1900s are ignored, as the related equation is estimated using data dated round 1900. For instance, while the eradication of smallpox may have raised the minimum health status, the emergence of HIV/AIDS has the opposite effect, making it harder than it was in 1900 to achieve a given level of health.

Box 2.3 **Measuring the efficiency of national health care systems**

The WHO measures health status by disability adjusted life expectancy (DALE), which is the average number of equivalent healthy life years that a newly born member of the population could expect to live. This measure accounts for the impact of both premature death and ill health on wellbeing. The number of life years abstracted from disability is estimated by advancing a population cohort to each age based on estimated birth rates and age-cause-specific death rates. Life expectancy is adjusted for the prevalence of disability. The adjustment is translated into an equivalent time loss by using a set of weights to account for the varying degrees of reduction in functional capacity due to diseases at each age. For instance, disability can make a substantial difference to wellbeing as limitations caused by injury, blindness, or paralysis may strike afflicted patients severely over their lifetime.

Health care is measured by national health expenditure per capita. The use of this aggregated input measure allows for variations in the input mix due to differences in relative prices across countries. To adjust for exchange rates and differing domestic price levels, health expenditure is expressed in purchasing-power-parity (PPP) terms.

The efficiency score is a ratio of actual health gains to maximum health gains that a country could expect to achieve with an efficient health care system:

$$\frac{\text{Actual DALE} - \text{Minimum DALE}}{\text{Maximum DALE} - \text{Minimum DALE}}$$

In this formula, all DALE measures are specific to a country. Maximum DALE establishes an upper limit corresponding to the greatest possible health gains for the given inputs. To project maximum DALE, a frontier equation is estimated relating adjusted life expectancy to health expenditure per capita and average year of schooling in the adult population. The latter variable is a proxy for factors beyond the health sector but influencing health. The inclusion of this variable in the frontier equation is based on the strong empirical association between health and literacy.

Minimum DALE sets a lower bound corresponding to a hypothetical scenario assuming the current literacy rate but that the health sector does not exist. Without health care in a modern sense, people would still be born, live and die, though with a shorter life expectancy. The turn of the last century is taken to be the starting point of modern health care. Under such an assumption, the WHO uses data for years around 1900 from a sub-sample of countries to estimate the relationship between life expectancy and educational attainment without modern health care. Minimum DALE is projected by substituting the current literacy level into the estimated equation.

Within the range between maximum DALE and minimum DALE, the actual DALE reflects the extent to which health gains have been realised from the input of health care throughout the past century. To improve efficiency means to move away from the minimum base and come close to the maximum potential. Nevertheless, every country has a different scale of achievement to calibrate performance.

Sources: WHO (2000); Evans et al. (2000).

Table 2.1 displays the WHO's efficiency estimates together with data on health expenditure and health status for a selection of developed countries including Australia. It appears that health expenditure has little correlation with health status or health care efficiency. While all the listed countries achieve a life expectancy in excess of 75 years (or 69 years adjusted for disability), health expenditure per capita varies widely across countries from a low of US$1200 for the United Kingdom and Ireland to over US$3700 for the United States.

Table 2.1 **Relative efficiency of the health sectors in selected countries**
Panel estimates, 1993 – 1997

	Health expenditure per capita				
	in international dollars [a]	*as per cent of GDP*	*Life expectancy at birth* [b]	*Disability adjusted life expectancy* [b]	*Efficiency index* [c]
Italy	1824	9.3	78.8	72.7	0.976
France	2125	9.8	79.3	73.1	0.974
Japan	1759	7.1	81.0	74.5	0.945
Austria	1960	9.0	77.4	71.6	0.914
Norway	1708	6.5	78.6	71.7	0.897
Netherlands	1911	8.8	78.1	72.0	0.893
Sweden	1943	9.2	79.5	73.0	0.890
United Kingdom	1193	5.8	77.2	71.7	0.883
Switzerland	2644	10.1	79.3	72.5	0.879
Belgium	1738	8.0	77.9	71.6	0.878
Luxembourg	1985	6.6	78.0	71.7	0.864
Ireland	1200	6.2	75.8	69.6	0.859
Canada	1836	8.6	79.1	72.0	0.849
Australia	*1601*	*7.8*	*79.5*	*73.2*	*0.844*
Germany	2365	10.5	76.9	70.4	0.836
Finland	1539	7.6	77.1	70.5	0.829
Denmark	1940	8.0	75.5	69.4	0.785
United States	3724	13.7	76.8	70.0	0.774
New Zealand	1393	8.2	76.6	69.2	0.766

a 1997 US dollars adjusted for purchasing power differences across countries. **b** Based on 1999 data. **c** The reported efficiency index for each country is an average of scores derived from numerous regression models used to specify the production frontier. In each regression, at least one country has an efficiency index of one. But across all regressions, no one country is consistently the most efficient in the sample. On average, Oman (not shown here) has the highest efficiency index calculated at 0.992.

Sources: WHO (2000); Evans et al. (2000).

The health sector in Italy is assessed to be relatively efficient, with the efficiency index estimated at 0.98 (out of a possible one). This compares to an estimate of just over 0.76 for the United States and New Zealand.

For Australia, the population is estimated to have achieved 84 per cent of the life expectancy that should be made possible given the spending of US$1600 per person

per year on health care and other substantial expenditure on education. This suggests some room for improving the performance of domestic health services, in spite of the country's high level of health attainment as indicated by the life expectancy that is the world's longest after Japan's.

Färe et al. (1997) also estimate the level of overall efficiency in health services for a sample of OECD countries. But their model does not incorporate a proper measure of health gains or make necessary adjustment for the effects of non-health factors. These deficiencies in model specification undermine the reliability of their results. Appendix C compares their method and results with those of WHO (2000).

As Evans et al. (2000) comment, the WHO's results almost certainly understate the full potential for each country to become more efficient. The fact that some countries have an efficiency index close to one does not mean that these countries have little scope for improving their health services. Rather it means that compared to these relatively efficient countries, other countries in the sample have greater room for improving their current level of population health. As efficiency is measured in relative terms, the assessment is unable to determine the full extent of inefficiency existing among the best performing countries. As a case in point, developed countries are all under pressure to reduce indiscriminate use of health care technology. Using data from these countries to estimate the production frontier fails to capture this common cause of inefficiency and, as a result, artificially lowers the benchmarking standard. However, micro level economic evaluation studies may suggest greater potential for improving efficiency.

2.4 An application of the 'bottom-up' approach

Baily and Garber (1997) report on a study launched by McKinsey and Company to compare relative efficiency *at the disease level* among the health sectors of the United States, the United Kingdom and Germany. The study uses aggregated national data on the inputs and treatment outcomes for four diseases — diabetes, gallstones (cholelithiasis), breast cancer, and lung cancer. As these health problems are common in the studied countries, the lessons that emerge from the analysis may shed light on the general performance of the health sectors compared. Moreover, by focusing on variations in the diagnostic and treatment protocols across countries, the study relates estimated differences in health care efficiency to provider incentives and regulatory settings in the health sector of each country.

To estimate the inputs used, the study identifies for each intervention the major care procedures involved, the key technology choices and clinical decisions that providers face at each step, and the resulting resource implications. Inputs are

measured in physical units for labour (from physicians, nurses, technicians and other medical staff), supplies (such as medications, surgical instruments and X-ray film), and capital (such as diagnostic equipment and hospital facilities). Administrative inputs are omitted. Based on standardised cost estimates, the inputs are aggregated into a composite input measure for each disease treatment in each country.

To estimate the treatment outcomes, it is necessary to quantify health status expected for each disease with and without treatment. Health status is quantified using either survival rates or calculations of the quality of life, depending on the primary goal of intervention (table 2.2). For the diabetes case, treatment outcomes are adjusted for differences in the rate of disease incidence between countries. This is done by comparing outcomes among white patients only, as non-whites are known to have higher rates of diabetes and of diabetic complications than whites. For the other disease cases, external factors — such as lifestyle, culture, genetics and so on — are assumed to have the same impact on the health status of patients in each of the studied countries. Under this assumption, untreated health status would be the same across countries and the estimated health status of treated patients is taken to provide a reasonable basis for outcome comparisons.

As the sample is limited to only three countries, the study cannot determine the health production function for each disease case. Nevertheless, the assumption of *diminishing returns in health production* is made. That is, patients respond to additional units of disease management or treatment with successively smaller health improvements. Further, as severely ill patients tend to receive care first, an expansion in health services to treat less severe patients generates diminished additional health benefits. Without estimating the production functions, estimates of inputs and outcomes can only be used to make crude comparisons of relative efficiency. For instance, a country is relatively efficient if it achieves better outcomes while using less of the inputs than other countries. Also, a country is more efficient if it achieves better outcomes and has a greater outcome–input ratio than another country. However, efficiency comparisons between two countries is indeterminate if one of them achieves better (worse) outcomes but has a lower (greater) outcome–input ratio than the other country. This is because the lower (greater) outcome–input ratio may indicate diseconomies (economies) of scale in health production, rather than any difference in productive efficiency.

Table 2.3 summarises the study results. The United States appears to be the most efficient among the sample countries in the management and treatment of lung cancer and gallstones. The United Kingdom is relatively efficient for diabetes. For breast cancer, the efficiency comparison between the United States and the United Kingdom is indeterminate; but both countries outperform Germany.

Table 2.2 **Measurement of treatment outcomes for specific diseases**

Disease case	*Intervention goals*	*Outcome measures*
(1) Diabetes There is no cure for this disease. Complications are frequent, can significantly diminish quality of life, and can even be life threatening. Common complications include heart and kidney diseases, visual impairing (which may lead to blindness), and foot ulceration (which may require amputation). Treatment consists of controlling the patient's diet and exercise habits, and taking oral medications to control blood glucose levels.	Delay or prevent some of the accompanying complications.	Complication rates are calculated for: diabetic ketoacidosis and hyperosmolar coma; retinopathy; blindness; and lower extremity amputation. Overall outcome measure is based on the complication rates and a set of utility ratings assigned to gauge the expected impact of each complication on a diabetic's quality of life.
(2) Gallstones (Cholelithiasis) This disease can cause abdominal pain and other symptoms that are mostly mild, transient, and slow in their development. Serious complications are uncommon, although they can lead to life-threatening conditions such as acute cholecystitis and gallbladder cancer. To remove gallstones, two approaches to cholecystectomy are commonly used, namely surgical removal of the gall bladder with its contents and laparoscopic operation.	Displace gallstones and prevent recurrence of symptoms.	Outcomes are measured to reflect the impact of symptom relief on each patient's quality of life, taking into account the frequency of symptomatic pain episodes and the utility ratings assigned for each pain episode before, during, and after surgery. Outcomes from operation are assumed to be similar between countries on the ground that both treatment options have similarly high success rates.
(3) Breast cancer If left untreated, this disease is often fatal as it has the tendency to spread from the breast to distant tissues. Disease management procedures are performed in four phases: screening; assessment; therapeutic; and follow-up. In the therapeutic phase, the treatment is aimed to remove the primary tumor while it is still localised to the breast and to prevent or halt its spread.	Treat the disease upon diagnosis of its existence and upon any relapse.	Outcomes are measured by the age-adjusted survival rate for women diagnosed with breast cancer at the fifth year after diagnosis.
(4) Lung cancer This disease is often incurable due to its poor prognosis at an early stage and the tendency for the cancer to spread widely to other organs. Symptoms include persistent cough, breathing difficulty, abnormal sputum, chest pain, and repeated attacks of bronchitis or pneumonia. Disease management is enacted in three phases: diagnosis and staging; curative care; and palliative care. Curative care is warranted in only a minority of cases. As an alternative to a complete cure, palliative care is often offered to ameliorate symptoms and maintain patient comfort without active therapy.	Save or prolong the patient's life.	Outcomes are measured by the life years saved over a five-year period after diagnosis, under the assumption that the impact of treatment on the patient's quality of life is comparable across countries.

Source: Baily and Garber (1997).

Medical knowledge and technologies available in the studied countries are fairly similar, hence unlikely to be responsible for any differences in efficiency. Instead, the study finds that differences in the treatment protocols used for particular diseases give rise to different levels of efficiency. Detailed comparative analysis of the health sector characteristics suggests that provider incentives and regulatory constraints explain much of the variations in health care efficiency (box 2.4).

Table 2.3 **Estimates of relative efficiency in treating selected diseases**

Indexes based on estimated outcome–input ratios[a]

	United States	United Kingdom	Germany
Diabetes	0.52	1.00	na
Gallstones	1.00	0.54	0.64
Breast cancer	0.88[b]	1.00[b]	0.56
Lung cancer	1.00	0.73	0.55

a For each disease case, the country that has the highest outcome–input ratio is assigned an efficiency index equal to one. The efficiency indexes are estimated in proportion to the outcome–input ratios. **b** Under the assumption of diminishing returns in health production, these estimates cannot be used to draw inferences on relative efficiency between the United States and the United Kingdom. na: Not available.

Source: Baily and Garber (1997).

Reconciling the 'top-down' and 'bottom-up' approaches

The findings from the study of specific diseases, particularly the favourable assessment of the US health sector, seems inconsistent with the common perception of the aggregate picture, like the one suggested by WHO (2000).

Among the possible explanations for such discrepancies are the different rates of disease incidence between countries, the focus of the case study on inpatient services, the selected sample of disease cases, the measurement of technical efficiency only, and problems in reconciling aggregated and disaggregated data.

Another important reason is the omission of the 'system costs' in the case study. These costs are incurred mainly in the management and administration of health care facilities and the insurance and reimbursement system. It is hard to divide these costs by disease. As Baily and Garber (1997) observe, the US health sector probably consumes a relatively large amount of administrative inputs due to a high degree of fragmentation in its provider markets and insurer markets, and the resulting complexity of its insurance and reimbursement system. Increases in these administrative costs may have eroded much of the efficiency advantage in disease treatment.

Box 2.4 **Major country differences in the treatment approaches**

In the United States, most physician services are negotiated and compensated by payers, including Medicare and private insurers, on a fee-for-service basis. Service providers compete for patients by offering relatively high standards of care, while payers are inclined to adopt extended coverage in their health insurance products to meet the demand for new treatment methods. Competition in both care provision and health coverage, coupled with the threat of malpractice suits against physicians, has led to increased treatment intensity and rapid technology adoption. The aggressive approach to treatment results in large improvements in outcomes for lung cancer, breast cancer (as opposed to screening), and gallstones compared to the associated increments in inputs. For instance, technology adoption in cholecystectomy has contributed significantly to the improved efficiency in removing gallstones.

The US third-party payment system encourages the use of mammographic screening technology. Despite the substantial costs involved, such screening often generates false positives, hence resulting in limited outcome improvements for breast cancer patients. The low degree of cost-effectiveness in diagnostic services partially offsets the technology advantage in therapeutic services. Further, private insurers have an incentive to avoid diabetic patients because they make above-average claims. Adverse selection in health insurance makes routine care less accessible to diabetic patients. The repayment schedules established for doctor visits exacerbate the difficulties in treating this disease — for instance, Medicare and Medicaid do not cover the cost of a diabetologist. The country has comparatively high rates of diabetic complications, which contributes to considerable reductions in quality of life over a diabetic's lifetime.

The UK health sector is characterised by tight control of physician supply, salary-based contracts for specialist services, and regulation by directive of clinical practices. Capacity constraints on hospital services force providers to be selective in choosing patients to treat and to substitute procedures that conserve resources. Limited incentives to providers and limited funding for capital investments both contribute to slow technology adoption. The country is notable for doing less screening and surgery for the cancers and less surgery for removing gallstones. Patients with lung cancer or gallstones are more likely to be given palliative care, hence smaller outcome improvements from treatment, than in the United States.

The success of the UK approach to diabetes seems to come from training patients in self-care methods, rather than from increasing treatment intensity. This approach is judged to be an important example of treating a chronic disease, using a specialised clinic and hence reduced resources to achieve better outcomes in the long term.

German hospital services are negotiated and compensated on a per day basis. Moreover, German hospitals face the threat of regulatory review and capacity cuts if their utilisation falls below a required level. These factors create incentives for service providers to enrol inpatients rather than outpatients and to provide an increased amount of care, particularly to prolong hospital stays. The regulatory regime and the incentive structure thus partly explain the efficiency gaps relative to the other countries.

Source: Baily and Garber (1997).

Only the ‘top-down’ approach to measuring efficiency can capture the impact of the ‘system costs’ on sectoral performance and provide a comprehensive assessment of overall efficiency for different health services. However, by comparing efficiency at a micro level, the ‘bottom-up’ approach as applied in this study is useful for identifying the sources of inefficiency in different types of health services.

3 Economic evaluation of health programs

Economic evaluation is the comparative analysis of *costs* versus *consequences* of a set of *health programs*. It provides a guide for service providers to select cost-effective care procedures — improving technical efficiency, and for policy makers to determine health care priorities — improving allocative efficiency.

