DOCTOR OF PHILOSOPHY

Measurement of hospital performance

Gerard O'Callaghan

2011

Aston University

MEASUREMENT OF HOSPITAL PERFORMANCE

GERARD O'CALLAGHAN

Doctor of Business Administration

ASTON UNIVERSITY

November 2011

Aston University

Measuring Hospital Performance

Student: **Gerard O'Callaghan**
Degree: **Doctor of Business Administration**
Year: **2011**

Summary

The rationale for carrying out this research was to address the clear lack of knowledge surrounding the measurement of public hospital performance in Ireland. The objectives of this research were to develop a comprehensive model for measuring hospital performance and using this model to measure the performance of public acute hospitals in Ireland in 2007.
Having assessed the advantages and disadvantages of various measurement models the Data Envelopment Analysis (DEA) model was chosen for this research. DEA was initiated by Charnes, Cooper and Rhodes in 1978 and further developed by Fare et al. (1983) and Banker et al. (1984). The method used to choose relevant inputs and outputs to be included in the model followed that adopted by Casu et al. (2005) which included the use of focus groups.
The main conclusions of the research are threefold. Firstly, it is clear that each stakeholder group has differing opinions on what constitutes good performance. It is therefore imperative that any performance measurement model would be designed within parameters that are clearly understood by any intended audience.
Secondly, there is a lack of publicly available qualitative information in Ireland that inhibits detailed analysis of hospital performance.
Thirdly, based on available qualitative and quantitative data the results indicated a high level of efficiency among the public acute hospitals in Ireland in their staffing and non pay costs, averaging 98.5%. As DEA scores are sensitive to the number of input and output variables as well as the size of the sample it should be borne in mind that a high level of efficiency could be as a result of using DEA with too many variables compared to the number of hospitals. No hospital was deemed to be scale efficient in any of the models even though the average scale efficiency for all of the hospitals was relatively high at 90.3%.
Arising from this research the main recommendations would be that information on medical outcomes, survival rates and patient satisfaction should be made publicly available in Ireland; that despite a high average efficiency level that many individual hospitals need to focus on improving their technical and scale efficiencies, and that performance measurement models should be developed that would include more qualitative data.

Key words: Data Envelopment Analysis, hospital efficiency, technical efficiency, scale efficiency, focus groups

Measuring Hospital Performance

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

The increasing cost of health care is a major concern for most economies. Health care costs have shown substantial increases in most western countries over the last fifty years. *"It is generally accepted that, under the pressure of cost increasing technological change and increases in demand due to epidemiological and demographic factors, this trend will continue"* (Maniadakis and Thanassoulis, 2000: 1578). This has resulted in an increasing emphasis being placed on the reform of health care systems with the objective of cost containment through increases in efficiency and productivity. *"Changes in the structure of the U.S. health care industry have forced decision-makers to look for ways to become more productive and cost efficient"* (Ferrier and Valdmanis, 2004: 1071).

In Ireland it was also clear that the efficient use of health resources had become a priority for policy makers. Mary Harney, Minister for Health and Children, stated:

> *The economic backdrop against which health and personal social services are delivered remains very challenging and we are facing severe resource pressures for some time to come. Targeting and the efficient use of resources that are available is thus of paramount importance.* (Department of Health and Children Annual Report 2009: 4)

Similarly in the UK policy makers were clearly focussed on increasing efficiency in the use of health resources. The Right Honourable Alan Johnson, Secretary of State for Health, stated:

"Alongside the increased investment, and in view of the current economic climate, it is right that the NHS, alongside other public services, is tasked with continuing to deliver increased efficiency in its use of resources." (Department of Health, Departmental Report 2009: 2)

The Wanless Report which was published in the UK on the 17th April 2002 set out projections of how much it would cost to deliver high quality services throughout the NHS for the next twenty years. A few hours later the Chancellor of the Exchequer Gordon Brown announced in his budget speech that funding for the NHS would increase by 7.4% annually in real terms over the next five years, but that this funding would be linked to further reforms that would make the NHS more responsive to the needs of patients. On the next day the Secretary of State for Health Alan Milburn published an outline of these reforms in England in the document "Delivering the NHS Plan." In this plan he set out specific productivity gains that he would like to see in response to the additional funding. These included:

- Major expansion in NHS activity.
- Structural changes with further power devolved to front line NHS, creation of foundation hospitals financial systems that will follow the patient.
- Financial penalties for delayed discharges.
- Establishment of Commission of Health Audit and Inspection.
- Further co-operation with the private sector.

While it is clear that policy makers do want to make efficiency gains the question that needs to be asked is why they do not make more use of efficiency studies.
"The main reason that efficiency analyses are little utilised by policy makers appears to stem from concern about their reliability." (Hollingsworth and Street, 2006: 1057)
If policy makers are to make more use of efficiency analyses they need to be confident that the analytical results are reliable. Policy makers require relevant, timely and reliable efficiency analyses. This is also true for decision making units such as hospitals. Hospitals seeking to improve their efficiency levels require comparative efficiency analyses. However, due to the limited disclosure of information these analyses may not always be accurate. Even when the efficiency analyses are accurate they may not provide the specific information required about the efficiency of the production processes in the hospital in order to take action. This research highlights the limited availability of relevant performance measurements in Ireland. To address these issues further development of analytical techniques is necessary. In particular there needs to be greater attention to model construction. The analysis also needs to be more specific in identifying the nature and form of any inefficiency. In order to achieve these objectives academic researchers need to work more closely with policy makers and decision making units.

Table 1.1 sets out the change in health expenditure as a share of Gross Domestic Product (GDP) in OECD countries between 1960 and 2009. This table clearly highlights the enormous increase in health expenditure as a share of GDP across the OECD countries. During this period the OECD average increased from 3.8% to 9.5%. A slightly higher increase occurred in Ireland where health expenditure as a share of GDP increased from 3.7% to 9.5%, which now places it at the OECD average, while the UK is above the average at 9.8%. The country with the highest share is the USA at 17.4%, followed by Netherlands at 12% and France at 11.8%. It is also worth noting that the USA has shown the greatest increase during this period, increasing its share from 5.1% to 17.4%.

Table 1.1

Health Expenditure as a share of GDP, from 1960 to 2009, in OECD countries

***Source: OECD Health Data 2011*.**

Figure 1.1 which sets out total health expenditure as a share of GDP between 1960 and 2009 for selected OECD countries again highlights the upward trajectory in health expenditure as a share of GDP in these countries.

Figure 1.1

Total health expenditure as a share of GDP, 1960-2009, selected OECD countries

Aston University

Illustration removed for copyright restrictions

Source: OECD Health Data 2011.

The performance of hospitals has also come under intense scrutiny. Increasing emphasis is being placed on measuring their relative performance to ensure that all resources are utilised in the best possible way. Given the vast amount of funds provided to hospitals there is a growing interest in ensuring that they operate as efficiently as possible. The current economic downturn has resulted in hospital efficiency becoming an even more crucial topic.

This research is aimed at developing an organisational performance measurement model within the acute hospital sector that will incorporate all key performance measures. In this introduction the main concepts that surround this area are firstly highlighted. What is meant by organisational performance and how one measures that performance is then addressed whilst also discussing the critical concepts of efficiency and effectiveness. The organisation models that exist and the framework within which performance is measured are next discussed and these are followed by an examination of the conceptualisation of productivity and the measurement topics that must be considered when measuring this. There are also a number of important design criteria for a productivity measurement system that are then considered. The performance measurement model that is used to analyse the data collected is then described and the introduction concludes with a discussion on the approach used in the collection of the data for this research.

1.2 What is meant by organisational performance?

Pritchard importantly distinguishes between performance and productivity. He states that *“performance is simply output”* (Pritchard, 1990: 467). He defines productivity *“as an index of output relative to goals (effectiveness) or output relative to inputs (efficiency)”* (Pritchard, 1990: 447). While outputs are part of productivity, productivity also includes inputs or a measure of outputs relative to objectives or goals. Likewise productivity is not profitability. *“Profitability includes measurement of the degree of cost recovery; productivity does not”* (Kendrick, 1984; Mali, 1978).

Concepts like production capacity or output capability are measures of the potential outputs of a system and are not measures of productivity as productivity deals with actual output. Based on these definitions the concept of productivity more clearly defines what I am trying to measure in this research. The concept of performance does not go far enough in defining my objectives as it will be necessary to measure actual output against a defined goal or outputs relative to inputs and performance measures do not necessarily do this.

The concept of productivity can be viewed from a different perspective by different disciplines. Pritchard (1990) discusses these perspectives in turn. An economist sees productivity as the efficiency of the transformation of inputs into outputs (Kendrick, 1977, 1984; Kopelman, 1986; Mahoney, 1988; Silver, 1984). Sometimes economists argue that a definition of efficiency is a snapshot, whereas productivity is efficiency measured over time. An accountant measures productivity based on an accounting perspective (e.g. Denison, 1984; Hurst, 1980) and attempts to describe and improve the financial performance of the organisation. The industrial engineer (e.g. Norman and Bahiri, 1972; Rosow, 1981) views productivity as the efficiency of throughput as measured by output to input ratios. The psychologist focuses primarily on the aspects of productivity that the individual can control, i.e. behaviour (Campbell and Campbell, 1988c; Guzzo, 1988; Guzzo, Jette and Katzell, 1985; Ilgen and Klein, 1988, Schneider, 1984). The manager takes the broadest but the least precisely identifiable perspective (Tuttle 1981, 1983). In this perspective (e.g. Preziosi, 1985; Shetty and Buehler, 1985), productivity includes all aspects of the organisation seen as important to effective organisational functioning. It includes efficiency and effectiveness but also includes quality of output; work disruptions; absenteeism, turnover and customer satisfaction. It is on this latter perspective on productivity that I will be basing my research.

1.3 Productivity

It is clear that the proper conceptualisation of productivity depends on the purpose of measurement. Once the purpose is identified, decisions about which approach to take or which unit of analysis to use become easier.

Pritchard presents the following definition of productivity:

> *"Productivity is how well a system uses its resources to achieve its goals"* (Pritchard, 1990: 455).

The one issue that needs to be borne in mind in relation to this definition is that we need to guard against a situation where an organisation deems itself to be highly productive only because it has set itself very low goals. There needs to be control on the level of the goals set.

There are a number of important points contained in Pritchard's definition. Firstly, that productivity is a systems concept that can apply to various entities. Secondly, that productivity is a description of how well the system does something and as such is an evaluative concept. However given that different systems with different functions exist within organisations one measurement system cannot evaluate all systems and functions. The developer or user of the productivity measurement system must ask what system and what functions are to be evaluated. Thirdly, that both efficiency and effectiveness are part of productivity. Both efficient use of inputs to produce outputs and producing outputs that meet organisational goals are included. Fourthly, that productivity accepts a goal-oriented model of organisations, with some revisions for natural systems and the multiple constituency models. The definition assumes that all systems in organisations have survival as their primary goal. The definition also agrees that it is inappropriate to assume that all organisational goals are totally the product of rational decision making and are the sole determinants of organisational actions. Pritchard states that *"the determination of objectives is a developmental, evolutionary and highly political process that is less than totally rational and objectives must be set for an unknowable future"* (Pritchard, 1990: 456).

The specific goals that the productivity measurement system will be based on depend on the objectives of those developing the system.

1.4 Why measure performance?

"Clearly a principal objective of performance measurement is to enhance various notions of efficiency" (Thanassoulis, 2001: 1).

The information derived from any measurement will be dependent on the aim of the assessment and the objectives of the organisation concerned.