In the health field, the term '*health program*' is used interchangeably with the term '*health care intervention*'. Health programs encompass a broad range of service activities for health promotion, prevention of disease, establishing diagnosis, curing disease, rehabilitation, and palliative care. Economic evaluation studies have been conducted to appraise new drugs, surgical procedures, diagnostic technologies, screening for cancer, and injury prevention activities.

This chapter introduces the basic principles of economic evaluation (section 3.1). Economic evaluation studies are distinguished by a number of key features of the study design and prerequisite (section 3.2). Various evaluation methods are available and can be adapted to suit the decision making context (section 3.3).

3.1 Basic principles

Economic evaluation has three guiding principles:

- accounting for both costs and consequences;
- comparing a specific set of health programs; and
- using outcome indicators to represent consequences.

For a health program to be worthwhile, the value of its benefits must exceed the value of its costs. 'Benefits' and 'costs' are broad concepts, referring to a range of factors including and beyond monetary concerns. Benefits stem from the consequences of a health program on the wellbeing of individuals and the community at large. Costs refer to the value of opportunities or benefits foregone as resources are consumed for a particular health program.

Economic evaluation compares two or more alternative health programs for their effectiveness in generating health benefits to service recipients. It is used to investigate two different sources of inefficiency in health services. The first source relates to the question on technical efficiency: Given a program that is worth doing, what is the best way of providing it? This question seeks the best health program available for a condition- or disease-specific group of patients. The second source of inefficiency relates to the question(s) on allocative efficiency: Is a particular health program worth doing and, if affirmative, what is the optimal level of that program — that is, how large should it be?

Depending on the study setting, different measures of health benefits are used in comparing alternative health programs. To assess technical efficiency, health benefits are expressed in terms of changes in life expectancy, functional health status, or the health related quality of life. Such outcome indicators are commonly used in cost minimisation analysis, cost–effectiveness analysis, or cost–utility analysis. To assess allocative efficiency, outcomes from dissimilar health programs are converted into a common denominator, such as the dollar value used in cost–benefit analysis and the composite utility-based measure used in cost–utility analysis. Details of these techniques are in section 3.3.

Studies not belonging to economic evaluation

It is instructive to identify several forms of partial evaluation that are related to, but do not belong to, economic evaluation (figure 3.1):

- Burden- or cost-of-illness studies — estimate the burden of health effects and economic costs on the community due to particular diseases.
- Epidemiological studies — address the causes and distribution of particular diseases, and the efficacy of health programs in the treatment of those diseases.
- Description studies — estimate either the costs or outcomes associated with a single health program.

These studies provide useful baseline information for economic evaluation, but have important limitations relative to full economic evaluation. Burden- or cost-of-illness studies do not provide evidence on the effectiveness and availability of health programs to treat the disease(s) under consideration. Epidemiological studies do not compare relative costs between different health programs. Partial evaluation studies also do not provide information on relative opportunity costs, as alternative health programs are not considered.

Figure 3.1 **Distinguishing characteristics of economic evaluation**

Is there a comparison of two or more alternatives?	Are both costs (inputs) and consequences (outcomes) of the alternatives examined?		
	No		**Yes**
	Only consequences are examined	Only costs are examined	
No	**Partial evaluation:** Outcome description	**Partial evaluation:** Cost description	**Partial evaluation:** Outcome and cost description
Yes	**Partial evaluation:** Epidemiological analysis	**Partial evaluation:** Burden- or cost-of-illness analysis	**Full economic evaluation:** Cost minimisation analysis; Cost–effectiveness analysis; Cost–utility analysis; Cost–benefit analysis

Source: Drummond et al. (1997).

3.2 Key features

Irrespective of the technique applied, economic evaluation should examine two key aspects for each of the health programs under consideration:

- *Efficacy* of the health program — Can it work? Does it do more harm than good to individuals who fully comply with the treatment and program recommendations?
- *Effectiveness* of the health program — Does it work? Does it do more harm than good to individuals to whom the program is provided. That is, efficacy of the program and its acceptance by individuals contribute to effectiveness.

Use of clinical evidence

To answer the above questions, economic evaluation provides clinical and epidemiological evidence on the comparative ability of different health programs to improve health. Such information is primarily obtained from randomised controlled clinical trials, particularly for comparisons of medical treatments. Randomised controlled trials use appropriate statistical designs to estimate the contribution of a specific health program to patient health.

Clinical trials provide evidence on either the efficacy or the effectiveness of health programs. Efficacy evidence assesses health programs under certain laboratory conditions. Such information is crucial for establishing the safety of a health program, hence whether it should be considered for use. However, efficacy evidence can be an imperfect indicator of the success of a health program in practice. For instance, actual rates of drug compliance and screening uptake may be less than ideal. Realised health improvements can be, sometimes substantially, lower than those found under laboratory conditions. By drawing on effectiveness evidence from clinical trials, economic evaluation can more accurately predict health improvements from the implementation of health programs. However, effectiveness data are not always obtainable. As such, it is common to infer effectiveness from efficacy evidence under certain assumptions about the conditions in which a health program is implemented.

Four evaluation stages

Typically economic evaluation is conducted in four stages (figure 3.2):

- define the study question and perspective;
- determine the way to identify and measure costs and benefits;
- compare costs and benefits; and
- derive the decision rule.

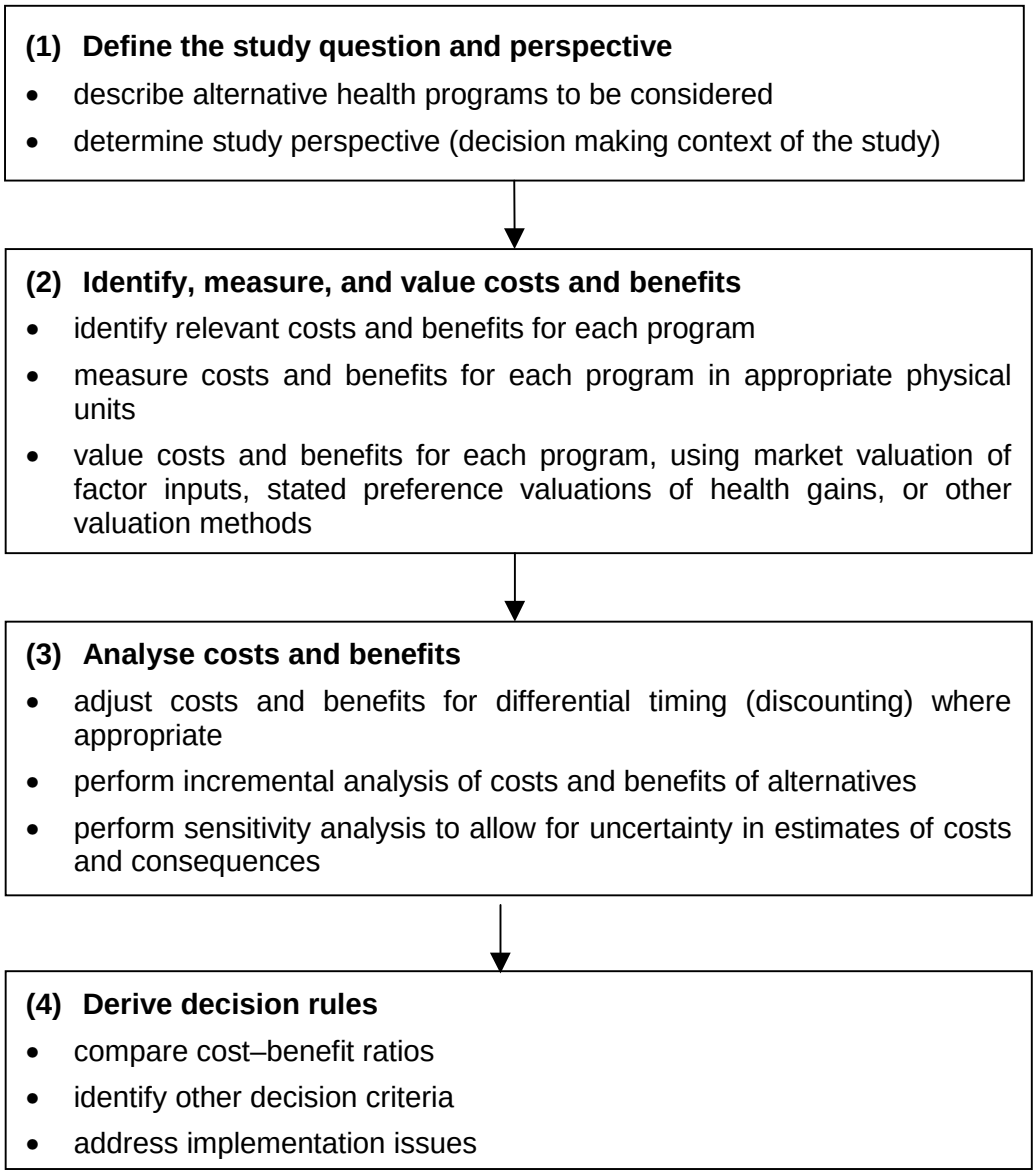
A brief account of these evaluation stages is given here. The major methodological aspects relating to the study perspective, the treatment of costs, and the valuation of benefits are discussed in greater depth in section 3.3.

Study question and perspective

Any economic evaluation study has a research question to address, which should identify a specific set of health programs for comparison, including the option of "doing nothing" or the current baseline program. Examples of the study question include "Is a hospital-based haemodialysis preferable to a home-based haemodialysis?" or "Is community-based bowel screening a worthwhile use of resources?"

The study perspective — that is, the decision making context of the study — has implications for what costs and consequences are considered and what decision rules are used in economic evaluation. Possible study perspectives include the service provider (like the hospital), the health department, the third-party payer, the government, and the society as a whole.

Figure 3.2 **Four stages of economic evaluation**



Identification and measurement of costs and benefits

Depending on the study perspective, relevant cost items include costs borne by the health sector, patients and other sectors. Similarly, benefits may be identified relating to therapeutic effects, changes in the quality of life for patients and their families, and resource savings subsequent to the implementation of a particular health program.

Costs and consequences should be measured in appropriate physical units. This allows study results to be replicated and generalised to other locations or time

periods, as resource prices may differ to the study setting. Examples of appropriate cost units are hours of nursing time, number of physician visits, pharmaceuticals used, and operating theatre hours used. Appropriate units of health benefits are years of life saved, number of deaths averted, and improvements in clinical measures of emotional, physical or social functioning. Benefits may also be quantified in a specific context, such as the reduction in serum cholesterol.

Costs and benefits should then be valued appropriately for comparison. This is further discussed in section 3.3.

Analysis of costs and benefits

Costs and consequences of alternative programs are compared at one point in time. But costs may be incurred over a period of time and benefits may accrue into the future. Where costs and consequences do not occur entirely in the present, they should be discounted on the basis of time preference.

Costs and consequences should always be compared on an incremental basis. For a meaningful comparison, it is necessary to examine the additional costs that one health program imposes over another, compared to the additional benefits provided (Drummond et al. 1997). Box 3.1 illustrates the importance of incremental analysis.

Sensitivity analysis of results should be performed to test assumptions made in the analysis. Threshold values for key variables should be identified where decision rules change and the robustness of the measurement for those variables explored further if small changes in their values affect a particular decision rule.

Decision rules

A core decision rule typically uses outcome–cost ratios to rank the health programs under consideration. Other relevant decision criteria should assess the distribution of health benefits and the practicality of health programs. For instance, consideration may be given to whether older patients should be favoured over the young, or whether existing infrastructure is capable of supporting any change to existing services.

3.3 Evaluation techniques

Economic evaluation studies adopt similar methods to measure costs, but differ in the study perspective and the valuation of benefits.

Box 3.1 **An example of incremental analysis**

Hull et al. (1981) estimated the cost per correct diagnosis for a sample of 516 patients with deep vein thrombosis. Two alternative diagnostic programs were examined: impedance plethysmography (IPG) alone, versus IPG plus outpatient venography if IPG was negative. Venography is an unpleasant procedure but represents the 'gold standard' diagnostic tool for deep vein thrombosis. Results were:

	Cost (US$)	Outcome (Number of correct diagnoses)	Ratio of cost to outcome (US$ per correct diagnosis)
Program (i): IPG alone	321 488	142	2264
Program (ii): IPG plus outpatient venography if IPG negative	603 552	201	3003
Incremental change between programs (i) and (ii)	282 064	59	4781

The average cost per correct case detected from program (ii) is US$3003. However, the true measure of the cost-effectiveness of this program is the cost per additional case detected over program (i). It turns out that program (ii) only detects a further 59 cases, and the cost for each additional case detected is US$4781, higher than the average cost.

This example shows a difference of greater than 50 per cent between average cost and incremental cost. In some other studies, it has been shown that the incremental benefits of a program are so small relative to costs that whilst average cost per unit of benefit appear 'reasonable', incremental cost is several hundred times higher.

Source: Hull et al. (1981).

Similar treatment of costs

Economic evaluation accounts for four types of costs:

- direct health care costs — incurred by service providers in treatment, health promotion, prevention, rehabilitation, and palliative care programs (such as the costs of surgery, physician time, and heating and lighting in hospitals);
- direct personal costs — incurred by individuals in receiving treatment or participating in the health program (such as travel costs and out-of-pocket fees);
- direct non-health costs — incurred by agencies outside the health sector (such as welfare and housing services, police, courts, informal care, and Workcover);

- indirect costs — incurred by patients and the society due to the net loss of productive time.

A health program may result in both resource burdens and resource savings when compared to alternative programs. To avoid double counting, the convention is to subtract resource savings from resource burdens and thereby express the overall cost on a net basis. As an example, for running a quit-smoking television campaign, the resource burdens are the costs of producing and broadcasting commercials, and the resource savings are the reduction in expenditure for treating lung cancer in the future.

While the concept of opportunity costs requires costing all resources consumed directly and indirectly in the provision of health services, opportunity costs are not all measurable in practice. As a result, actual resource consumption is commonly used as a proxy for opportunity costs. The underlying assumption is that the value of benefits from the next best alternative use of those resources is equal to the measured value of those resources.

Resources are commonly valued in market prices. Market price information for factor inputs is readily available for many types of health care resources, such as consumables and staff. Market prices would indicate the opportunity costs of those resources under the assumption of competitive factor markets. But market distortions prevail in the health sector. For example, physician numbers and licences are primarily limited by the intake of medical schools but also directly regulated by government in some circumstances. This may inflate physicians' salaries above their opportunity costs. As another example, under a needle exchange program, the opportunity cost of providing syringes is the unsubsidised cost of syringes rather than the cost to participants, which is frequently zero. However, economic evaluation cannot account for all of such supply-side distortions, and adjustments are justified only if they are likely to have a significant impact on study results.

Some resources do not command a market price, such as the time of volunteers and family members in the provision of care. To value these resources may require imputing opportunity costs through the estimation of shadow prices. For persons not in employment, the shadow price for their time may be taken to be the marginal wage rate for those persons if they were employed, on the ground that the carer could have spent time earning a wage. Alternatively, the opportunity cost of their time may be taken to be zero, especially if they are not actively seeking employment.

Further, where resource use cannot be directly attributed to a particular program, as often is the case for overheads and capital, it is necessary to assume some allocation of costs between and within programs.

Different study perspectives

The choice of the study perspective mainly affects the range of costs to consider in economic evaluation. However, there are no definitive rules on which costs should be included under a particular perspective. Table 3.1 provides some examples of the types of costs included under different perspectives.

Table 3.1 **Costs under different study perspectives**

	Societal	*Patient and family*	*Self-insured Employer*	*Public or Private Insurer*	*Managed Care Plans*
Cost element:					
(1) Medical care	All medical care costs	Out-of-pocket expenses	Covered expenses	Covered payments	Covered services
— Unit	*All units*	*Those paid out-of-pocket*	*Those covered*	*Those covered*	*Those covered*
— Price	*Opportunity cost*	*Amount paid out-of-pocket*	*Amount paid plus administrative cost*	*Amount paid plus administrative cost*	*Marginal costs*
(2) Patient time cost for treatment	Cost of all time used	Opportunity cost to patient	Paid sick time and administrative cost	na	na
(3) Professional care giving	All costs	Out-of-pocket expenses	Covered payments	Covered payments	Covered payments
(4) Informal care giving	All costs	Opportunity cost to care giver	na	na	na
(5) Transport and other non-medical services	All costs	All costs	na	na	na
(6) Sick leave, disability, other transfer payments	Administrative costs only	Administrative costs only	Amount paid by employer plus own administrative costs	Amount paid by insurer plus own administrative costs	If any paid

na: Not applicable.

Source: Gold et al. (1996).

In general, a social perspective considers all individuals affected by a health program, and all significant outcomes and costs from that program irrespective of who would experience the outcomes or bear the costs. Depending on the program scope, costs and benefits may be assessed for a small group of individuals or for a large population. For instance, costs and benefits of hernia repair are assessed for a patient and a family carer at home, whereas the assessment of a water fluoridation scheme is taken for the entire population.

Other study perspectives omit some costs and consequences from consideration on the ground that they are not relevant for the decision making context. A perspective of third-party insurer may ignore the costs that are not borne directly by the insurer, such as patients' out-of-pocket expenditures. Similarly, a government perspective may include direct outcomes for the individual, costs directly borne by the government, and productivity losses to the individual and the economy as a result of illnesses.