In the 1860's Florence Nightingale highlighted the differences in mortality rates in London hospitals whilst in 1917 Ernest Codman complained that fellow surgeons failed to publish their results because of fear that the public might not be impressed. The interest in system performance was given impetus by the World Health Report 2000 produced by the WHO entitled "Health Systems: improving performance". While this has resulted in many countries over recent years starting to publish comparative information about health care performance and others planning to do so in the near future, it also gave rise to much discussion and criticism. Much of the criticism related to the rankings received by some countries in the report which appeared to contradict general perceived views of the countries' health systems. Blendon et al. (2001) compared the WHO rankings for seventeen industrialised countries with the perceptions of their citizens and found that their results showed that there were little relationships between the WHO rankings and the satisfaction of their citizens. Their findings suggested that both public and expert views should be considered in international rankings. Similarly Richardson et al. (2003) concluded that country rankings based upon the model were unreliable. However, they also concluded that despite the problems with the model that the study was a landmark in the evolution of system evaluation but one which required significant revision. On the other side of the argument Murray et al. (2001) defended the WHO model. They argued in response to Blendon et al.'s article that satisfaction with the way health care runs in a country is not conceptually comparable with overall health system performance or attainment and only partly comparable to responsiveness.

The Wanless Report (2002) in the UK had as a key assumption that productivity would improve over time and that it should at least match the productivity performance in the rest of the service sector of the economy.

The USA has the greatest experience of publishing comparative health performance data. The best known reporting system is the Health Plan Employer Data Information Set (HEDIS), which is produced by the National Committee for Quality Assurance

(NCQAD). Another high profile reporting system is the New York Cardiac Surgery Reporting System which publishes hospital and surgeon specific risk adjusted coronary artery bypass graft surgery (CABG) mortality.
The OECD has also published comparative health data over many years. Table 1.2 sets out OECD health data showing the change in average life expectancy between 1960 and 2009 in Ireland, the UK and the OECD average.

Table 1.2

Life Expectancy Data between 1960 and 2009 for Ireland, the UK and the OECD average

Aston University

Illustration removed for copyright restrictions

Source: OECD Health Data 2011

It is clear from this table that life expectancy has increased considerably in Ireland and in the UK during this period. Average OECD life expectancy has shown an even greater percentage increase over this period. It could be argued that this improvement in health status justifies the enormous increase in health expenditure across the OECD countries between 1960 and 2009.

However the publication of any information collected, particularly that relating to public sector organisations, may result in unintended consequences. Goddard and Smith (2002) for example identified nine enemies of virtuous performance measurement. These are:

- Tunnel vision, i.e. concentration on areas that are included in the performance scheme, to the exclusion of other important unmeasured areas.
- Measure fixation, i.e. pursuant of success as measured rather than as intended. For example the employment of a “hallo” nurse to address an accident and emergency waiting time criteria, with no impact on patient satisfaction or outcome.

- Sub-optimisation, i.e. pursuit of narrow local objectives by managers, at the expense of organisational objectives.
- Myopia, i.e. concentration on short term issues to the exclusion of long term considerations.
- Complacency, i.e. lack of ambition for improvement brought about by adequate comparative performance.
- Misinterpretation, i.e. incorrect inferences about performance brought about by the difficulty of accounting for the full range of potential influences on a performance measurement.
- Misrepresentation, i.e. the deliberate manipulation of data by provider staff, including creative accounting and fraud, so that reported behaviour differs from actual behaviour.
- Gaming, i.e. altering behaviour so as to obtain a strategic advantage, particularly prevalent when targets are based on year-on-year improvements.
- Ossification, i.e. organisational paralysis brought about by an excessively rigid system of measurement.

Understanding the performance of any organisation is complex and necessitates a thoughtful, tailored and explicit approach. Furthermore, for performance measures to have any real use, as well as being both explicit and relevant, they should also be comprehensible and manageable (Carter, 1991), in terms of collection, analysis and informing future activity.

1.5 How does one measure organisation performance?

An important conceptual issue that arises when measuring productivity is whether an efficiency or effectiveness approach should be used. Efficiency is a ratio of outputs to inputs and effectiveness is defined as the outputs relative to some standard or expectation.

Whilst productivity scholars agree that efficiency is part of the concept of productivity there is some disagreement on whether effectiveness is also part of productivity. Many of them see productivity as just efficiency (e.g., Campbell and Campbell, 1988c; Craig and Harris, 1973; Kendrick, 1984; Muckler, 1982; Werther et al., 1986). The majority, however, believe that productivity should include both efficiency and effectiveness. (e.g., Balk, 1975; Bullock and Batten, 1983; Coulter, 1979; Deprez, 1986; Guzzo, 1988; Pritchard et al., 1988, 1989, Tuttle, 1981, 1982, 1983).

Both efficiency and effectiveness have their advantages and disadvantages (e.g. Deprez, 1986; Kendrick, 1984; Kopelman, 1986; Norman and Bahiri, 1972; Tuttle, 1981). Efficiency is defined as a ratio of outputs to inputs. It is easy to calculate, easy to understand and is accepted by organisational personnel. In Data Envelopment Analysis (DEA) the best ratio of output to input forms the benchmark and then the efficiency measure of each unit is based on how close the ratio of its output to input comes to the benchmark ratio. There are, however, a number of disadvantages to efficiency measurements. One important disadvantage is that being highly efficient in the short term may be highly dysfunctional in the long term. An organisation may appear efficient but in doing so may be allowing quality of product to fall, allowing equipment to deteriorate or failing to cultivate customers. Not expending resources on such needs will be to the long-term detriment of the organisation. Other disadvantages of efficiency measures can be that they ignore demand for products or services, that they may fluctuate due to factors beyond the control of the organisation and that it is difficult to include quality in an efficiency measure. It should be possible however to build in for exogenous factors such as these in the performance measurement model.