The influential Washington Panel on Cost Effectiveness in Health and Medicine has endorsed the use of a social perspective in economic evaluation (Gold et al. 1996):

> We might reasonably prefer a system in which decisions about health interventions reflected the seriousness of the problem and the ability of the intervention to do something about it, without reference to the specific individuals with the problem or to particular budgets or special interests (p.7).

In spite of the above recommendation, the majority of economic evaluation studies have adopted a relatively narrow perspective. To the extent that many economic evaluations are commissioned or funded by public and private sector health organisations, the choice of the study perspective is influenced by their corporate objectives and decision making contexts. For example, a hospital that commissions an economic evaluation is less likely to be concerned with costs borne outside of its budget than with the efficient use of its own budget. Gold et al. (1996) suggest that, in such situations, particularly in the case of publicly funded programs, a social perspective analysis should be run alongside analysis based on a narrower study perspective.

Different treatment of benefits

The various forms of economic evaluation approach the valuation of benefits in different ways (table 3.2):

- Cost minimisation analysis (CMA) requires no explicit measurement of benefits, which are assumed to be identical between the health programs compared.
- Cost–effectiveness analysis (CEA) does not explicitly derive the values of benefits, which are left in natural units such as life years gained or number of cases detected.
- Cost–utility analysis (CUA) represents benefits in terms of a gain in life years weighted by utility indices reflecting the health related quality of life.
- Cost–benefit analysis (CBA) represents benefits in terms of money values of health improvements.

Table 3.2 Treatment of benefits in alternative evaluation methods

	Nature of benefits considered	*Measurement of benefits*	*Selection criterion*
Evaluation method:			
(1) Cost minimisation analysis	Identical outcome from alternative health programs, such as the same number of lives saved	Valuation of outcomes not necessary	Minimise program cost; rank alternative programs
(2) Cost–effectiveness analysis	Only one 'dimension' of outcome (single effect)	Natural units, such as lives saved or cases of a disease detected	Minimise cost per unit of outcome or maximise outcome per unit of cost; rank alternative programs
(3) Cost–utility analysis	Several dimensions of outcome (multiple effects); the impact of illnesses on quality of life explicitly quantified	Quality adjusted life years (QALYs)	Minimise cost per QALY gained or maximise QALYs per unit of cost; rank alternative programs
(4) Cost–benefit analysis	Several dimensions of outcome (multiple effects); the impact of illnesses on quality of life explicitly quantified	Dollars	Benefits exceed costs

Source: McKie et al. (1998).

Cost minimisation analysis

CMA is the appropriate study design where clinical evidence demonstrates the same outcomes for alternative health programs, as the only relevant difference between them is in their costs. Examples include comparisons of:

- surgical and drug interventions that both return an individual to full health following treatment;
- alternative birthing options for women with no medical complications; and
- alternative health promotion strategies using media campaigns and school education programs to achieve the same reduction in smoking incidence and prevalence rates among teenagers.

Since it is difficult to validate the assumption of outcome equivalence on the basis of clinical evidence, CMA applications are rare. For instance, the mother may not only be concerned with the prevention of medical complications during childbirth, but also consider the environment in which childbirth takes place, the level of pain, and drug use as important outcomes in their own right.

Cost–effectiveness analysis

Variations in the level of outcomes between alternative programs have led researchers to apply a CEA study design. In a sense, CMA is a special case of CEA where it has been shown that outcomes from alternative programs are identical.

Under CEA, health outcomes from alternative programs are measured in terms of a common physical or natural unit of health gains. Outcome measures commonly used in CEA are number of lives saved, life years gained, reductions in disease incidence, and improvements in ability to perform daily tasks. Where measuring effective outcomes is infeasible, measures of intermediate outcomes are used, such as number of individuals screened, positive diagnostic tests, and successful operations.

The main strength of CEA is its ability to determine the least costly way to treat a given condition for different levels of health outcomes. The most effective health program is identified according to a cost–effectiveness ratio — that is, the ratio of net resource costs to net benefits from a given health program, where net costs and net benefits are relative to any baseline program.

CEA has three main limitations.

- The requirement to measure outcomes in natural health units limits efficiency comparisons to a relatively small set of similar programs or conditions. CEA cannot be used to compare dissimilar programs. For example, a program aimed at restoring the sight of patients is not comparable to a program for treating foot ulcers under CEA.
- Natural health units capture only one dimension of program outcomes. For example, a measure of life years gained ignores the quality of life as another dimension of outcomes.
- CEA provides evidence on which of the alternatives considered is more efficient in achieving the desired end, but not on whether any of the options should be provided at all.

The judgement on the optimal level of health budget has to be made independently from CEA, or using a CBA study design as discussed below.

Cost–utility analysis

Variations in both the level and dimension of outcomes between alternative health programs have led to the development of CUA study designs. Under this approach, program outcomes are measured in terms of changes in the health related utility for

those receiving services. The valuation of health related utility explicitly combines the effects of a health program on mortality (life years gained) and morbidity (quality of life) aspects of health. The use of generic, utility-based measures of health allows comparison across a diverse range of health programs that have effects on all types of health states for a range of conditions. Examples of CUA study design include the comparison of 'surgery plus radiotherapy' versus 'surgery plus radiotherapy plus chemotherapy' treatments for cancer, and the comparisons across programs for hip and heart problems, depression and diarrhoea, or prevention and palliative care.

Quality adjusted life year (QALY) is the most commonly used measure of health related utility. In its simplest form, QALY weights a life year by a utility index reflecting the health related quality of life. Such utility index is often expressed as a fraction between zero (death) and one (full health). It may be assigned negative values for health states that are regarded as being worse than death — being bedridden with severe pain. Box 3.2 illustrates the QALY procedure and table 3.3 shows some QALY estimates for different health states.

Box 3.2 **An illustration of the QALY procedure**

Haemodialysis provides a potentially life saving treatment for end stage renal failure, but additional life years are commonly associated with poorer quality of life. If a haemodialysis patient lives an additional period of 20 years, but those years are judged to be worth only 60 per cent as much as years spent in full health, the additional life years are discounted as 12 (= 20 × 0.6) QALYs.

Table 3.3 **An example of health state ratings**

Health state	*Relative utility level*
Healthy (reference state)	1.00
Life with menopausal symptoms	0.99
Side effects of hypertension treatment	0.95–0.99
Kidney transplant	0.84
Some physical and role limitation with occasional pain	0.67
Hospital dialysis	0.56–0.59
Anxious/depressed and lonely much of the time	0.45
Being blind, deaf or dumb	0.39
Hospital confinement	0.33
Dead	0.00
Confined to bed with severe pain	<0.00
Unconscious	<0.00

Source: Torrance (1987).

Besides QALY, there are numerous other health related utility measures available. Examples include the Assessment of Quality of Life (AQOL), the Euroqol 5 Dimension (EQ5D), the Healthy Year Equivalent (HYE), the Save Young Life Equivalent (SAVE), and the SF6D instruments. Reviews of various outcome measures can be found in Gold et al. (1996), Brazier et al. (1999), and Nord (1992).

While CUA allows for comparisons of dissimilar health programs, it cannot be used to compare between health and non-health sector activities. This is because utility-based outcome measures do not capture changes in overall wellbeing of individuals or the community, given that health is only one aspect of social wellbeing.

Cost–benefit analysis

CBA study designs use a single metric — money, most commonly — for measuring all potential benefits from alternative programs, whether health or non-health related. This allows comparisons across all health care activities as well as non-health sector activities.

A wide range of CBA applications are possible. First, CBA can address both technical and allocative efficiency concerns in the production of health gains. Second, it may be used to compare programs across sectors — housing, education, transport, and health, for instance. Third, by comparing the monetary valuations of benefits and opportunity costs, in principle, CBA can be used to determine the optimal size of each sector to the point that no further wellbeing gains can be achieved from shifting resources between sectors. Practical applications of CBA to address such macro-level allocation problems are rare. But the technique has had numerous applications in comparing programs within the health sector (Mushkin 1962; Klarman 1965; Weisbrod 1971).

CBA has not been used as frequently as CEA and CUA because of the difficulty in obtaining acceptable monetary valuations for health outcomes and, in particular, lives. The latter problem has been the most controversial issue in the economic evaluation literature. Two competing valuation methods have been mostly employed in CBA studies:

- human capital approach; and
- willingness-to-pay approach.

Human capital approach

Under this approach, the value of a human life is treated as external to the individual. It is measured in terms of the present value of the individual's expected

lifetime earnings, in the same way that productive capital is commonly valued (Robinson 1986). The approach is appealing primarily because it is simple and requires relatively few data to implement.

However, the approach has drawn a number of serious criticisms:

- The approach discriminates against less productive people, particularly those who are disadvantaged due to their race and education attainment.
- The focus on productive value is too narrow, ignoring an individual's value to family, friends, and the community.
- The strength of an individual's preferences is not accounted for.

McKie et al. (1998) consider the above criticisms fatal to the human capital approach. They suggest that the technique, at best, is defensible as a measure of a person's contribution to Gross Domestic Product, and that this is only a part of the value of a human life.

Willingness-to-pay approach

Monetary valuations of outcomes in CBA have mostly been based on willingness-to-pay techniques. This approach uses an individual's *stated preferences* — not *revealed preferences* — to determine the relative values of program outcomes. Stated preferences are responses to hypothetical questions, while revealed preferences require observing an individual to actually trade good health for poor health. Because there is no market for human life, valuation of health outcomes is necessarily based on stated preferences (box 3.3).

Box 3.3 **Valuation of health outcomes based on stated preferences**

Information on stated preferences is mainly obtained from individuals' responses to questions about their willingness to pay for decreases in the risk of death and morbidity, or required compensation for increases in the risk of death and morbidity. For example, if an individual is prepared to accept $10 000 for one per cent increase in the risk of death, the imputed value of a life is $1 million (= $10 000 ÷ 1 per cent).

Alternatively, econometric methods may be used to derive the value of life from income data for workers in 'high risk' industries, such as high-rise steel workers, firemen and policemen.

Estimates for the value of life and health states are prone to significant influence from model specification and statistical procedures applied. Besides, as McKie et al. (1998) point out, it is problematic to extrapolate the value of life from the context of a low risk of death. The contexts of low and high risks of death may be extremely

different, and affect individuals' preferences in different ways. Those preferences are, by their nature, subjective and susceptible to such factors as hope, fear and anxiety, which vary with the risk level. The authors conclude that willingness-to-pay techniques based on a simple extrapolation of the value of human life are flawed and at best measure the value of life contaminated by the value of risk.

A comparative appraisal

All economic evaluation techniques employ outcome–cost ratios for comparing a specific set of health programs. Their main differences are in the study perspective and the valuation of health outcomes. Such differences lead to a trade-off between the scope for potential use and the practicality of various evaluation methods.

In particular, CBA is capable of comparing technical efficiency between alternative health programs, and addressing allocative efficiency concerns between health and non-health activities. But measuring health outcomes in money terms remains problematic and controversial. As a consequence, CBA is not often used in the health field despite its potential. By using a generic measure of health related utility, CUA is useful for comparing health programs over a range of health conditions. By comparison, CEA and CMA represent even more restrictive forms of assessment, but are more popular by virtue of their simplicity in comparing alternative health programs that achieve particular health outcomes.

The caveat to the above discussion is that all evaluation techniques used in the health sector, whether ad hoc or based on explicit economic and ethical criteria, necessarily imply or impute a value for life or health conditions. When resources are scarce and there is a possibility of saving life, even at tremendous cost, placing a value on life is inevitable. Indeed such valuations are routinely made and have been made for many years. When decisions are made not to devote resources to a life saving activity, a value is instantly placed on life. For instance, the decision not to fund a subsidised seat belt campaign for coaches either implicitly or explicitly involves comparing the value of life — that is, the additional lives saved and the morbidity avoided through reduced injuries in road traffic accidents — to the opportunity cost of using those resources elsewhere. Equally, expenditure is frequently allocated to activities that increase the risk of death. The difference between CBA compared to CUA, CEA, CMA, and ad hoc decision rules is not that a money value is placed on life in CBA, but that CBA attempts to do so in an explicit and transparent manner.

4 Benchmarking analysis of service providers: illustrated by hospital studies

Unlike economic evaluation comparing a specific set of health programs under certain laboratory conditions, benchmarking analysis of service providers accounts for the diversity of service activities performed and the operational aspects in organising and administering such activities.

This chapter examines a number of benchmarking techniques for their adaptation to health services and health production. These techniques are grouped as:

- simple ratio analysis;
- unit cost analysis;
- stochastic frontier analysis (SFA); and
- data envelopment analysis (DEA).

These techniques have in common the concept of frontier, although not all of them involve the estimation of an explicit frontier equation — only SFA and DEA do so. Techniques based on the concept of frontier conform to the definition of productive efficiency. Efficient producers are identified to be those operating on the estimated production or cost frontier, whereas inefficient producers operate below the production frontier or above the cost frontier. By contrast, non-frontier regression techniques assume that all observations are efficient but exhibit nonsystematic variations in performance. These techniques estimate the central tendency of cost and production relationships rather than pinpoint the level of efficiency for individual producers. They are not recommended for benchmarking purposes (Sherman 1984; Barrow and Wagstaff 1989) and are excluded from this study.

In this chapter, a collection of hospital studies are examined to illustrate the major methodological issues in benchmarking analysis, particularly the treatment of heterogeneous services. Among the published studies of efficiency in the health field, hospitals have received most attention. This is not surprising given the role of hospitals as a key health care delivery mechanism and the accessible data on

hospital operations. To a degree, the methods for measuring efficiency in hospitals can be extended to other types of service providers once the required data become available.

Specifically, the selected benchmarking studies of hospital efficiency are examined in two respects:

- whether the technique is equipped to capture the essential features of hospital services, including the pattern of joint production and the influence of external factors on productive performance; and
- whether the technique is applied to derive efficiency measures consistent with the performance of service providers in attaining the greatest possible health improvements for service recipients.

These two main criteria are elaborated in sections 4.1 and 4.2 respectively. Section 4.3 discusses two other issues in specifying hospital services, namely the measurement of inputs and the adjustment of efficiency for environmental factors. Section 4.4 presents a few benchmarking studies of domestic hospitals.

4.1 Benchmarking techniques

The various benchmarking techniques adopt different ways to:

- represent joint production with multiple input and output variables;
- account for the effect of external factors on productive performance; and
- measure efficiency using either a single indicator or a suite of indicators.

Simple ratio analysis

Each simple ratio is limited to a comparison of two variables, one measuring an input quantity and another measuring an output quantity. To deal with situations where multiple outputs are produced using multiple inputs, a suite of ratios are calculated. The selection of input and output variables, and hence partial efficiency indicators, depends largely on the intended breadth of analysis and data availability.

By itself, ratio analysis cannot identify the appropriate set of comparators for a particular observation. Like-with-like comparisons rely on the grouping of comparable production units based on prior information about them. To determine any effect of environmental factors on efficiency, ratio estimates need to be analysed using other statistical tools — such as regression or variance analysis.

Moreover, ratio analysis provides no objective means to integrate multiple ratios into a single indicator of efficiency and thus can be inconclusive in making efficiency comparisons. One production unit may appear relatively efficient according to one group of ratios but inefficient according to another. Individual ratio estimates are subject to changes in the input or output mix not reflecting any variation in the level of productive efficiency, such as the substitution of inputs for producing the same amount of outputs.

However, ratio analysis is a useful precursor to other sophisticated techniques. Ratio estimates provide a preview of data structure and may bring insights into the relationship between data patterns and efficiency estimates. For example, DEA calculates the efficiency score as a function of individual ratio estimates (discussed below). Besides, by virtue of its simplicity and minimal requirement on the sample size — ratio comparisons are possible for as few as two observations, ratio analysis can be the only possible form of efficiency analysis when data are scarce or highly aggregated.

Unit cost analysis

Unit costs are input costs divided by an output variable. A unit cost measure can be viewed as a single indicator of efficiency using input prices as weights to combine multiple ratios of different inputs to a particular output. Depending on the coverage of inputs, unit cost measures may relate to recurrent costs, capital costs, or total costs. Multiple outputs need to be aggregated in calculating the unit cost for the whole process of joint production.

This approach requires price or cost data, which are harder to collect on a consistent basis than quantity data required for ratio analysis. Price differences may cause variations in unit cost estimates to be mistaken for efficiency changes. Where factor markets are not competitive — such as the regulated market of physician services, unit cost estimates are affected by price distortion and, as a result, may become biased indicators of efficiency. Jacobs and Baladi (1996) identify various kinds of bias in common hospital cost measures.

Stochastic frontier analysis

SFA employs advanced econometric methods to estimate a stochastic frontier equation, which is either a production or a cost function with an error term.

Several factors affect the choice between estimating a production or cost function. Direct estimation of the production function is hampered by econometric problems

arising from the correlation between different input variables and between any input variable and the error term. Further, the production function includes only one output as the dependent variable and thus requires an aggregated output measure to represent multiple outputs. As noted by Lovell et al. (1987), output aggregation amounts to the imposition of restrictive conditions on the estimated frontier equation and may affect efficiency comparisons.

By assuming cost minimising (but not necessarily profit maximising) behaviour of producers, the production function can be transformed into the cost frontier retaining all the original production technology characteristics. The cost frontier relates total cost to outputs, input prices, and environmental factors. To estimate the cost frontier, output aggregation is not required and correlation of output variables is less of a computational problem. This is an advantage particularly for the analysis of hospitals as they have limited ability in choosing the level and mix of service outputs. Input prices are assumed to be statistically independent from other explanatory variables. But data on input prices and total cost are needed.

A variety of functional forms may be used to specify the production or cost function, such as the translog function and other variants. They differ in the specification of model parameters and the error term. Model parameters are specified reflecting prior knowledge and econometric evidence of production characteristics. The error term is composed of two components. The first component has a two-sided distribution capturing data measurement errors and random shocks to the production process. It explains any transitory deviation from the performance target projected using the frontier equation. The second error component has a one-sided distribution, measuring the level of efficiency relative to the estimated frontier.

SFA offers rich specification of the production process, including environmental factors. It also allows for statistical testing of hypotheses about the production frontier and constructing confidence intervals around the efficiency measure.

However, SFA has two major limitations. The first is the requirement of a large sample size in order for the applied econometric technique to work well. The second is the reliance on an assumed distribution of efficiency estimates. The distinction between inefficiency and transitory deviation from performance targets crucially hinges on the assumed probability distributions for the two error components. As Coelli (1995) points out, prior justification or powerful statistical tests do not always exist to guide the selection between alternative patterns of probability distribution. Efficiency or otherwise may be determined largely by a nontestable assumption about the error term. For instance, a narrowly dispersed probability distribution assumed for the efficiency term would lead to small estimated efficiency differences.

Data envelopment analysis

Fried et al. (1993) and Charnes et al. (1994) provide details of DEA methodology and applications. This section examines the main strengths and weaknesses in relation to hospital studies.

DEA employs mathematical programming methods to estimate a deterministic production frontier. It can derive a single indicator of efficiency while allowing for multiple output and input variables and without using price or cost data.

DEA does not specify a functional form for the production frontier. This suits a complicated production process that is unlikely to be well described by any mathematical equation — as in the case of hospital services. Using the linear programming method, the frontier is estimated as a piecewise linear equation, which provides a close approximation for any unspecified, flexible production function (Banker 1993). However, without the aid of an econometric procedure, DEA requires data to be free of measurement errors and cannot estimate confidence intervals around the efficiency measure.

DEA can be viewed as a form of *activity analysis*, in which the production process is characterised by discrete input and output levels. This means that simple ratios are instrumental in providing a detailed description of the production process. According to Jomini and Chan (2000), DEA measures efficiency as a function of individual output–input ratios. As a sophisticated form of ratio analysis, DEA is superior to simple ratio analysis mainly for its built-in ability to make like-with-like comparisons. The modelling procedure divides the sample into different peer groups each containing observations of similar input structure and output composition. Efficiency comparisons are made within and between such peer groups.

In addition, the technique provides several means to analyse the relationship between operating environments and efficiency. Relevant environmental factors are captured through categorical or continuous input variables in DEA (Banker and Morey 1986). Alternatively, posterior regression analysis may be applied to relate the efficiency measure to these environmental factors (Fried et al. 1999).

Two weaknesses of DEA may potentially affect the usefulness of some hospital studies using this technique. The first problem is peculiar to linear programming, as this method tends to estimate efficiency in two parts — an efficiency score and slack variables. The efficiency score indicates the scope for contracting all inputs or expanding all outputs by the same proportion. Slack variables arise when, after the proportionate contraction of inputs or expansion of outputs, it is still possible to reduce some inputs or increase some outputs in isolation. In other words, efficiency scores alone do not reveal the full magnitude of inefficiency. For instance, a high

efficiency score associated with large slack variables does not necessarily indicate better performance than a low efficiency score with small or zero slacks. The hospital study by Webster et al. (1998) is an example where slack variables dominate efficiency comparisons. To address this problem, Thrall (2000) modifies the efficiency measure to combine all forms of inefficiency. Fried et al. (1999) propose that, to account for the effect of environmental factors, both the score and slack variables should be adjusted. However, Førsund (1998) argues that slacks appear due to the default extrapolation of certain segments of the production frontier and would diminish by incorporating more information about the production technology in the programming procedure.

The second problem with DEA is related to the inclusion of meaningless output–input ratios in the efficiency measure. In hospitals, some inputs are shared among different services while other inputs have specific uses. However, by including multiple input and output variables, DEA allows any number of output–input ratio(s) to determine the efficiency score. This highlights the possibility of measuring efficiency by any arbitrary ratio that is not a sensible partial indicator of efficiency. Table 4.1 contains some examples of meaningless output–input ratios that are included in various DEA studies of hospital efficiency. To better represent the pattern of input sharing in measuring efficiency, Jomini and Chan (2000) suggest DEA models be modified to exclude irrelevant partial efficiency indicators from the efficiency measure. Cook et al. (2000) propose deriving separate efficiency measures for different service areas and using such component-based efficiency measures to form the aggregate efficiency score.

Table 4.1 **Examples of meaningless output–input ratios**

Output–input ratio included in DEA	*Hospital study* [a]
Psychiatric discharges per pediatrics bed	Hogan and Wroblewski (1993)
Surgery inpatient days per total admission	Webster et al. (1998)
Emergency room visits per birthing room	Hogan and Wroblewski (1993)
Newborns per dollar of teaching expenditure	Morey et al. (1995)
Outpatient days per bed	Magnussen (1996); Ferrier and Valdmanis (1996); Hogan and Wroblewski (1993); Mobley and Magnussen (1998); Holvad and Hougaard (1993)

a See appendix D for a detailed list of input and output variables included in these studies.

Source: As quoted.

Comparative advantages of DEA

Compared to other techniques under review, DEA is the most versatile partly for its flexible representation of joint production and partly for its ability to account for factors that have potential effects on productive efficiency. Further, DEA models

are highly tractable, particularly for the close link with simple output–input ratios. However, unlike SFA, DEA does not allow for statistical noise in the production process. But the estimation of a stochastic frontier is feasible only at a price, namely the tautological relationship between the assumed error distributions and efficiency estimates derived from SFA. As judged by Hollingsworth et al. (1999), on balance, DEA is probably the most appropriate technique currently available for measuring efficiency in health services. Nevertheless, due attention should be directed to a number of modelling problems identified in previous DEA applications, such as the implication of slack variables for a comprehensive assessment of efficiency and the specification of input sharing patterns.

4.2 Relating hospital efficiency to treatment outcomes

The definition of output or outcome variables in various hospital studies determines to what extent efficiency is measured in line with the objective of improving patient health. This section examines the following purposes of these model variables:

- classify hospital services on the basis of patient illness characteristics;
- aggregate service outputs by weights of case complexity;
- measure quality of care; and
- measure health gains to patients.

The discussion draws on a selection of hospital studies that are summarised in appendix D.

Classification of hospital services

Many hospital studies include gross measures of hospital services, such as the total numbers of patient days, patient cases, surgeries performed, and discharges. These highly aggregated variables bias efficiency comparisons against hospitals that treat a relatively high proportion of severely ill patients.

Other studies control for patient health conditions mainly in measuring acute care services. Patients are classified by a Diagnosis Related Group (DRG) system that identifies patients with similar illness conditions and care procedures (box 4.1). The DRG system currently used in Australia covers about 600 patient classes. Nevertheless, nonrandom variations in patient illness conditions exist within a DRG class. As a case in point, teaching hospitals may treat a higher proportion of patients within a particular DRG with more severe conditions than non-teaching hospitals. To define homogenous outputs, some hospital studies subdivide DRG classes using

additional clinical information. Age is often considered to be a key factor affecting resource use for treating patients. Discharge destination — including death, home, nursing home, and other health care facilities — is identified as one dimension of hospital outputs. Clinical indicators of comorbidity and treatment complications have also been used to refine the DRG classification. For example, Long et al. (1990) apply the Body System Count Methodology for gauging the number of body systems diagnosed as requiring treatment.

Box 4.1 **Casemix classification**

The DRG system was initially developed in the United States and has now been used extensively for hospital reimbursement and analysis purposes. It reflects the distribution of patients — that is, the casemix — based on morbidity, patient characteristics such as age, and aspects of the treatment received. Each DRG represents a class of patients with similar clinical conditions, hence requiring similar resource usage for treatment.

Patient-class-specific cost estimates are derived from cross-section regression studies of discharge records or applications of information theory (Butler 1995; Hardwick 1986). These cost weights are averages for a group of hospitals as determined under a funding scheme, and may differ from actual costs incurred by individual hospitals.

Methods and sources of information for preparing casemix data in Australia can be found in Palmer and Wood (1984), Palmer (1986, 1991), and Eager and Hindle (1994).

For non-acute services, casemix information has yet to be developed. These services are mostly measured by simple counts, such as the number of outpatient visits.

Box 4.2 illustrates the treatment of teaching and research functions in hospital studies. Box 4.3 describes the deficiency of using output measures in comparing hospital efficiency.

Box 4.2 **Teaching and research**

Popular output measures for teaching and research are the numbers of trainees and research publications, respectively. As an exception, Morey et al. (1992) and Morey et al. (1995) consider trainees as an input to hospital services, and measure teaching output by the value of medical education provided. One way to avoid using an input-based output measure for teaching is to include a dummy variable indicating the teaching responsibility as part of the operating environment of hospitals. But this approach makes no allowance for variations in the level of such activity that is taking place. Linna (1998) represents teaching as both an output and an operational feature.

Box 4.3 **Outputs or intermediate inputs?**

Whatever definition of outputs is adopted, output variables should not be input-based. Otherwise, no meaningful measure of efficiency can be developed.

In the context of hospital services, output measures are intermediate inputs to the production of health. They have appeared either as input or output variables in different studies. For instance, the number of hospital days is a common output measure but the length of stay in hospital is sometimes taken to gauge resource consumption. While the numbers of patient admissions and discharges may be obvious output variables, a few studies break with convention by including them as input variables. Likewise, the definition of teaching output in hospital studies is ambiguous, as discussed in box 4.2.

Aggregation of service outputs and adjustment for case complexity

Patient classes are typically consolidated into broad categories prior to modelling. However, the level of aggregation varies between studies. Some studies employ a single casemix adjusted output variable to represent all inpatient cases. Others include several casemix adjusted output variables for different broad patient categories (box 4.4). Long et al. (1990) is one exception in which detailed casemix data are analysed without output aggregation.

Box 4.4 **Examples of casemix adjusted output variables**

- Sherman (1984) defines two output variables measuring the number of hospital days for inpatients less than and over 65 years of age, respectively.
- Hogan and Wroblewski (1993) aggregate acute cases across 470 DRGs into 6 broad categories: (i) pediatric discharges; (ii) adult medical discharges; (iii) adult surgical discharges; (iv) obstetrics discharges; (v) psychiatric/substance abuse discharges; and (vi) newborn discharges. These broad categories are each subdivided on the basis of clinical indicators, yielding a total of 26 output variables.
- SCRCSSP (1997) classifies acute services in three categories reflecting resource intensity: (i) minor; (ii) moderate; and (iii) complex. The number of discharges in each of these categories is adjusted for casemix variation using the Weighted Inlier Equivalent Separations (WIES) index, which is a series of DRG cost weights adjusted for the expected patient length of stay.

Sources: Sherman (1984); Hogan and Wroblewski (1993); SCRCSSP (1997).

Palmer (1991) compares the aggregation of patient cases to the calculation of total value of hospital services, where 'output prices' are measured in terms of health gains instead of dollar amounts. For instance, a cancer should be assigned a larger

weight of health gains than a leg fracture. In practice, output aggregation is commonly weighted by DRG cost data. These cost data primarily indicate the average level of case complexity, hence resource intensity, rather than relative health gains due to the treatment (Palmer 1986). To the extent that cost weights are not proportionate to health gains — for instance, as a result of not using cost-effective care procedures, casemix adjustment introduces bias to the evaluation of hospital efficiency in respect of health production. A hospital may be assessed as efficient in maximising reimbursement under the DRG funding scheme but not necessarily efficient in improving patient health.

Among studies that take simple counts of patient cases, crude indicators of case complexity are often included. The most common examples are the casemix index and the high-technology index. The casemix index is the DRG-weighted average number of admissions. The high-technology index measures the scope of high technology services offered in a hospital, including electrocardiography, open heart surgery, and others. Less frequently used proxy indicators of case complexity are the availability of cancer program, the average length of stay, and the proportion of emergency visits in total patient cases.

Measurement of quality of care

Quality of care refers to aspects of hospital services that affect the process of care and the effectiveness of hospital treatment (box 4.5). Two types of crude measures are used: general outcome indicators and process-of-care indicators. Examples of hospital-wide outcome indicators are unplanned readmission rates, hospital misadventure rates, and unexpected death rates. A variety of process-of-care indicators have been used, including the degree of specialisation, the adoption of clinical guidelines, the number of daily visits by doctors, the existence of appraisal programs for professional staff, and the extent of certification for medical staff.

Measurement of health gains to individual patients

Health gains to individual patients are measured by changes in health status between admission and discharge. A patient's health condition prior to hospitalisation is assessed using an illness severity scoring procedure (box 4.6). Treatment outcomes are assessed in terms of the survival status at discharge (including death) or the recovery from illnesses or injuries to a normal health status.

Take Puig-Junoy's (1998) DEA study of a sample of intensive care units (ICUs) to illustrate. The survival probability of each critically ill patient is modelled as a non-controllable input to the ICUs. It is determined prior to admission by a clinical

scoring procedure. Two outcome variables are included, measuring the number of days surviving in the hospital and the discharge status for individual patients respectively. The latter is a binary variable relating to the exclusive outcome between death and survival. Using these outcome variables, efficiency is measured in relation to the attainment of health gains — that is, the number of lives saved and the length of life prolonged — for a given amount of resources used and after adjustment for the initial patient health status (table 4.2).

Box 4.5 **Profiling quality of care in hospital services**

Quality of care is embodied in various aspects of hospital services, including the structure, the process, and the treatment outcomes.

The first two aspects indicate resource utilisation rather than service efficiency, and are based on the assumption that resources and practice patterns always contribute to health improvements. Structural indicators measure the availability of resources, such as staff per bed. Process indicators focus on the compliance with accepted standards of hospital care. Inpatient process indicators monitor the surgical approach, the use of medication, and the type of diagnostic tests performed. For ambulatory care, the assessment involves monitoring whether diagnostic tests, procedures or referrals are carried out as recommended.

The outcome aspect refers to the effectiveness of treatments at the hospital level. Outcome indicators include the rates of in-hospital mortality, complications, and misadventures. These measures are calculated on a general or disease-specific basis. General outcome indicators integrate all aspects of a patient's care into one measure, such as mortality and readmission rates. Disease-specific outcome measures monitor a particular part of the caring process, like the dislocation for joint replacement points.

Partial information on various countervailing arrangements is often used to infer service quality. These arrangements include accreditation, codes and ethics, and mandated or voluntary quality assurance schemes.

Sources: McNeil et al. (1992); Folland et al. (1997).

Box 4.6 **Patient illness severity**

Illness severity is one dimension of patient health status and refers to such patient conditions as the intensity of disease manifestation, the degree of dependency, and the response to therapy.

'Off-the-shelf' illness severity scoring methods are commercially available in the United States. Landon et al. (1996) compare 14 scoring methods that employ different data sources and procedures to calibrate the level of illness severity. Most of these methods use standard data from hospital discharge abstracts, such as diagnosis and procedure codes. But some of them require detailed clinical information.

Sources: Palmer (1986); Landon et al. (1996).

Table 4.2 **Selected outcome–input ratios included in the efficiency measure for hospital services**[a]

Life of ICU patient saved per physician (=0 for death, =1 for survival)
Days surviving in ICU per physician
Life of ICU patient saved divided by mortality risk at admission (=0 for death, >0 for survival)
Days surviving in ICU divided by mortality risk at admission

[a] See appendix D for a full list of model variables included. Following the analysis of Jomini and Chan (2000), the DEA efficiency measure can be transformed into a function of individual outcome–input ratios.

Source: Puig-Junoy (1998).

Similar outcome measures are used elsewhere for efficiency analysis. For instance, Finkler and Wirtschaffer (1993) assess the outcomes of obstetric services by measuring the risk adjusted foetal and perinatal mortality rates. Landon et al. (1996) and Chassin et al. (1996) compare the risk adjusted outcomes of coronary artery bypass graft surgeries. (These studies are not listed in appendix D, as they use non-frontier estimation techniques.)

Assessing efficiency in terms of health outcomes can capture a systematic form of inefficiency associated with clinical decisions, including the use of ineffective health technology and redundant procedures. Outcome measures come close to reflecting the effect of health care on patient wellbeing (Weisbrod 1992). However, this approach confines the measurement of efficiency to homogeneous hospital services and requires in-depth clinical knowledge for preparing health data.