Effectiveness is a much broader concept because it includes other factors such as standards, objectives of the organisation, expectations of interested parties and the viability of the organisation relative to its competition. Quality can also be readily included and it does not have the problem of getting accurate and meaningful inflation-adjusted prices for all inputs and outputs of the unit. However, how quality would be included in the analysis has long been a source of debate. Dyson et al. (2001) highlighted a number of pitfalls when incorporating qualitative variables into an analysis. They raised two distinct problems when measuring customer perception of quality. The first being that these measures are often treated as conforming to conventional data or interval scales while it is difficult to assert in many cases that the quantification techniques used yield anything more than ordinal data. Secondly, the measurement of qualitative data is often highly subjective, as the value scales of those involved in the ratings may differ from decision making unit to decision making unit. They used as an example the differing expectations of customers from various bank branches and stated that the same satisfaction rating in different branches may correspond to different levels of service quality delivery.

> *Using data from surveys often attempting to characterise qualitative variables, may result in an unfair DEA evaluation, as the underlying scale used by the different customers depends on their expectations, and therefore is not identical for all decision making units*
>
> (Dyson et al., 2001: 251).

They advised using surveys with care, attempting to cover a large number of respondents with an instrument designed to reduce the effect of subjectivity on the measurement process.

Effectiveness measures also have their disadvantages. To use effectiveness measures, it must be possible to identify meaningful organisational goals and develop measures that are consistent with these goals. In meeting these goals it is also important to consider whether the quantity of resources used in doing so are in the best interests of the organisation. An organisation could meet its goals very well but use far too many resources in so doing. Thus effectiveness can be just as dysfunctional as efficiency when used alone.

1.6 Organisation Model

Another issue that arises in the productivity literature has to do with the model organisations used as the basis for conceptualisation and measurement. Pritchard (1990) focussed on three models:

(a) the natural systems model (Campbell, 1977), (b) the multiple constituency model (e.g. Connolly et al., 1980; Keeley, 1978; Pennings and Goodman, 1977) and (c) the goal oriented model (Campbell, 1977).

The natural systems model assumes that the demands on an organisation are so complex and changing that it is not possible to identify a finite set of organisational goals that are definable in any meaningful way. Instead this model assumes that the overall goal of the organisation is survival.

The multiple constituency model considers the organisation as being influenced by groups of individuals internal and external to the organisation, such as managers, employees, customers and so forth, each with their own goals based on their own self-interests.

The goal-oriented model assumes that the organisation is run by a set of rational decision makers who have a manageable set of goals for the organisation that can be defined well enough to be understood and that it is possible to develop a strategy to achieve these goals. Organisational effectiveness can be thought of as the degree to which these goals are met.

In considering which model to use it is necessary to decide if a single set of usable goals exist and then to try to identify them from the organisational members. As a single set of usable goals does not exist for the natural systems model or the multiple constituency model, the effectiveness approach to productivity is not really possible. Efficiency measures have a similar problem as they assume a fixed and usable goal of producing the most of what the organisation produces with the least amount of organisational resources. The majority of the literature on productivity assumes the goal-oriented approach to organisations as the organisational model to use in conceptualising productivity.

Another core issue in the conceptualisation of organisational productivity is what unit of analysis to use. Some authors (e.g. Campbell and Campbell, 1988c; Gabris et al., 1985) go so far as to say that productivity should not be used at the individual level of analysis. An argument in support of this position is that the vast majority of work is done interdependently and thus it is difficult if not impossible to identify the contributions of individuals to the joint process. The contrasting position (e.g. Kopelman, 1986) would be that it is just as conceptually meaningful to discuss the efficiency or effectiveness of an individual as an organisation or country. At the end of the day the unit of analysis to use will be dependent on the context involved.

It is also important to look at the impact of teams at the level of the organisation. In recent research West et al. (2006) presented data from over 500 health care organisations that suggested that the extent of team-based working could positively influence organisational-level outcomes when teams had clear and appropriate inputs and processes, but that team-based working in which inputs and processes were unclear could have negative consequences for organisations (e.g. overall performance, medical errors, patient mortality).

1.7 Framework

Identifying all of the above issues should provide a framework for understanding where each approach to productivity is applicable in any given situation. Pritchard (1990) and others (Belcher, 1982; Campbell and Campbell, 1988c; Mahoney, 1988) would argue that no one true conceptualisation and measurement approach to productivity exists. Productivity is an index of how well an organisation is operating. It is however necessary to identify and agree on what functions within the organisation we are interested in before we can agree on how to measure them. Pritchard (1990) sets out five major possible purposes for measuring productivity. These are:

(a) comparing large aggregations of organisations
(b) evaluating the overall productivity of individual organisations for comparison with each other or with some standard
(c) gaining management information
(d) controlling parts of the organisation
(e) use as a motivational tool

Each of these purposes may require different productivity measurement systems. The aim of this thesis was to evaluate the overall productivity of organisations, which are hospitals in this case, for comparison with each other. The objectives of the research were to develop a comprehensive model for measuring hospital performance and using this model to measure the performance of public acute hospitals in Ireland in 2007.

1.8 Measurement

There are a number of measurement topics that need to be considered when dealing with productivity. Scope of the measurement system is important. It is crucial that the system includes measurement of all of the important functions in the organisational unit. It is also of great importance that care should be taken in what is measured so as not to result in unintentional negative consequences. This could arise where feedback results in improvements in one unit of the organisation to the detriment of other units.

Reliability, validity and generalisability are standard criteria for good measures.

> *The concept of reliability has to do with how well you have carried out your research project. Have you carried it out in such a way that, if another researcher were to look into the same questions in the same setting, they would come up with essentially the same results (although not necessarily an identical interpretation)? If so, then your work might be judged reliable.* (Blaxter et al., 1996: 200)

Reliability relates to the consistency of the construct measurement and the extent to which it is free of error.