4.3 Other issues in specifying hospital services

This section discusses the following issues:

- how to represent inputs to the provision of hospital services; and
- how to represent patient characteristics and hospital operating environments.

Measurement of hospital inputs

In principle, by accounting for all inputs used (as well as outputs produced), efficiency is compared in a comprehensive manner. In practice, however, variables are selected to represent key inputs driving efficiency, assuming that they are used in certain fixed proportions to the omitted inputs. The following discusses input variables in three categories: labour, materials and hired services, and capital.

Labour

The input of labour is often measured in terms of staff number. But such a measure of employment reveals little about the use of the workforce, as it obscures the mix of full-time, part-time and casual workers as well as staff working overtime. The number of hours worked is the preferred unit of measurement. Where work hours are not recorded, the number of full-time equivalent staff is probably the best way of measuring labour. In lieu of quantity measures of labour, some studies use labour cost measures, either assuming uniform wage rates across the whole sample of hospitals or adjusting nominal labour costs by wage indexes.

By classifying labour into different categories, many studies recognise the differing skill requirements for nurses, physicians, specialists, administrators, and ancillary staff. A few studies also include a separate input variable for interns and trainees. Visiting or contracted physicians have been accounted for in several ways — as a separate input variable, combined with other hospital staff, or included as part of the hired services.

Materials and hired services

It is common to aggregate purchased materials and hired services in one cost variable, as quantity data on individual items are scarce. But a few studies manage to separate energy and drug use from other expenses.

Capital

Capital input refers to the service flow derived from a durable capital stock. The annual cost of capital has two components: depreciation and real interest cost of current capital value. Under the assumption that service flow is directly proportional to stock quantity, capital stock may be used to represent capital input.

Flow measures of capital are based on the calculation of depreciation cost, interest cost, or both. The number of beds is the most commonly used stock variable in hospital studies. Some studies distinguish hospital beds between different service units, such as paediatrics, acute care, ICU, and so on. Other stock variables used are the numbers of birthing rooms, surgical operating rooms, and bassinets. Aggregated measures of capital stock have also been used, including the net plant asset value and the capital index for buildings, plants and equipment combined.

Accounting for patient and hospital characteristics

Factors that affect patient health or constrain hospital operations need to be accounted for in making meaningful comparisons of efficiency between hospitals.

Patients' socioeconomic background

To account for the effects of patients' education and personal efforts on their own health, Hughes and Yaisawarng (2000) differentiate the need for hospital resources between patients in different education and income classes. On the same rationale, Hogan and Wroblewski (1993) impute patients' socioeconomic background from hospital types and whether patients are fee-paying or public-funded.

Hospital operating environments

Operating environments are related to hospital ownership, market structure, funding arrangement, teaching activity, affiliation, location, occupancy rate, scale of operations, and scope of services. These factors affect hospital efficiency through their effects on the cost structure and the patient profile.

Zuckerman et al. (1994) emphasise the cost effect of environmental factors. Hogan and Wroblewski (1993) also account for some of these factors but reason differently. They hypothesise that hospitals in urban locations tend to attract severely ill patients. Also, public patients are assumed to be older or economically disadvantaged and hence harder to treat than private patients, other things being equal. In the same vein, church-affiliated hospitals are thought to attract underprivileged patients who likely require resource intensive treatments.

4.4 Benchmarking studies of hospitals in Australia

Sophisticated benchmarking techniques have been used extensively for evaluating hospital efficiency overseas. Hollingsworth et al. (1999) identify nearly 90 DEA applications to hospitals in the United States and Europe. For Australia, however, benchmarking studies of hospitals are few.

Table 4.3 summarises the results of three studies sampling hospitals in Australia: SCRCSSP (1997), Hughes and Yaisawarng (2000), and Webster et al. (1998). Appendix D shows details of the sample coverage, the technique(s) applied and the model variables included. These studies each look at one segment of the hospital sector and all have limitations. They offer no conclusive evidence on the comparative level of efficiency in domestic hospitals.

Table 4.3 **Selected benchmarking studies of hospitals in Australia**

Study	*Sample*	*Technique*	*Technical efficiency scores* [a]
SCRCSSP (1997)	109 public acute care hospitals in Victoria in 1994–95	DEA	*Metropolitan/large country hospitals* Mean: 0.90 to 0.92 (various models) Lowest: 0.63 to 0.71 (various models *Small country hospitals* Mean: 0.68 to 0.86 (various models) Lowest: 0.44 to 0.48 (various models)
Hughes & Yaisawarng (2000)	70 public acute care hospitals in NSW in 1995–96	DEA	*Metropolitan hospitals* Mean: 0.98 Lowest: 0.87 *Non-metropolitan hospitals* Mean: 0.96 Lowest: 0.79
Webster et al. (1998)	301 private hospitals in Australia in 1991–92 to 1994–95	DEA	Mean: 0.39 to 0.90 (various models) *Efficiency changes* Mean: 0.67 to 0.91 per cent per year (various models) *Technological changes* Mean: 2.40 to 9.50 per cent per year (various models)
		SFA (production function)	Mean: 0.71 to 0.79 (various models)
		SFA (cost function)	Mean: 0.77 to 0.96 (various models)

a Efficiency scores are relative to a best practice score of one. Some models produce efficiency estimates on a different scale of measurement; those estimates have been rebased in this table.

Sources: SCRCSSP (1997); Hughes and Yaisawarng (2000); Webster et al. (1998).

Several points can be made in comparing the model designs of these studies. First, while Webster et al. (1998) experiment with different benchmarking techniques, model specifications and data definitions, such sensitivity tests cannot decisively determine the appropriateness of any of the models tested. An appropriate technique and model specification should be chosen in light of a sound understanding about the nature of hospital services and the properties of alternative techniques. Second, SCRCSSP (1997) and Hughes and Yaisawarng (2000) include casemix adjusted output variables, while Webster et al. (1998) measure hospital services in gross terms only. Third, SCRCSSP (1997) is the only study that relates the measurement of efficiency to treatment outcomes measured at the hospital level. But the study by Hughes and Yaisawarng (2000) is unique for its detailed analysis of environmental factors. Fourth, the DEA results from Webster et al. (1998) are potentially impaired by the inclusion of meaningless output–input ratios and the dominance of input slacks. These potential shortcomings point to areas for further research.

5 Towards an integrated approach

Until recently, economic evaluation and benchmarking analysis have remained two largely separate approaches to measuring efficiency in health services. Several factors are responsible for this, including the evolution of these techniques from different research disciplines and the lack of common data for their application. However, with some of the recent advances in estimation methodology and data collection outlined in chapters 3 and 4, there is now scope for integrating both techniques into a unified efficiency analysis framework. This integrated framework enables outcome-based analyses of efficiency to match with output-based analyses, and micro-level studies to match with macro-level studies. The benefits of this integration are to:

- indicate under what conditions output-based assessments are adequate proxies for the performance of a health care system in improving the health status of the population (not just in carrying out particular health care procedures); and
- identify the specific causes of inefficiency that are attributable to efficiency differences measured at an aggregate level.

This chapter explores the linkages between benchmarking and economic evaluation techniques (section 5.1). Since both techniques account for inputs in similar ways, the linkages are primarily based on the ability to reconcile the use of output measures in benchmarking analysis and the use of outcome measures in economic evaluation. The techniques offer synergies in that economic evaluation:

- identifies which procedures are cost-effective, which is one source of potential improvement in health care efficiency; and thereby
- provides a basis for determining the appropriate weights for aggregating different service outputs in benchmarking analysis.

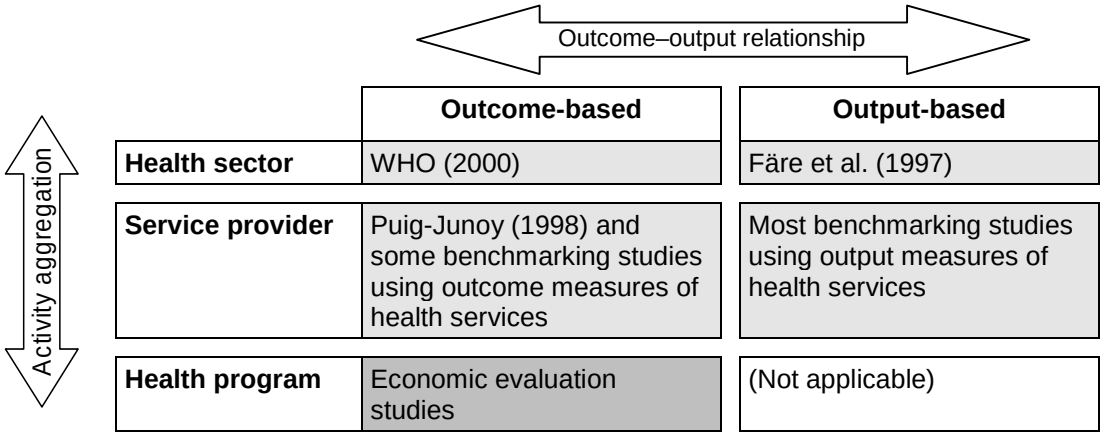
The discussion underscores a number of key findings addressing the objectives of this study as set out in chapter 1 (section 5.2).

5.1 Linkages between measurement techniques

Studies of efficiency in the health sector can be linked in two dimensions (figure 5.1):

- through the cascade of activity levels; and
- through consistent comparisons of health care efficiency based on the measurement of service outputs and health outcomes.

Figure 5.1 **Linking economic evaluation and benchmarking analysis**



Economic evaluation studies of individual health programs are outcome-based. Benchmarking studies of service providers are mostly output-based; but some have used measures of health outcomes. An implicit assumption ties outcome- and output-based efficiency studies. The assumption is that cost-effective treatments are applied and that the cost estimates used for aggregating outputs reflect the lowest possible costs for obtaining health outcomes. If this assumption is not satisfied, the output-based approach yields different efficiency estimates from the outcome-based approach.

The output–outcome relationship between economic evaluation and benchmarking analysis can be established using a model of health production as introduced in chapter 2. The analysis in Garber and Phelps (1997) shows that economic evaluation provides the conceptual basis for *converting* and *aggregating* service outputs into an outcome measure (box 5.1).

Theoretical basis of the outcome–output relationship

An outcome measure captures the combined effects of health services on patients' survival probabilities and quality of life during their lifetime. This allows for comparisons of health services that generate outcomes (or incur costs) at different points in time and comparisons of different service areas — such as acute and preventive care. An output is a measure of treatment procedures performed and typically a measure of throughput in the context of health production.

By analysing allocative efficiency, economic evaluation can determine the *implicit relative output prices* for different health services, which are equal to the relative health outcomes from cost-effective care procedures used in delivering these health services (equation 3 in box 5.1). In deriving the implicit output prices, economic evaluation mimics a competitive market mechanism, which does not exist in the health sector in reality. If cost-effective care procedures are used, health outcomes are proportional to unit service costs, and so are implicit output prices (equation 4). In this circumstance, unit service costs are appropriate for output aggregation purposes in benchmarking analysis (equation 6). Further, the resulting aggregated output measure is proportional to the sum of health outcomes from these services (equation 7). Conversely, the adoption of non-cost-effective care procedures means that service costs would overstate health outcomes (equation 5).

Box 5.1 **Relating outcomes to outputs: an example of disease treatment**

Using the framework in section 2.2, a patient's wellbeing is maximised under two conditions:

- First, for each disease, a care procedure evaluated with the highest outcome–cost ratio is adopted — that is, cost-effectiveness or technical efficiency is achieved:

$$\left(\frac{\Delta H_{ik}}{C_{ik}}\right) \leq \left(\frac{\Delta H_i}{C_i}\right)^* \qquad (1)$$

  where ΔH_{ik} is a measure of health outcomes from applying treatment procedure k to disease i and C_{ik} the corresponding incremental cost, and * indicates a cost-effective care procedure.

- Secondly, across different diseases, the outcome–cost ratios associated with the cost-effective care procedures are all equal at the margin — that is, allocative efficiency is achieved:

$$\left(\frac{\Delta H_i}{C_i}\right)^* = \alpha \quad \text{for all } i \qquad (2)$$

where α reflects the marginal health benefit of an extra dollar spent on treating disease i, which is required to the same across all cost-effective treatments.

The implicit relative output price between the services for treating diseases i and j can be expressed as a ratio of the relative contributions of respective cost-effective procedures to the patient's health:

$$\frac{C_i^*}{C_j^*} = \frac{\Delta H_i^*}{\Delta H_j^*} \qquad (3)$$

(continued)

Box **5.1 (continued)**

Equation (2) implies a proportionate relationship between outcomes and costs for cost-effective care procedures, where α^{-1} converts health outcomes (for example, improved survival probabilities for the patient) into dollar values:

$$\alpha^{-1} \Delta H_i^* = C_i^* \quad (4)$$

From (1) and (2), however, the benefit from a non-cost-effective procedure, expressed in terms of the implicit price for the cost-effective treatment, is less than the cost:

$$\alpha^{-1} \Delta H_{ik} < C_{ik} \quad (5)$$

The outcome measure corresponding to the treatment of a number of patients (measure of output, or Q_i) is proportional to the cost of the cost-effective procedure performed:

$$\alpha^{-1} \Delta H_i^* Q_i = C_i^* Q_i \quad (6)$$

This disease-specific outcome measure can be aggregated across a set of diseases or a set of service providers to derive an aggregate measure of health outcomes (for example, improved life expectancy for the population), which is proportional to the total number of patient cases weighted by the costs of cost-effective treatments:

$$\alpha^{-1} \left(\sum_i \Delta H_i^* Q_i \right) = \left(\sum_i C_i^* Q_i \right) \quad (7)$$

Source: Garber and Phelps (1997).

Linking outcome- and output-based studies

The rationale behind economic evaluation highlights the need for systematic evidence to justify the use and allocation of health care resources, since the market mechanism plays a restricted role in guiding the use of cost-effective care procedures. Through the analysis of allocative efficiency, economic evaluation (CUA, to be specific) can determine the relative worth of health services to service recipients and thereby establish the outcome–output relationship. The adoption of cost-effective care procedures is a prerequisite for consistency between outcome and output based assessment of health care efficiency.

To use outcome measures in efficiency analysis, the main methodological issues are the problems of causal attribution — that is, gauging the change in health conditions due to a particular health service and accounting for the effects of non-health inputs on health production. Different techniques offer different solutions to these problems. Economic evaluation uses clinical trial designs to control for causal relationships, ensuring that observed changes in health status are actually caused by the health program tested and not due to other confounding factors. In

benchmarking service providers, outcome measures may be prepared as part of the clinical procedure to monitor changes in health conditions of service recipients. In both economic evaluation and benchmarking analysis, statistical methods are used to isolate the specific contributions of health and non-health factors to health improvements.

One way to convert outputs into an outcome measure is to aggregate outputs with weights that reflect the relative health gains from these services. DRG cost weights do not suit this purpose, as they may not correspond with the relative outcomes from cost-effective services. For measuring efficiency in disease treatment, a likely promising development lies in the WHO-sponsored Global Burden of Disease and Injury Study, which produces a comprehensive set of estimates of the impact of premature death and disability due to each disease or injury type (HSPH 2001; Murray and Lopez 1996). These estimates have the potential to become model parameters for converting and aggregating classified hospital services into a single outcome measure.

Linking studies at different levels of activity aggregation

For a micro-level analysis, health programs are typically evaluated as the fundamental production process in treating particular health states. At a higher activity level, service providers are analysed as executing and administering a range of health programs, using either output or outcome measures of health services. At a highly aggregate level, the health sector is analysed as a collective organisation of service providers caring for a diverse range of health states. It is also possible to evaluate sector-wide efficiency in the treatment of specific diseases, like the study by Baily and Garber (1997) in which data are aggregated across service providers involved in the treatment of selected disease cases in each country.

Efficiency comparisons between service providers and between health programs are useful for identifying the patterns and sources of inefficiency. The main advantage of comparing efficiency at the sectoral level is comprehensiveness — that is, the inclusion of all health services and their interaction effects. But sectoral assessment may obscure varying performance and the influence of differing environments on performance in different parts of the health sector.

To reconcile the measurement of efficiency at different activity levels, the main methodological issues are about the feasibility of undertaking formal assessments for all health services and of obtaining conformable data compiled at different activity levels. The discrepancy between the cross-country comparison results of WHO (2000) and Baily and Garber (1997) cannot be fully explained unless the latter study extends its scope to include other major, representative disease cases.

However, when measuring efficiency in the provision of homogeneous services, it is easier to link benchmarking analysis and economic evaluation. This is because aggregation is less of a problem and, by using outcome measures, comparable measures of efficiency are used. Puig-Junoy (1998) demonstrates how to measure the efficiency of ICUs in a similar way to economic evaluation of life saving health programs.

Implementing the efficiency analysis framework

The way chosen to implement the efficiency analysis framework reflects the perception of various conceptual and methodological issues, and the vision for developing analytical tools and data to address these issues. The WHO and the NHPC in Australia propose two different ways of implementing the framework.