> *Validity, from a realist perspective, refers to the accuracy of a result. Does it really correspond to, or adequately capture, the actual state of affairs? Are any relationships established in the findings true, or due to the effect of something else?* (Robson, 2002: 100)

> *Generalisability refers to the extent to which the findings of the enquiry are more generally applicable, for example in other contexts, situations or times, or to persons other than those directly involved.* (Robson, 2002: 100)

Unless a measure is reliable, it cannot be valid. However, while reliability is necessary, it is not sufficient to ensure validity. Unreliability may have various causes such as participant error and participant bias as well as observer error and observer bias. Participant error can arise where a participant’s performance might fluctuate from occasion to occasion on a more or less random basis. There are ways of ensuring that these types of fluctuations do not bias the findings, particularly when specific sources of error can be anticipated. A more problematic issue from a validity perspective is sources of participant bias. This can arise where participants may make a particularly strong effort at improved performance in order to ensure a good result. Here it can be difficult to decide if the result is due to this short-term effect or whether it will be more long lasting. Observer error could lead to random errors. Again there are ways of dealing with this type of error. Observer bias like participant bias can lead to problems of interpretation. This area can be addressed with procedures including a “blind” assessment or the use of independent assessors. Having made a serious effort to get rid of participant and observer biases and demonstrated the reliability of the measure the next question to be addressed is whether it has validity. Does it have construct validity, predictive criterion validity, internal validity or external validity. Does it measure what you think it measures, i.e. does it have construct validity. Whilst there is no easy single way of determining construct validity one might look for what seems reasonable, look at possible links between results or look at how well the results predict future performance. This latter measure is called predictive criterion validity.

Any one way of measuring or gathering data is likely to have its weaknesses. One way of addressing this is to use multiple measures. Getting similar results from using different measurement methods increases confidence in the validity of the results. Using multiple measurement methods, however, is not a panacea for all methodological ills as each method can raise its own theoretical problems and in many cases can be impracticable.

Having demonstrated satisfactorily that we have a valid measure we need to find out whether the study can plausibly demonstrate a causal relationship between the treatment and the outcome, i.e. having internal validity. We need to find out whether the treatment involved in the research question actually caused the outcome. The term “internal validity” was introduced by Campbell and Stanley (1963) who provided an analysis of possible threats to internal validity. These threats are other things that might happen which confuse the issue and make us mistakenly conclude that the treatment caused the outcome. Campbell and Stanley (1963) suggested eight possible threats to internal validity and Cook and Campbell (1979) developed this analysis, adding four further threats. All twelve possible threats are:

(1) History. Things that have changed in the participants’ environments other than those forming a direct part of the enquiry.

(2) Testing. Changes occurring as a result of practice and experience gained by participants on any pre-tests.

(3) Instrumentation. Some aspect(s) of the way participants were measured changed between pre-test and post-test.
(4) Regression. If participants are chosen because they are unusual or atypical, later testing will tend to give them less unusual scores.
(5) Mortality. Participants dropping out of the study.
(6) Maturation. Growth, change or development in participants unrelated to the treatment in the inquiry.
(7) Selection. Initial differences between groups prior to involvement in inquiry.
(8) Selection by maturation interaction. Predisposition of groups to grow apart.
(9) Ambiguity about causal direction. Does A cause B, or B cause A?
(10) Diffusion of treatments. When one group learns information or otherwise inadvertently receives aspects of a treatment intended only for a second group.
(11) Compensatory equalisation of treatments. If one group receives special treatment, there will be organisational and other pressures for a control group to receive it.
(12) Compensatory rivalry. As above but an effect on the participants themselves. This is referred to as the “John Henry” effect. John Henry was a legendary black railroad steel worker who swung his hammer in competition with a steam drill, which was introduced experimentally to replace human steel drivers. While he outperformed the steam drill he died from overexertion. “*The John Henry effect is used to describe the above average performance by a control group placed in competition with an experimental group using an innovative procedure which threatens to replace the control procedure”.* (Saretsky, 1972: 579)

In general there are two strategies for dealing with these threats. The first is that if you know what the threat is you can take specific tests to deal with it. The second strategy is the use of randomisation, which helps to offset many unforeseen factors. If we can rule out these threats we establish internal validity. We will have shown that a particular treatment has caused a certain outcome.

Campbell and Stanley (1963) use the term external validity to describe generalisability. Internal and external validity tend to be inversely related in the sense that the various controls imposed in order to bolster internal validity often fight against generalisability. LeCompte and Goetz (1982) have provided a classification of threats to external validity similar to those given for internal validity. These threats are:
(1) Selection. Findings being specific to the group studied.

(2) Setting. Findings being specific to, or dependent on, the particular context in which the study took place.
(3) History. Specific and unique historical experiences may determine or affect the findings.
(4) Construct effects. The particular constructs studied may be specific to the group studied.

There are two general strategies for addressing these potential threats. These are “direct demonstration” and “making a case”. Direct demonstration involves carrying out further study involving some other type of participant, or in a different setting etc. Making a case is putting forward persuasive argument that it is reasonable from the results to generalise, by showing that the group studied or setting or period is representative. A study may also be repeated with a different target group or in a deliberately different setting to assess the generalisability of its findings.