The WHO defines efficiency in terms of achieving health improvements. It recognises the complementary role of efficiency analysis at the sectoral and micro levels. Results presented so far are for analysis at the sectoral level. Micro-level analysis is being pursued, which involves creating databases on the costs and outcomes of a feasible set of health programs and policy options. It is expected that, in a few years, the micro approach will be used for estimating the frontier of health improvements by country (AIHW 2001).

The NHPC's (forthcoming) proposed framework focuses on the measurement of performance at the sectoral level, although it is also intended to support benchmarking analysis of service providers in different service areas, such as community health, general practice and public health, apart from acute care. Due to the perceived difficulties in measuring health outcomes and attributing them to specific factors, it is suggested that data developments in the short term should cater for the analysis of the output–input relationship. Over a longer period, efficiency measures may be developed to show any association between outcomes and outputs. This could be complemented by the directions identified by WHO (2000) and Garber and Phelps (1997) to establish the outcome–output relationship through the conduct of economic evaluation.

5.2 Key findings of the study

This section summaries the key findings for each of the study objectives.

State of knowledge of efficiency analysis techniques

Economic evaluation of health programs and benchmarking analysis of service providers are two complementary approaches that can be used to identify opportunities for improving efficiency in the health sector (section 2.1).

In the field of economic evaluation, the last decade or so has seen the development of composite outcome measures incorporating fatal and nonfatal conditions into the assessment of health status. This has expanded the scope for comparing dissimilar health programs (sections 2.3, 2.4, 3.3).

Meanwhile, benchmarking techniques have been extensively applied in the study of relative efficiency of service providers, hospitals in particular. Most of the techniques applied are based on the concept of the production frontier. Among these frontier techniques, DEA is considered to be the most useful. DEA is based on a transparent programming procedure that is readily adaptable to the joint-production nature of health services. Further, DEA extends simple ratio analysis and produces more informative efficiency measures than partial efficiency indicators (section 4.1).

Newly developed benchmarking techniques have been applied, complementing the efficiency analysis framework. Outcome studies at both the service provider and sectoral levels were previously thought to be impossible; but new methods have been created in the past few years to estimate the production frontier for health gains. These recent outcome studies reinstate the focus on outcome assessment in the study of health care efficiency (sections 2.4, 4.2).

Methodological issues

The distinctive features of health and health services give rise to specific methodological issues in the measurement of health care efficiency. It is important to distinguish between these issues as they call for different solutions (section 5.1).

A comprehensive analysis of efficiency in the health sector requires measuring efficiency for different service areas and at different activity levels. In this respect, the primary issues are about the representation of joint production and the reconciliation of assessments made at different activity levels and service areas (section 2.1).

On the other hand, the difference between using output and outcome measures in efficiency analysis arises from a conceptual issue — that is, the potentially weak link between service provision and health production. This reflects the particular

characteristics of the health care market, and leads to the ambiguous representation of health services as either outputs or intermediate inputs in different contexts of the production process. An economic perspective requires evaluating health services in terms of their health effects (section 2.2). To implement this, two practical solutions are suggested. The first solution requires using outcome variables to gauge changes in health status due to service provision. Also, the efficiency measure is adjusted for the effect of non-health inputs on health outcomes. This method is mostly applied for evaluating homogeneous services. The second solution allows for the use of output variables to represent heterogeneous services, but incorporates appropriate output weights or broad outcome indicators to qualify the output-based efficiency measure (section 4.2).

Model extensions and data developments

In using DEA, the inclusion of multiple input and output variables in the frontier equation does not guarantee a reasonable description of the joint production process. In particular, the pattern of input sharing may not be adequately described in the frontier equation. Inadequate model specification leads to spurious benchmarking results based on meaningless comparisons of inputs and outputs. The efficiency measure needs to be formulated by restricting any irrelevant output–input ratio. Moreover, slack variables in DEA need to be accounted for in deriving efficiency estimates. The necessary model extensions can be drawn from advanced DEA applications in other fields (section 4.2).

The conceptual and methodological issues identified above point to the importance of collecting and using information on health outcomes for efficiency analysis purposes. In Australia and overseas alike, the existing wealth of operational data and dearth of outcome data means that research efforts should be directed to improving the collection of data on health outcomes. This may require some adjustments in the information systems used. For instance, overseas hospitals have installed scoring procedures to measure the severity of illnesses for individual patients (box 4.6). In Australia, hospital-wide outcome indicators are being developed, pending the establishment of a national reporting system for rigorous epidemiological measurement (NHPC 2000).

Implications for the study of hospital efficiency

For hospital studies, the use of DRG data to classify inpatients is an important development in health statistics. Indeed, rarely has such an elaborate list of services ever been compiled and used in efficiency analysis in any other sector. However, it remains problematic to determine the proper weighting factors for respective DRG

classes, as no systematic, bias-free information is yet available to link treatments to health improvements (section 4.2). A long-term solution depends on the promotion of evidence based and proven cost-effective care procedures as guided by economic evaluation results, as this would align the cost structure of hospital services with the relative health gains (section 5.1). A near-term solution may depend on the WHO's initiative in setting up databases on treatment outcomes for an extensive set of diseases and injuries (WHO 2000).

A Role of health care

This appendix discusses the role of health care relative to other health-enhancing activities in bringing about health improvements.

In a historical perspective, health care interventions have probably played a relatively minor role in improving mankind's health status. As Folland et al. (1997, ch. 4) observe, effective medicine was a fairly recent phenomenon occurring well into the twentieth century. Since the mid-1750s, the world population had begun a distinctively strong pace of expansion as a result of the contemporary dramatic decline in mortality. The reduced mortality is attributed primarily to the spread of knowledge about the sources of diseases, the improvement in living environment, and the greatly increased supply of foodstuffs and sanitary water that became available due to the agricultural and industrial revolutions.

Folland et al. (1997) cite the work of McKeown (1976), a medicine historian, to show that the cure and prevention of infectious diseases accounted for most of the mortality reduction in England and Wales since the mid-1800s and that a major portion of such reduction in mortality occurred before 1900. Another cited study, undertaken by McKinlay and McKinlay (1977), confirms that a large part of mortality reduction in the United States from 1900 to 1973 occurred prior to the emergence of effective medical interventions in the country. Fuch (1974) provides further evidence on the pattern of infant deaths in New York City since 1900. The improved living standards and the spread of literacy and education are believed to be responsible for the marked reduction in infant death rates prior to the introduction of anti-microbial drugs in the 1930s, which accelerated the fall in infant mortality in subsequent years.

A number of econometric studies suggest that, over the past few decades, health care remained a weak contributor compared to other determinants such as education and lifestyle. Using 1960 data across states in the United States, Auster et al. (1969) find that increases in per capita health care expenses had no discernible effect on mortality rates. Hadley (1982) uses US data for 1970 and decomposes mortality by age, sex, and race. He estimates the elasticity of mortality rates with respect to health care expenditures at below 0.2. Hadley (1988) updates his earlier study to 1980 and obtains greater estimates of the mortality elasticity than previously, at above 0.5 for some population subgroups. The use of disaggregated data in Hadley's studies may partly explain his estimation of increased mortality elasticity.

Besides mortality, the impact of health care on other aspects of health has also been examined. Newhouse and Friedlander (1980), Brook et al. (1983), and Valdez et al. (1985) measure health outcomes by morbidity rates, physiological indicators, or self-reported health status scores. Their results show that health care is statistically not a significant explanatory variable for the health status measures used.

The above evidence showing a secondary role of health care should not be taken to suggest that health care is an unimportant determinant of health. Some of the empirical studies cited may not have considered any interaction effects between health care and other determinants of health.[1] As emphasised by DaVanzo and Gertler (1990), effective health services depend on behavioural responses reflecting the impact of health policy on people's decisions to take up these services. The performance of health services hinges on both the behavioural influences and the biomedical effects of health care interventions. For example, clinical trials may suggest a vaccine can reduce the incidence of measles. But the effectiveness of a vaccination program depends on parents taking their children to use the service. Numerous causal factors affect parents' behaviour to seek health care, including their perceived costs and benefits of the service, education level, and religious attitude.

Along with economic development, formal health care has become an increasingly important contributor to health. The health sector expands with new functions, and takes over care activities that used to be performed by households, such as nursing home care.

[1] By including health status as an endogeneous variable, the econometric studies may have ignored the feedback effects of health on health inputs. Such misspecification of the production function can lead to biased estimates for the link between health status and health care. For example, healthier people have better access to education and health care, which in turn contributes to better health.

B Link between planning and clinical decisions

This appendix illustrates how health care policy can influence clinical practices through the effects on technology use and incentives to service providers.

Based on the findings of a European Community project on economic appraisal of health technology, Drummond (1987) raises numerous examples of expensive diagnostic devices, and frequently used 'small ticket' technologies alike, that have not been properly evaluated for their worthiness, hence resulting in increased health care costs. In the United Kingdom, for instance, computerised tomography (CT) scanning is performed in clinical applications where the detection rate is low relative to the costs. On the contrary, by requiring some form of economic evaluation in the funding process, heart transplant operations have been shown to represent relatively good value for money compared to other open-heart treatments.

To demonstrate the significance of incentives in affecting clinical performance, Drummond (1987) draws on the American experience during the 1980s in moving towards prospective reimbursement and managed competition for many health services. In essence, the reforms were designed to confront service providers with incentives to be cost conscious, hence making scope for cost saving through reductions in hospitalisation and increased use of broad medical management approaches such as preventive care, lifestyle and dietary measures.

Weisbrod (1991) provides an in-depth analysis of the dynamic interplay between the incentive structure of the US health care insurance system and the development and adoption of health technology in the country. The proliferation of new technologies creates new demand for health care and leads to changes in the form and extent of insurance coverage. Conversely, technological changes are influenced by incentives for the research and development sector to develop particular kinds of treatment procedures and technologies. Incentives and constraints associated with the third-party payment systems prevalent until the early 1980s are found to be responsible for the rapid, indiscriminate adoption of new health technology. Service providers faced incentives to create demand for diagnostic testing, surgery and other clinical procedures that deliver little material health benefits but perhaps some psychological effects. Further, redundant procedures and 'defensive medicine' were given in order to reduce the perceived risks of malpractice litigation.

Teplensky et al. (1995) apply regression techniques to identify the dominant motivations for over 500 US hospitals to employ magnetic resonance imaging equipment — a capital-intensive medical technology. Results confirm that clinical requirements were only one of the motives while the hospitals' investment behaviour was strongly linked to profit considerations and strategic positioning as a technological leader in a competitive market environment.

Finkler and Wirtschafter (1993) empirically investigate the causes of differences in the Cesarean-section practice among a number of hospitals offering obstetric services under the same health care insurance program in Southern California. The researchers estimate a regression model relating perinatal mortality rates to direct hospital costs and casemix adjusters, which control for clinical differences in the mothers and babies cared for. The risk adjusted cost-effectiveness comparisons reveal that high cesarean-section rates were unrelated to either costs or perinatal outcomes but likely caused by the more generous reimbursement scheme offered by insurers for cesarean sections than for vaginal deliveries, which is a less costly method of delivery.

In an Australian context, PC (1999) recognises that the domestic regulatory environment, health policy and commercial relationships between hospitals, doctors and health funds generally affect hospital efficiency.

At the global level, WHO (2000) notices that the payment systems in many countries contain distorted incentives that lead to mismanagement of health care resources. Lack of the skills to assess technology and control quality is another reason for inefficiency in health services.

C Different assessments of efficiency in the health sectors of selected OECD countries

In this appendix, WHO's (2000) efficiency estimates for selected OECD countries are compared with those derived by Färe et al. (1997).

Färe et al. measure relative efficiency as a distance from the estimated production frontier. Efficiency scores are decomposed into two components, one measuring technological changes (that is, shifts of the frontier over time) and another measuring efficiency gaps relative to the best practice frontier for a given year. However, only estimates of overall efficiency for 1989 — the latest year in the study period — are examined here.[1]

Färe et al. use two different models to generate efficiency estimates. The 'outcome' model represents health outcomes by two variables, namely the life expectancy of women at age 40 and the infant mortality rate. The 'output' model includes the number of hospital inpatient days and inpatient discharges as output variables. Both models have the same set of input variables: doctors and inpatient hospital beds. Due to data gaps in some countries, the 'outcome' model is applied to only a subset of the sample countries.

The WHO's efficiency estimates are derived from a 'outcome' model (see chapter 2).

Table C.1 displays the efficiency scores obtained by the two studies for different time periods. For ease of comparison, Färe et al.'s estimates are rebased assigning a score of one for Denmark, which has the highest estimated level of efficiency.

Results vary considerably between the two 'outcome' models. Färe et al. identify Denmark and the United States as relatively efficient. This is in sharp contrast to the WHO's assessment and contradicts the anecdotal evidence of inefficiency in the US health sector.

[1] To be specific, Färe et al. use a linear programming model to derive the Malmquist efficiency index. Such technique bears a close relationship to data envelopment analysis (DEA). See chapter 4 for an introduction to these techniques for estimating the production frontier.

Table C.1 **Comparison of efficiency estimates for the health sectors in selected OECD countries**

	WHO (2000)	*Färe et al. (1997)*	
	Outcome model 1993–97	*Outcome model 1989*	*Output model 1989*
Italy	0.976	0.763	0.627
France	0.974	0.583	0.814
Japan	0.945	0.831	0.480
Austria	0.914	0.568	0.839
Netherlands	0.893	na	0.706
Sweden	0.890	0.822	0.629
Belgium	0.878	na	0.609
Luxembourg	0.864	na	0.633
Canada	0.849	na	0.757
Germany	0.836	0.395	0.732
Finland	0.829	0.779	0.607
Denmark	0.785	1	1
United States	0.774	0.978	0.909
New Zealand	0.766	0.628	0.462

na: Not available.

Source: WHO (2000); Färe et al. (1997).

Färe et al.'s model deviates from the conceptual framework of health production in two ways. First, the outcome variable does not measure changes in health status. Second, determinants of health other than health care are not accounted for. Such deficiencies in model specification undermine the usefulness of their results. However, the assessments by the two studies may differ due to the different sample coverage, which affects the estimated production frontier. Also, the level of efficiency for a particular country may have changed beyond the timeframe analysed.

Färe et al.'s 'outcome' and 'output' models are two dramatically different assessments of relative efficiency across countries. Their 'outcome' model assesses France, Austria and Germany to be moderately efficient. But these countries obtain rather low efficiency scores from their 'output' model. Conversely New Zealand fares much better under their 'outcome' model than under their 'output' model. These comparisons confirm that output measures cannot be used as close proxy for outcome measures.