It is difficult to be sure that a piece of qualitative research is valid. It is possible to recognise situations which make validity more likely whilst at the same time it is difficult to state unequivocally that it is accurate, correct or true.
“Validity in qualitative research has to do with description and explanation, and whether or not a given explanation fits a given description. In other words, is the explanation credible?” (Denzin and Lincoln, 1998: 50)

“Discussions of validity concern the philosophy that the researcher uses, and a broader philosophy of what constitutes truth” (Perakyla, 1997: 50)
It is much easier to come up with factors that are likely to lead to invalid research. As with quantitative research these can be thought of as threats to validity. Maxwell (1992) has presented a useful typology of the kinds of understanding involved in qualitative research. The main types are description, interpretation and theory, each of which has particular threats to its validity. The main threat to providing a valid description of what you have seen or heard lies in the inaccuracy or incompleteness of the data. This suggests that audio- or video-taping should be carried out wherever feasible. The main threat to providing a valid interpretation is that of imposing a framework or meaning on what is happening rather than this occurring or emerging from what you have learned during your involvement with the setting. Maxwell (1996) stated how one might go about demonstrating the validity of your interpretation:

"In my view, validity of interpretation in any form of qualitative research is contingent upon the end product including a demonstration of how that interpretation was reached". (Maxwell, 1996: 150)

The main threat to validity of theory is not considering explanations or understandings of the phenomena you are studying. This can be countered by actively seeking data which are not consonant with your theory.

1.9 Design

Design concerns the various things that should be thought about and kept in mind when carrying out a research project. Robson (2002) puts forward the following components for a research model:

(1) Purpose. What is this study trying to achieve? Why is it being done? Are you trying to assess the effectiveness of something?

(2) Theory. What theory will guide or inform this study? How will you understand the findings? What conceptual framework links the phenomena you are studying?

(3) Research questions. To what questions is the research geared to provide answers? What do you need to know to achieve the purpose(s) of the study? What is it feasible to ask given the time and resources that you have available?

(4) Methods. What specific techniques will you use to collect data? How will the data be analysed? How do you show that the data are trustworthy?

(5) Sampling strategy. From whom will you seek data? Where and when? How do you balance the need to be selective with the need to collect all the data required?

A good design framework will have high compatibility between each of these aspects. A fixed design strategy calls for a tight pre-specification before you reach the main data collection stage. If you cannot pre-specify the design you should not use the fixed approach. Data are almost always in the form of numbers; hence this type is commonly referred to as a quantitative strategy. A flexible design evolves during data collection. Data are typically non-numerical; hence this type is often referred to as a qualitative strategy.

Traditional fixed design strategies would be experimental and non-experimental. Robson defines an experimental strategy as: *"The central feature is that the researcher actively and deliberately introduces some form of change into the situation, circumstances or experience of participants with a view to producing a resultant change in their behaviour".* (Robson, 2002: 88)

Robson (2002) defines a non-experimental strategy as: *"The overall approach is the same as in the experimental strategy but the research does not attempt to change the situation, circumstances or experience of the participants".* (Robson, 2002: 88)

Three traditional flexible design strategies would be case study, ethnographic study and grounded theory study. Case study is where a researcher develops a detailed intensive knowledge about a single case, or of a small number of related cases. Ethnographic study is where a researcher seeks to capture, interpret and explain how a group, organisation or community live, experience and make sense of their lives and their world. Grounded theory study is where the central aim is to generate theory from data collected during the study.
Each of these strategies represents different ways of collecting and analysing empirical evidence and each has its own particular strengths and weaknesses. The research questions themselves will influence the choice of strategy. The research question in this thesis followed a fixed design non-experimental strategy.

There are also a number of important design criteria for a productivity measurement system. Firstly the option of using either a single index or multiple indices of productivity needs to be considered. Secondly as different activities within an organisational unit are not of equal importance there needs to be some method of weighting the importance of each activity. However, this information may not be necessary if the DEA model is used. Thirdly there is frequently not a linear relationship between the level of input an organisational unit puts into an activity and the contribution that level of activity makes to the organisational unit. This can arise for example where the value of the unit's output gets higher and higher until it reaches a point of diminishing returns and from that point on further increases in output quantity are not as valuable. Pritchard et al. (1988, 1989) found that none of the forty five indicators in their study of five organisational units were linear. This issue needs to be borne in mind when designing any productivity measurement system. The productivity measurement and enhancement system (ProMes) as developed by Pritchard et al. (1989) takes account of non-linearity. Finally the productivity measurement system should be capable of aggregating the measurement system of different units into a single broader measurement system.

1.10 Measurement Model

The technical efficiency of the hospitals included in this research was assessed utilising data envelopment analysis (DEA) methodology. DEA is a non-parametric linear programming technique which identifies best practice within a sample and measures efficiency based on differences between observed and best practice units, and is typically used to measure technical efficiency.

1.11 Conclusion

Efficiency, effectiveness and quality measures are of critical importance and must be reflected in any measure of organisational productivity. The research problem was to find a method of measuring hospital productivity using Pritchard's (1990) definition of productivity, i.e. how well a system uses its resources to achieve its goals. DEA methodology was used to carry out this assessment. The research involved collecting not only patient activity data, that is readily available, but also subjective data based on people's perception of relevant performance measures. Much of the patient outcome and activity data were available from secondary sources, whilst, other measurements such as stakeholders' opinions on relevant input and output measures had to be sought through the use of questionnaires and focus groups.

In using any qualitative methodology I needed to ensure that I recognised that I am part of the research process and that I accounted for my own feelings, emotions and reflections in the design protocol and in the subsequent interpretation of the results. In order to ensure the validity of the choice of input and output measures I issued questionnaires to relevant stakeholders and established expert focus groups to help me decide on the most relevant input and output measures to be included in the final performance measurement model.

In this introduction the research objectives have been set out, the main concepts surrounding performance measurement have been discussed and the approach to the research has been outlined. In chapter 2, the Irish health care context will be described.

Chapter 2

Irish Healthcare Context

2.1 Introduction

The Irish hospital system comprises three types of hospitals. These are statutory hospitals, owned and administered by the Health Service Executive (HSE), voluntary hospitals, owned and operated by lay boards or religious orders, and private hospitals. This research will only be addressing the statutory and voluntary hospitals, as private hospitals typically do not participate in the provision of data that will form the basis for this analysis. Statutory hospitals can be broken down into regional hospitals, county hospitals, district hospitals and what would be classified as special hospitals. Voluntary hospitals have their origins in the 18^{th} century and were generally run by philanthropic individuals who recognised the need to provide hospital services for the poor. These hospitals were funded largely through general fundraising and contributions from wealthy individuals. This type of service had not been available in Ireland to the public since the closure of the monasteries during the reformation period, unlike the situation in Britain where, under Elizabethan legislation, a system of rate supported public parochial assistance had been devised.