D Selected studies of hospital efficiency

Table D.1 **Efficiency measures and model variables included in ratio and frontier analyses**

Study / sample coverage	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Ratio analysis				
1. Duckett and Jackson (2000): public and private hospitals in Australia in 1996–97	• Non-capital cost per casemix adjusted separation	• Casemix adjusted number of separations	• Total hospital cost adjusted for discrepancies in the accounting rules adopted by hospitals	
2. SCRCSSP (2000): public hospitals in Australia in years 1995–96 to 1997–98	• Recurrent cost per casemix adjusted separation	• Casemix adjusted number of separations	• Total recurrent hospital cost	
	• Total cost per casemix adjusted separation	• Casemix adjusted number of separations	• Total cost, including recurrent expenditures, depreciation and user cost of buildings and equipment	
	• Average length of stay by DRG	• Number of separations by DRG	• Total number of hospital days	
	• Average cost per non-admitted occasion of service	• Number of non-admitted patient cases	• Expenditure allocated to non-admitted patients	
3. PC (1999): private hospitals in Australia in 1996–97	• Average cost per casemix adjusted separation	• Casemix adjusted number of separations	• Total recurrent hospital cost	
	• Revenue per employee	• Total hospital revenue	• Number of hospital staff	
	• Average length of stay by DRG	• Number of separations by DRG	• Total number of hospital days	
4. Long et al. (1990): 646 US non-federal short-term hospitals in years 1980 to 1985	• Various casemix adjusted partial measures of input intensity by discharge status	• Casemix adjusted numbers of separations by discharge status (DRGs augmented by the Body System Count Methodology to measure patient illness severity)	• Lengths of stay, for pre- and post-operative periods • Number of laboratory tests • Number of diagnostic tests • Number of chest X-rays • Number of drugs	• 5 discharge destinations: home, home health agency, skilled nursing facility, short-term hospital, and death

Table D.1 **(continued)**

Study / sample	*Efficiency measures* [b]	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
5. Puig-Junoy (1998): 993 critical care patients from 16 intensive care units (ICU) in Catalonia, Spain in 1991–92	• 2-stage DEA: first stage to compute overall efficiency scores by risk patient group and ICU; second stage to relate efficiency scores to environmental variables by means of a regression model	• Number of days for critically ill patients to survive in intensive care units (ICU) • Surviving discharge status	• ICU and hospital days weighted by factors reflecting varying resource consumption over the period of hospitalisation • Available nurse days per patient • Available physician days per patient • Available ICU technologies	• Mortality risk at admission • Mortality probability variation in the first 24 hours in ICU • Hospital ownership • Intensity of market competition • Hospital size, in terms of beds • Degree of specialisation in ICU treatment, measured by the share of patients treated in ICU in different risk groups • Frequency in using ICU clinical guidelines • Existence of evaluation program for medical staff • Daily visits per patient • Teaching status • Patient age • Urgent admission status
6. Hughes and Yaisawarng (2000): 70 acute care hospitals in NSW in 1995–96	• 4-stage DEA: first stage to compute DEA efficiency scores; second stage to regress input slacks on environmental variables; third stage to adjust data for environmental effects using regression estimates; fourth stage to recompute DEA efficiency scores using adjusted data	• Numbers of separations for acute and subacute inpatients weighted by DRG factors • Number of non-admitted patient occasions of service weighted by DRG factors • Number of junior medical officers in training	• Salaried medical officers • Visiting medical officers • Nurses • Other staff • Cost of goods and services • Depreciation costs	• Admitted patients' socioeconomic background, indicated by their area of residence • Hospital size and caseload complexity, measured by the number of diagnoses • Hospital type and location (metropolitan or district, principal or referral)

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
7. Hogan and Wroblewski (1993): 300 selected US non-federal short-term hospitals in 1983 and 1984	• 2-stage, weight-restricted DEA: first stage to determine optimal input/output weights by minimising average cross-efficiency score for all sample hospitals; second stage to compute efficiency score for each hospital using the predetermined input/output weights; post hoc analysis to relate efficiency scores to environmental variables	• Pediatrics discharges by comorbidity group • Adult medical discharges by comorbidity group • Adult surgical discharges by comorbidity group • Obstetrics discharges by age, method of delivery, and comorbidity group • Psychiatric / substance abuse discharges by age group • Newborn discharges by birth condition • Subacute / long-term care patient days • Trainees • Emergency room visits • Outpatient visits • Outpatient surgical procedures	• FTEs of medical professionals • FTEs of registered nurses • FTEs of licensed practising nurses • FTEs of nursing aides • FTEs of administrators • FTEs of ancillary staff • FTEs of other staff • Energy cost • Employee benefits • Depreciation and interest cost • Supplies • Pediatrics beds • Acute beds • ICU beds • Obstetrics beds • Subacute / long term scare beds • Birthing rooms • Bassinets • Surgical operating rooms	• Medical training status • Nursing training status • Availability of residency program • Casemix index • Availability of cancer program • Number of services offered • Location • Patient background, indicated by the share of public patients • Church affiliation • Occupancy rate • Ownership • Availability of ambulatory surgery program • Average length of stay • Hospital size, measured by number of beds • Regulatory regime, indicated by the payment system • Participation in multi-hospital system • Use of contract management
8. Ferrier and Valdmanis (1996): 360 rural US hospitals in 1989	• 2-stage DEA (estimating cost and technical efficiency)	• Acute inpatient days • Subacute days • Intensive days • Surgeries performed • Discharges • Outpatients	• FTE personnel excluding doctors • Beds • Unit FTE labour cost (input price variable) • Net asset value per bed (input price variable)	• Total patient days as a proxy for hospital size • Occupancy rate • Share of outpatients in total patients • Share of intensive care days in total patient days • Ownership type • Risk adjusted rate of unexpected hospital death

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
9. Holvad and Hougaard (1993): 80 Danish hospitals in 1990	• 2-stage DEA, FDH [b]	• Discharges • Outpatient visits • Patient days	• Doctors • Nurses • Other health care personnel • Administrative personnel • Beds • Total current net expenditure	• Share of outpatients in total patients • Degree of centralisation, indicated by the ratio of beds in a hospital to total beds available in the county of that hospital • Average length of stay, as a proxy for casemix complexity • Capacity utilisation, indicated by the average time a bed is empty • Share of emergency cases in total patient cases • Share of outpatient cases in total non-emergency cases • Labour intensity, measured by the ratio of employees to beds
10. Grosskopf and Valdmanis (1993): 82 public and private not-for-profit hospitals in California, US in 1982	• DEA	• Casemix adjusted and unadjusted acute care inpatient days • Casemix adjusted and unadjusted intensive care inpatient days • Casemix adjusted and unadjusted numbers of surgeries • Ambulatory and emergency room visits • Casemix index	• Physicians and house staff • FTEs of non-physician labour • Net plant assets	
11. Mobley and Magnussen (1998): 50 Norwegian public hospitals and 178 short-term hospitals in California, US in 1991	• DEA	• Numbers of inpatient days by age group • Outpatient visits • Casemix index for patients aged at or above 65	• FTEs of physicians and residents • FTEs of other staff • Beds	

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
12. Valdmanis (1992): 41 public and private not-for-profit hospitals operating in urban areas of Michigan, US in 1981	• DEA	• Pediatric inpatient days • Non-pediatric inpatient days by age group • Acute inpatient days • Intensive care inpatient days • Surgeries • Emergency room visits • Ambulatory visits	• Housestaff • Physicians • FTEs of nurses • FTEs of other labour • Total admissions • Beds • Net plant assets	
13. SCRCSSP (1997): 109 public hospitals in Victoria in 1994–95	• DEA	• Numbers of weighted inlier equivalent separations (WIES) by case severity • Unplanned readmission rate within 28 days of discharge	• FTEs of medical staff • FTEs of non-medical staff • Non-salary costs • Non-medical salaries • Medical salaries	
14. Sherman (1984): 7 teaching hospitals in Massachusetts, US (year unknown)	• DEA	• Numbers of patient days for patients at or about, and below 65 years • Nurses trained • Numbers of interns and residents	• FTEs of medical and surgical staff • Purchased supplies and services • Bed days available	
15. Banker et al. (1986): 114 hospitals in North Carolina, US in 1978	• DEA	• Numbers of inpatient days by age group	• Nursing services • Ancillary services (including x-ray, anethesiology, laboratory labour) • Administrative and general services • Depreciation and interest costs	
16. Hollingsworth and Parkin (1995): 75 Scottish acute care hospitals in 1992–93	• DEA	• Numbers of hospital days for medical and surgical inpatients • Total accident and emergency attendances • Outpatient attendances • Obstetrics and gynaecology inpatient days • Other specialty inpatient days	• Staffed beds • Nurses • Total of professional, technical, administrative and clerical staff • Total of non-nursing medical and dental staff • Cost of drug supply • Capital charge	

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
17. Morey et al. (1992): 300 US non-federal short-term hospitals in 1983	• DEA	• Inpatient discharges • Newborns delivered, by birth condition • Rehabilitation patient days • Subacute patient days • Emergency room visits • Clinic visits • Ambulatory surgeries performed • Medical education in dollars • Rate of unexpected hospital mortalities • Risk adjusted rate of re-admission • Risk adjusted rate of post-surgical complications • Risk adjusted mortality rate within specified periods following admission	• Beds • Cost-of-living adjusted salaries and benefits for nursing personnel • Cost-of-living adjusted salaries and benefits for other personnel excluding physicians • Fees for purchased services • Interest expense • Cost-of-living adjusted salaries and benefits for interns and trainees • Other expenses, including expenditures on physicians	• Casemix index
18. Magnussen (1996): 46 Norwegian acute care non-teaching hospitals in years 1989 to 1991	• DEA	• Numbers of days for medical and surgical patients • Total patient days by case complexity group • Numbers of medical and surgical patients • Long-term care days • Outpatient visits	• FTEs of physicians and nurses • FTEs of other staff • Beds	
19. Burgess and Wilson (1993): 137 US veterans affairs hospitals in years 1985–1987	• DEA	• Inpatient days • Inpatient discharges • Outpatient visits • Ambulatory surgical procedures • Inpatient surgical procedures	• Acute care beds • Long-term beds • FTEs of clinical labour • FTEs of non-clinical labour • Physician hours	

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
20. Färe et al. (1989): 39 US hospitals in 1982	• DEA (incorporating measures of capacity)	• Acute care inpatient days • Intensive care inpatient days • Surgeries • Ambulatory and emergency visits	• Number of physicians • FTEs of non-physician staff • Beds • Number of admissions	• Location (urban/rural)
21. Morey et al. (1995): 314 US non-federal short-term hospitals in 1988	• DEA (estimating cost and technical efficiency)	• Casemix adjusted inpatient discharges • Newborns delivered • Subacute/rehabilitation patient days • Emergency visits • Clinic visits • Ambulatory surgeries • Dollar value of medical education	• Cost-of-living adjusted salaries and benefits for personnel • Fees paid for legal, auditing, consultancy and other contracted activities • Medical education expenditures • Expenditures on supplies and purchased services • Beds	• Risk adjusted rate of unexpected hospital death
22. Linna (1998): 43 Finnish public acute care hospitals in years 1988 to 1994	• DEA, Malmquist index [c]	• Number of emergency visits • Total scheduled and follow-up visits • DRG-weighted number of admissions • Bed days • Number of resident doctors receiving one year of training • Number of on-the-job training weeks of nurses • Number of impact-weighted scientific publications	• Net operating costs adjusted for price variation in wage rate and government health care expenditure deflator	
	• SCF	Same as above	• Net operating costs (including minor capital investments) • Number of beds (fixed input in the short-run cost function)	• Readmission rate within 60 days of discharge • Time variable, as a proxy for frontier shift • Teaching status

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
23. Webster et al. (1998): 301 Australian private non-psychiatric hospitals in years 1991–92 to 1994–95	• DEA, Malmquist index	• Acute care inpatient days, for psychiatric care, rehabilitation and medical care • Surgery inpatient days, for advanced surgery, standard surgery, minor surgery, and obstetrics • Non-inpatient occasions of service • Nursing-home type inpatient days • Surgical procedures • Acute care inpatient separations • Accident / emergency • Cost-weighted sum of inpatient bed days and non-inpatient occasions of service • Total inpatient revenue	• FTEs of salaried medical officers • Contract value of visiting medical officers • FTEs of nursing staff • FTEs of other staff • Beds • Material costs • Admissions	
	• SPF (1output per model)	• Total revenue • Cost-weighted sum of inpatient bed days and non-inpatient occasions of service	• FTEs of total staff • Total labour cost • Beds • Capital stock estimated using the Perpetual Inventory Method • Material costs • Materials and hired services, including contract medical officers	
	• SCF	• Total revenue • Cost-weighted sum of inpatient bed days and non-inpatient occasions of service	• Total non-capital expenditure • Unit labour cost (input price variable) • Deflator for materials and hired services (input price variable)	
24. Vitaliano and Toren (1996): 219 general care hospitals in New York, US in 1991	• SCF	• Patient days • Emergency room visits • Outpatient clinic visits	• Total cost • Wages of registered nurses (input price variable) • Wages of radiologists (input price variable)	• Casemix index • Technology index • Occupancy rate • Teaching status • Ownership

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
25. Chirikos and Sear (2000): 186 short-term acute care hospitals in Florida, US in years 1982 to 1993	• DEA	• Casemix adjusted admissions • Post-admission inpatient days by payer group • Tests and procedures for outpatients • Ambulatory services	• Wage and salary payments to medical staff • Wages and salary payments to administrative staff • Other expenses for patient care • Depreciation charges for plant and building • Depreciation charges for equipment • Other expenses including interest expenses on long-term loans	
	• SCF	Same as above	• Total cost	
26. Färe et al. (1993): 17 Swedish public non-teaching hospitals in years 1970 to 1985	• Malmquist index	• Inpatient discharges • Long-term inpatient bed day • Ambulatory care, measured by doctor visits	• Real labour input, including doctors, other medical staff, service labour and administrative labour • Real non-labour, non-capital input	
27. Burgess and Wilson (1995): 137 US veterans affairs (VA) hospitals and 2283 non-VA hospitals in years 1985 to 1988	• Malmquist index	• Acute care inpatient days • Casemix adjusted acute care inpatient discharges • Long-term care inpatient days • Outpatient visits • Ambulatory surgical procedures • Inpatient surgical procedures	• Acute-care beds weighted by a scope-of-services index • Long-term beds • FTEs of registered nurses • FTEs of licensed practical nurses FTEs of other clinical labour • FTEs of non-clinical labour • FTEs of long-term care labour	
28. Wagstaff (1989): 49 Spanish public hospitals in years 1977 to 1981	• SCF	• Cases per bed • Internal medicine cases • General surgery cases • Gynaecology cases • Paediatrics cases • Intensive care cases • Other cases	• Total operating cost (excluding capital expenditure) • Beds (proxy for capital in short-term cost function)	• Teaching status

Table D.1 **(continued)**

Study	*Efficiency measures*	*Output / outcome variables*	*Input / cost variables*	*Environmental / control variables*
Frontier analysis				
29. Zuckerman et al. (1994): 1600 US short-term general hospitals in 1986 and 1987	• SCF	• Inpatient admissions by payer group • Post-admission inpatient days by payer group • Outpatient visits	• Total cost • Average annual salary rate (input price variable) • Depreciation and interest costs per bed (input price variable)	• Share of beds in ICUs • Share of outpatient visits not involving surgery • Share of long-term admissions • Ratio of births to admissions • Ratio of residents to beds • Casemix index • Inpatient surgeries per admission • High technology index • Accreditation indicator • Patient age and sex • Share of admissions on weekends • Shares of admissions for psychiatric, rehabilitation, and alcohol related problems • Share of interstate admissions • Shares of discharges to other care facilities • Share of admissions transferred from another hospital or a long-term care facility • Severity of illness within a DRG • Stage of illness within a DRG • Rate of unexpected mortality within 30 days of admission • Unexpected rate of hospital misadventures

a DEA: data envelopment analysis; FDH: Full disposal hull; SPF: stochastic production frontier; SCF: stochastic cost frontier. **b** FDH is a special case of DEA, where only efficient observations are used as performance targets instead of linear combinations of them. **c** Using panel data, Malmquist indexes are derived as an extension of DEA to decompose estimated efficiency changes into shifts of the frontier over time and changes in resource waste relative to the frontier at a particular point of time.

Sources: As quoted.

References

ABS (Australian Bureau of Statistics) 1997, *National Health Survey: Summary of Results*, cat. no. 4364.0, AGPS, Canberra.

—— 2000, *Private Hospitals 1998–99*, cat. no. 4390.0, AusInfo, Canberra.

AIHW (Australian Institute of Health and Welfare) 2000a, *Australia's Health 2000*, AIHW, Canberra.

—— 2000b, *Australian Hospital Statistics, 1998–99*, cat. no. HSE 11 (Health Services Series no. 15), AIHW Canberra.

—— 2000c, *National Health Data Dictionary, Version 9*, National Health Data Committee, AIHW cat. no. HWI-24, AIHW, Canberra.

—— 2001, *Response by Australia and New Zealand to the World Health Report 2000*, AIHW, Canberra, http://www.aihw.gov.au/conferences/whr2000/index.html (accessed 17 April 2001).

Auster, R.D., Leveson, I. and Sarachek, D. 1969, 'The production of health, an exploratory study', *Journal of Human Resources*, vol. 4, no. 4, pp. 411–436.

Baily, M.N. and Garber, A.M. 1997, 'Health care productivity', *Brookings Papers: Microeconomics*, pp.143–202.

Banker, R.D. 1993, 'Maximum likelihood, consistency and data envelopment analysis', *Management Science*, vol. 39, no. 10, pp. 1265–1273.

——, Conrad, R.F. and Strauss, R.P. 1986, 'A comparative application of data envelopment analysis and translog methods: an illustrative study of hospital production', *Management Science*, vol. 32, no. 1, pp. 30–44.

Barrow, M. and Wagstaff, A. 1989, 'Efficiency measurement in the public sector: an appraisal', *Fiscal Studies*, vol. 10, no. 1, pp. 72–97.

Brazier, J., Deverill, M., Green, C., Harper, R. and Booth, A. 1999, 'A review of the use of health status measures in economic evaluation', *Health Technology Assessment*, vol. 3, no.9.

Brook, R.H., et al. 1983, 'Does free care improve adults' health? Results from a randomized controlled trial', *The New England Journal of Medicine*, vol. 309, pp 1426–1434.

Burgess, J.F. and Wilson, P.W. 1993, 'Technical efficiency in veterans administration hospitals', in Fried, A., Lovell C.A.K. and Schmidt S.S. (eds), *The Measurement of Productive Efficiency: Technique and Applications*, Oxford University Press, New York.

—— 1995, 'Decomposing hospital productivity changes, 1985–1988: a nonparametric Malmquist approach', *Journal of Productivity Analysis*, vol. 6, no. 4, pp. 343–363.

Butler, J. 1995, *Hospital Costs Analysis*, Kluwer Academic Publications, Dordrecht.

Charnes, A., Cooper, W.W. and Rhodes, E. 1978, 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, vol. 2, pp. 429–444.

——, ——, Lewin, A.Y. and Seiford, L.W. (eds) 1994, *Data Envelopment Analysis: Theory, Methodology, and Application*, Kluwer Academic.