Regional hospitals cater for a wide population base, tend to be major trauma centres, have specialised units and are major teaching hospitals. County hospitals have tended to have consultant led services for general medicine, general surgery, obstetrics and gynaecology. District hospitals are not included in this study as they are mainly catering for long stay, non-acute, patients. Those hospitals that have been classified as special would be single specialty hospitals that would include maternity, paediatric, orthopaedic and eye, ear and throat hospitals. Voluntary hospitals, which are generally located in large centres of population in Dublin, Cork and Limerick are general hospitals and often function as teaching hospitals.

Statutory hospitals are owned and funded by the HSE whilst voluntary hospitals are funded, supported but not owned by them. Prior to the establishment of the HSE in 2005 statutory hospitals outside of Dublin were funded by regional health boards, all hospitals in the Dublin region were funded by the Eastern Regional Health Authority and voluntary hospitals outside of Dublin were funded directly by the Department of Health and Children. Since 2005 the HSE controls the funding of all hospitals. This change has led to a new transparency where information regarding the operation of all hospitals has been made more readily available. This in turn has allowed more comparability between hospitals and exposed all hospitals to closer scrutiny. It is in this changing environment that it has become critical to be able to identify those factors that are affecting a hospital's performance.

In order to understand the size and complexity of the Irish hospital system, time series data on hospital bed complements, inpatient discharges, day case discharges, ratio of acute hospital beds to population, public health expenditure and the general hospital's programme non-capital expenditure are presented graphically in the following figures 2.1 to 2.10. In this chapter details of the acute hospital's Casemix Efficiency Model which is being used in Ireland has also been set out.

2.2 Bed Complement

The number of approved acute inpatient beds and day beds in the system are presented in table 2.1 and figure 2.1. The number of both types of bed increased over the period 1997 to 2006, inpatient beds by 13.1% and day beds by 96.2%.

Table 2.1

Total number of inpatient and day case beds: Ireland: 1997 – 2006

Year	Day Case Beds	Inpatient Beds
1997	610	11,121
1998	636	11,051
1999	673	11,058
2000	721	11,190
2001	771	11,373
2002	812	11,686
2003	909	11,806
2004	1132	11,887
2005	1253	12,094
2006	1197	12,574

Figure 2.1

Aston University

Illustration removed for copyright restrictions

Source: Health in Ireland: Key Trends
Department of Health and Children

Whilst both bed complements showed an increase, the higher percentage increase in day beds highlights a major change in the way acute healthcare is being provided. Increasingly, medical procedures that in the past required admission overnight, and longer, to a hospital can now be done as day procedures. This is largely due to advances in medicine but is also being driven by the increasing cost of healthcare and the pursuit of efficiencies. An important point to note, however, is that the number of day beds in the system was coming from a very low base and that the actual increase was from 610 to 1,197 beds, i.e. 587 additional beds. The 13.1% increase in inpatient beds equated to 1,453 additional beds, giving a total of 12,574.

Another important factor that needs to be recognised is that because the population increased by 15.7% between 1997 and 2006, as set out in table 2.2, the actual ratio of beds to the population increased by only 1.6%, from 3.20 to 3.25 beds per 1,000 people. This is graphically illustrated in figure 2.2. It is also important to note that whilst the ratio of day beds to the population increased by 64.7% the ratio of inpatient beds to the population actually reduced by 2.0%. This is illustrated in figures 2.3 and 2.4 respectively. There are therefore more day beds available for each person but fewer inpatient beds. This again highlights the movement from inpatient treatment to day procedures.

Table 2.2

Population (000s) by Age Group for Each Year, 1997 – 2006

Aston University

Illustration removed for copyright restrictions

Source: Central Statistics Office

Figure 2.2

Aston University

Illustration removed for copyright restrictions

Source: Health in Ireland: Key Trends
Department of Health and Children

Figure 2.3 shows that the ratio of day case beds to the population increased consistently between 1997 and 2005. There has however been a 6.9% reduction in the ratio between 2005 and 2006.

Figure 2.3

Aston University

Illustration removed for copyright restrictions

Source: Health in Ireland: Key Trends
Department of Health and Children

Whilst the ratio of inpatient beds to the population has reduced overall by 2% between 1997 and 2006 it is clear from figure 2.4 that it has fluctuated during this period. Between 1997 and 2000 the ratio reduced by 2.7%, between 2000 and 2002 it increased by 1%, between 2002 and 2005 it reduced by 1.9% and between 2005 and 2006 it increased by 1.3%.

Figure 2.4

Source: Health in Ireland: Key Trends**s**
Department of Health and Children

2.3 Hospital Discharges

As can be seen from Table 2.3 hospital inpatient discharges increased from 525,495 to 591,766 between 1997 and 2006, i.e. a 12.6% increase, and hospital day cases increased from 243,019 to 555,204, a 128.5% increase, over the same period. These increases closely reflect the bed complement changes and the obvious move towards day procedures.

Table 2.3

Hospital Inpatient Discharges 1997 – 2006

Aston University

Illustration removed for copyright restrictions

Source: 1997 –2005- Integrated Management Returns (IMRs), and Hospital Inpatient Enquiry (HIPE), Department of Health and Children; 2006 – National Hospital's Office, Health Service Executive.

It is clear from Figures 2.5 and 2.6 that whilst day cases increased consistently year on year the rate of inpatient increase was not as constant. Inpatient discharges only increased by 4.9% over the five year period 1997 to 2002. In fact they actually reduced between 1998 and 1999 and again between 2001 to 2002 before increasing consistently by 7.4% in the four year period from 2002 to 2006.