Chassin, M.R., Hannan, E.L. and DeBuono, B.A. 1996, 'Benefits and hazards of reporting medical outcomes publicly', *New England Journal of Medicine*, vol. 334, no. 6, pp. 394–398.

Chirikos, T.N. and Sear, A.M. 2000, 'Measuring hospital efficiency: a comparison of two approaches', *Health Services Research*, vol. 34, no. 6, pp. 1389–1408.

Coelli, T.J. 1995, 'Recent developments in frontier modelling and efficiency measurement', *Australian Journal of Agricultural Economics*, vol. 39, no. 3, pp. 219–245.

Cook, W.D., Moez, H. and Hans, J.H.T. 2000, 'Multicomponent efficiency measurement and shared inputs in data envelopment analysis: an application to sales and service performance in bank branches', *Journal of Productivity Analysis*, vol. 14, no. 3, pp. 209–224.

DaVanzo, J. and Gertler, P. 1990, *Household Production of Health: A Microeconomic Perspective on Health Transitions*, Rand Note no. N-3014-RC, Rand Corporation.

Drummond, M. 1987, 'Economic evaluation and the rational diffusion and use of health technology', *Health Policy*, vol. 7, pp. 309–324.

——, Hailey, D. and Smith, C.S. 1991, *Maximising the Impact of Health Technology Assessment: The Australian Case*, National Centre for Health Program Evaluation Working Paper no. 12, Melbourne.

——, O'Brien, B., Stoddart, G.L. and Torrance, G.W. 1997, *Methods for the Economic Evaluation of Health Care Programmes*, 2nd edn, Oxford University Press, New York.

DHAC (Commonwealth Department of Health and Aged Care) 2000a, *Annual Report 1999–2000*, DHAC, Canberra.

—— 2000b, *National Hospital Cost Data Collection 1998–99, Final Cost Report (Round 3)*, DHAC, Canberra.

Duckett, S.J. and Jackson, T.J. 2000, 'The new health insurance rebate: an inefficient way of assisting public hospitals', *Medical Journal of Australia*, no. 172, pp. 439–442.

E-QUAL Donovan Research 2000, *National Satisfaction Survey of Clients of Disability Services*, A report prepared for the Steering Committee for the Review of Commonwealth/State Service Provision and National Disability Administrators, AusInfo, Canberra.

Eagar, K. and Hindle, D. 1994, *A Short Introduction to Casemix*, National Casemix Education Series no. 1, Department of Human Services and Health, Canberra.

Eckermann, S. 1994, 'Hospital productivity: a contradiction in terms?' in Harris, A. (ed), *Economics and Health 1994: Proceedings of the Sixteenth Australian Conference of Health Economists*, Australian Studies in Health Service Administration series, no. 78, School of Health Administration, University of New South Wales, pp. 167–181.

Evans, D.B., Tandon, A., Murray, C.J.L. and Lauer, J.A. 2000, *The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries*, Global Programme on Evidence for Health Policy Discussion Paper no. 29, World Health Organisation, Geneva.

Färe, R., Grosskopf, S., Lindgren, B. and Poullier J. 1997, 'Productivity growth in health care delivery', *Medical Care*, vol. 35, no. 4, pp. 354–366.

——, ——, —— and Roos, P. 1993, 'Productivity developments in Swedish hospitals: a Malmquist output index approach,' in Charnes, A., Cooper, W., Lewin, A.Y. and Seiford, L.M. (eds), *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic.

——, —— and Valdmanis, V. 1989, 'Capacity, competition and efficiency in hospitals: a nonparametric approach', *Journal of Productivity Analysis*, vol. 1, pp. 121–138.

Ferrier, G.D. and Valdmanis, V. 1996, 'Rural hospital performance and its correlates', *Journal of Productivity Analysis*, vol. 7, no. 1, pp. 63–80.

Finkler, M.D. and Wirtschafter, D.D. 1993, 'Why pay extra for Caesarean-section deliveries?' *Inquiry*, vol. 30, no. 2, pp. 208–215.

Finkler, S.A., Brooten, D. and Brown, L. 1988, 'Utilization of inpatient services under shortened lengths of stay: a neonatal care example', *Inquiry*, vol. 25, no. 2, pp. 271–280.

Folland, S., Goodman, A.C., and Stano, M. 1997, *The Economics of Health and Health Care*, 2nd edn, Prentice Hall, New Jersey

Førsund, F.R. 1998, 'The rise and fall of slacks: comments on quasi-Malmquist productivity indices', *Journal of Productivity Analysis*, vol. 10, no. 1, pp. 21–34.

Fried, H.O., Schmidt, S.S. and Yaisawarng, S. 1999, 'Incorporating the operating environment into a nonparametric measure of technical efficiency,' *Journal of Productivity Analysis*, vol. 12, no. 3, pp. 249–267.

——, Lovell, C.A.K. and Schmidt, S.S. (eds) 1993, *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press, New York.

Fuchs, V.R. 1974, *Who Shall Live?* Basic Books Inc, New York.

Garber, A.M. and Phelps, C.E. 1997, 'Economic foundations of cost–effectiveness analysis', *Journal of Health Economics*, vol. 16, no. 1, pp. 1–31.

Gold, M.R., Siegel, J.E., Russell, L.B. and Weinstein, M.C. 1996, *Cost–effectiveness in Health and Medicine*, Oxford University Press, New York.

Grosskopf, S. and Valdmanis, V. 1993, 'Measuring hospital performance: a nonparametric approach', *Medical Care*, vol. 31, no. 6, pp. 525–532.

Grossman, M. 1972a, *The Demand for Health: A Theoretical and Empirical Investigation*, NBER Occasional Paper no. 119, Columbia University Press, New York.

—— 1972b, 'On the concepts of health capital and the demand for health', *Journal of Political Economy*, vol. 80, no 2, pp. 223–255.

Hadley, J. 1982, *More Medical Care, Better Health?* Urban Institute, Washington D.C.

—— 1988, 'Medicare spending and mortality rates of the elderly', *Inquiry*, vol. 25, no. 4, pp. 485–493.

Hardwick, J. 1986, 'Hospital case mix standardisation: a comparison of the resource need index and information theory measures', in Butler, J. R. G. and Doessel, D.P. (eds), *Economics and Health 1985: Proceedings of the Seventh Australian Conference of Health Economists*, Australian Studies in Health Service Administration series, no. 56, School of Health Administration, University of New South Wales, pp. 36–63.

Hogan, A.J. and Wroblewski, R. 1993, 'The determinants of hospital efficiency: a data envelopment analysis,' in Rhodes, E.L. (ed), *Applications of Management*

Science, Public Policy Applications of Management Science, JAI Press, vol. 7, pp. 115–144.

Hollingsworth, B. and Parkin, D. 1995, 'The efficiency of Scottish acute hospitals: an application of data envelopment analysis', *IMA Journal of Mathematics Applied in Medicine and Biology*, vol. 12, no. 3–4, pp. 161–173.

——, Dawson, P.J. and Maniadakis, N. 1999, 'Efficiency measurement of health care: a review of nonparametric methods and applications', *Health Care Management Science*, vol. 2, no. 3, pp. 161–172.

Holvad, T. and Hougaard, J.L. 1993, *Measuring Technical Input Efficiency for Similar Production Units: 80 Danish Hospitals*, European University Institute Working Paper ECO no. 93/96, San Domenico, Italy.

HSPF (Harvard School of Public Health) 2001, 'Executive summary', *The Global Burden of Disease and Injury Series*, Cambridge, http://www.hsph.harvard.edu/organizations/bdu/gbdmain.html (accessed 4 May 2001).

Hughes, A. and Yaisawarng, S. 2000, 'Operating environment adjusted efficiency of acute care public hospitals in New South Wales', in *Proceedings of the 2000 International DEA Symposium, Ipswich, 3–5 July, 2000*, University of Queensland.

Hull, R., Hirsh, J., Sackett, D.L. and Stoddart, G.L. 1981, 'Cost-effectiveness, clinical diagnosis, venography and noninvasive testing in patients with symptomatic deep-vein thrombosis', *New England Journal of Medicine*, vol. 304, pp. 1561–67.

Jacobs, P. and Baladi, J. 1996, 'Biases in cost measurement for economic evaluation studies in health care', *Health Economics*, vol. 5, no. 6, pp. 525–529.

Jomini, P. and Chan, C. 2000, 'Production technology, operating environment and quality issues in DEA: a case study of railway efficiency, in *Proceedings of the 2000 International DEA Symposium, Ipswich, 3–5 July, 2000*, University of Queensland.

Klarman, H.E. 1965, 'Syphilis control programs', in Dorfman, R. (ed) *Measuring Benefits of Government Investments*, Brookings Institution, Washington D.C.

Landon, B., Iezzoni, L., Ash, A.S., Shwartz, M., Daley, J., Hughes, J.S. and Mackiernan, Y.D. 1996, 'Judging hospitals by mortality rates: the case of CABG surgery', *Inquiry*, vol. 33, no. 2, pp.155–166.

Lauer, J.A., Tandon, A., Evans, D.B. and Murray, C.J.L. 2001, *Determinants of Health System Performance*, Global Program on Evidence for Health Policy Working Paper, World Health Organisation, Geneva.

Linna, M. 1998, 'Measuring hospital cost efficiency with panel data models,' *Health Economics*, vol. 7, no. 5, pp. 415–427.

Long, M.J., Chesney, J.D. and Fleming, S.T. 1990, 'A reassessment of hospital product and productivity changes over time', *Health Care Financing Review*, vol. 11, no. 4, pp. 69–77.

Lovell, C.A.K. 1993, 'Production frontiers and productive efficiency', in Fried, H., Lovell, C.A.K. and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency*, Oxford University Press, pp. 3–67.

——, Sarkar, A. and Sickles, R.C. 1987, *Output Aggregation and the Measurement of Productive Efficiency, Science*, Working Paper Series no. 87–6, Department of Economics, University of North Carolina.

Magnussen, J. 1996, 'Efficiency measurement and the operationalization of hospital production', *Health Services Research*, vol. 31, no. 1, pp. 21–37.

Marshall, P., Chan, C. and Trace, K. 2000, *Estimating the Costs of Container Transport — A Cost Model Approach*, Productivity Commission Research Memorandum no. MR113, Melbourne.

McKeown, T. 1976, *The Modern Rise of Population*, Academic Press, New York.

McKie, J., Richardson, J., Singer, P. and Kuhse, H. 1998, *The Allocation of Health Care Resources*, Dartmouth Publishing, Aldershot, England.

McKinlay, J.B. and McKinlay, S.M. 1977, 'The questionable contribution of medical measures to the decline of mortality in the United States in the twentieth century', *Milbank Memorial Fund Quarter/Health and Society*, vol. 55, pp. 405–428.

McNeil, B.J., Pedersen, S.H. and Gatsonis, C. 1992, 'Current issues in profiling quality of care,' *Inquiry*, vol. 29, no. 3, pp. 298–307.

Mehrez, A. and Gafni, A. 1989, 'Quality adjusted life years, utility theory, and healthy years equivalents', *Medical Decision Making*, vol. 9, no. 2, pp. 142–149.

Mobley, L.R. and Magnussen, J. 1998, 'An international comparison of hospital efficiency: does institutional environment matter?' *Applied Economics*, vol. 30, no. 8, pp. 1089–1100.

Morey, R.C., Fine, D.J., Loree, S.W., Retzlaff-Roberts, D.L., and Tsubakitani, S. 1992, 'The trade-off between hospital cost and quality of care: an exploratory empirical analysis', *Medical Care*, vol. 30, no. 8, pp. 677–698.

——, Ozcan, Y.A., Retzlaff-Roberts, D.L. and Fine, D.J. 1995, 'Estimating the hospitalwide cost differentials warranted for teaching hospitals, an alternative to regression approaches', *Medical Care*, vol. 33, no. 5, pp. 531–552.

Murray, C.J.L., and Lopez, A.D. (eds.) 1996, *The Global Burden of Disease: a Comprehensive Assessment of Mortality and Disability from Diseases, and Risk Factors in 1990 and projected to 2001*, Harvard University Press, Cambridge.

Mushkin, S.J. 1962, 'Health as an investment', *Journal of Political Economy*, vol. 70, no. 5, pp. 129–157.

Newhouse, J.P., and Friedlander, L.J. 1980, 'The relationship between medical resources and measures of health: some additional evidence', *Journal of Human Resources*, vol. 15, no.2, pp. 200–218.

NHPC (forthcoming), *National Health Performance Framework Report*, NSW Health Department, Sydney.

—— 2000, *Fourth National Report on Health Sector Performance Indicators — A Report to the Australian Health Ministers' Conferences*, NSW Health Department, Sydney,

Nord, E. 1992, 'An alternative to QALYs: the saved young life equivalent (SAVE)', *British Medical Journal*, vol. 305, pp. 875–877.

Palmer, G.R. 1986, 'Hospital output and the use of diagnosis related groups for purposes of economic and financial analysis', in Butler, J. R. G. and Doessel, D.P. (eds), *Economics and Health 1985: Proceedings of the Seventh Australian Conference of Health Economists*, Australian Studies in Health Service Administration series, no. 56, School of Health Administration, University of New South Wales, pp. 159–181.

—— 1991, 'The use of DRGs in the management and planning of hospital services', *Australian Economic Review*, no. 93, pp. 62–70.

—— and Wood, T. 1984, 'Diagnosis related groups: recent developments and their adaptation and application in Australia', *Australia Health Review*, vol. 7, no. 2, pp. 67–89.

PC (Productivity Commission) 1999, *Private Hospitals in Australia*, Commission Research Paper, AusInfo, Canberra.

Pliskin, J.S., Shepard, D.S. and Weinstein, M.C. 1980, 'Utility functions for life years and health status', *Operation Research*, vol. 28, pp. 206–224.

Puig-Junoy, J. 1998, 'Technical efficiency in the clinical management of critically ill patients', *Health Economics*, vol. 7, no. 3, pp. 263–277.

Robinson, J.C. 1986 'Philosophical origins of the economic evaluation of life', *Millbank Memorial Fund Quarterly*, vol. 64, no. 1, pp. 133–55.

Ross, J. 1995, 'The use of economic evaluation in health care: Australian decision makers' perceptions', *Health Economics*, vol. 7, no. 3, pp. 263–277.

Russell, L. 1992, 'Opportunity costs in modern medicine', *Health Affairs*, Summer, pp. 162–69.

SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision) 1997, *Data Envelopment Analysis, A Technique for Measuring the Efficiency of Government Service Delivery*, AGPS, Canberra.

—— 2001, *Report on Government Services 2000*, AusInfo, Canberra.

Sherman, H.D. 1984, 'Hospital efficiency measurement and evaluation: empirical test of a new technique', *Medical Care*, vol. 22, no.10, pp. 922–938.

Teplensky, J.D., Pauly, M.V., Kimberly, J.R., Hillman, A.L. and Schwartz, S. 1995, 'Hospital adoption of medical technology: an empirical test of alternative models', *Health Services Research*, vol. 30, no. 3, pp. 437–465.

Thrall, R.M. 2000, 'Measures in DEA with an application to the Malmquist index', *Journal of Productivity Analysis*, vol. 13, no. 2, pp. 125–137.

Torrance, G.W. 1987, 'Utility approach to measuring health related quality of life', *Journal of Chronic Diseases*, vol. 40, no. 6, pp. 563–600.

Valdez, R.B., Brook, R.H., Rogers, W.H., et al. 'Consequences of cost sharing for children's health,' *Pediatrics*, vol. 75, pp. 952–961.

Valdmanis, V. 1992, 'Sensitivity analysis for DEA models: an empirical example using public vs. NFP hospitals', *Journal of Public Economic*, vol. 48, no.2, pp. 185–205.

Vitaliano, D.F. and Torren, M. 1996, 'Hospital cost and efficiency in a regime of stringent regulation', *Eastern Economic Journal*, vol. 22, no. 2, pp. 161–175.

Wagstaff, A. 1989, 'Estimating efficiency in the hospital sector: a comparison of three statistical cost frontier models', *Applied Economics*, vol. 21, no. 5, pp. 659–672.

Webster, R., Kennedy, S. and Johnson, L. 1998, *Comparing Techniques for Measuring the Efficiency and Productivity of Australian Private Hospitals*, Australian Bureau of Statistics Working Paper cat. no. 1351.0, Canberra.

Weisbrod, B.A. 1971, 'Costs and benefits of medical research: a case study of poliomyelitis', *Journal of Political Economy*, vol. 79, no. 3, pp. 527–544.

—— 1991, 'The health care quadrilemma: an essay on technology change, insurance, quality of care, and cost containment', *Journal of Economic Literature*, vol. 29, no. 2, pp. 523–552.

—— 1992, 'Productivity and incentives in the medical care sector,' *Scandinavia Journal of Economics*, vol. 94, supplement, pp. 131–145.

World Health Organisation (WHO) 2000, *Health Systems: Improving Performance*, Geneva.

Zuckerman, S., Hadley, J. and Iezzoni, L. 1994, 'Measuring hospital efficiency with frontier cost functions', *Journal of Health Economics*, vol. 13, no. 3, pp. 255–280.