Figure 2.5

Aston University

Illustration removed for copyright restrictions

Source: Health in Ireland: Key Trends
Department of Health and Children

Figure 2.6

Aston University

Illustration removed for copyright restrictions

Source: Health in Ireland: Key Trends
Department of Health and Children

The ratio of Inpatient and Day Case discharges to population as set out in Figures 2.7 and 2.8 highlight even more starkly the move away from admitting patients overnight and towards day case admissions. During the period 1997 to 2006 the ratio of acute inpatient discharges reduced from 143 per 1,000 to 140 per 1,000 whilst at the same time day case discharges increased from 66 per 1,000 to 131 per 1,000. This is a 2% reduction in the inpatient ratio and a 98% increase in the day case ratio.

The average length of stay for inpatients between 1997 and 2006 as can be seen from Table 2.3 remained relatively constant at between 6.3 and 6.6 days. Over the period the figure reduced by 3.9%, from 6.5 in 1997 to 6.3 in 2006. One could argue that with advances in medical treatments that the average length of stay should be reducing at a faster rate. However, with the trend of carrying out the more straightforward cases as day cases one is left with the more complex cases that are not suitable as day cases and probably require a longer hospital stay, thus ensuring a higher overall average length of stay.

Interestingly, the reduction in the ratio of inpatient discharges to population coincides with the 2% reduction in the ratio of inpatient beds to the population over the same period. However, the increase in the ratio of day cases to the population exceeds the increase of 64.7% in the ratio of day case beds to the population. This highlights the added efficiency of carrying out medical procedures as day cases where day beds can be utilised more than once each day.

Figure 2.7

Source: Health in Ireland: Key Trends
Department of Health and Children

Figure 2.8

Aston University

Illustration removed for copyright restrictions

Source: Health in Ireland: Key Trends
Department of Health and Children

2.4 Health and Hospital Expenditure

Total public health expenditure in Ireland increased from €3,671 million in 1997 to €12,337 million in 2006. Without taking inflation into account this is an increase of 236% in total public health expenditure since 1997. During this period non-capital expenditure on General Hospitals increased from €1.8 billion to €5.4 billion, an increase of 199.6%. Figure 2.9 clearly highlights these spiralling costs. The increasing expenditure on healthcare is a major concern for most economies and Ireland is no different in this regard.

Figure 2.9

Aston University

Illustration removed for copyright restrictions

Source: Non-capital Expenditure – "Estimated Non-Capital Health Expenditure 1990 to 2006 Categorised by Programme and Service" – www.dohce.ie
Capital Expenditure – Revised Estimates for Public Services.

Figure 2.10 sets out the ratio of non-capital expenditure on the public general hospital's programme per head of population between 1997 and 2006. As can be seen from these figures the ratio of non-capital expenditure on this programme per head of population increased from €495 to €1,280 between 1997 and 2006. This is an increase of 159% over the period and an extraordinary average increase of 17.7% per annum.

Figure 2.10

Source: "Estimated Non-Capital Health Expenditure 1990 to 2006 Categorised by Programme and Service" – www.dohce.ie

2.5 Current Health Environment

The Irish Health Service is currently undergoing a period of change or as Ansoff (1985) might describe it “a period of Turbulence”. The first major changes in the structures and operations of the health service since 1970 are now underway. The government’s health strategy “Quality and Fairness” (2001), set out a framework for the development and reform of the Irish Health System. Three major reports issued during 2003 now form the basis for the current health reform programme. These are:

(1) “Audit of structures and functions in the health system” (2003) – Prospectus Report.

The recommendations of this report include:

- The creation of a single national health service executive to replace the existing health boards. This resulted in the establishment of the Health Service Executive (HSE) in January 2005.
- The strengthening of processes and capabilities to deliver value for money and to manage ongoing change.
- The strengthening of governance and accountability across the system.
- The reorganisation of existing agencies and their functions.

The main implications of this report for the acute hospital sector are:

- The acute hospital sector has now come under the auspices of the National Hospitals Office.
- The National Hospitals Office will, as well as providing all funding for acute hospital services, make recommendations in relation to the reorganisation and grouping of hospital services in each region.
- Funding for all acute hospitals will be based on contracts incorporating service agreements.

(2) “Commission on financial management and control systems in the health service” (2003) – Brennan Report.

This report included the following recommendations:

- The establishment of a health service executive.
- A range of reforms to governance and financial management control and reporting systems.

- Substantial rationalisation of existing health agencies.
- A range of changes to the current arrangements with medical consultants.

(3) "National task force on medical staffing" (2003).

This report made recommendations in relation to how the European Working Time Directive for non-consultant hospital doctors could be implemented. The directive states that the average number of hours that a doctor can work each week must not exceed 58 hours from the 1st of August 2004, and that this must be reduced to 56 from the 1st of August 2007 and to 48 hours from the 1 August 2009. Some of the main recommendations of this report are:

- Acute hospital services should be delivered by an integrated network of hospitals, currently serving populations of about 350,000.
- The organisation and staffing of acute hospitals must be restructured to allow for the safe provision of emergency and elective care.
- Substantially more medical consultants should be appointed as part of a move to a team-based consultant provided service.

These reports, which are gradually being implemented, are resulting in seismic changes to the Irish acute hospital system. The establishment of the Health Service Executive, in particular, in 2005 has radically changed hospital accountability and governance structures. In this climate, being able to accurately measure hospital performance has taken on even more importance.

2.6 Health of the population

As well as increased investment in health services the past decade has shown unprecedented improvements in health status and life expectancy. Average life expectancy in Ireland has increased to 80 years in 2009. This is up from 76.6 years in 2000 and is slightly above the EU average of 79.5 years. These details are set out in table 1.2.

Health has been defined by the WHO as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity. One method of assessing this measure is to survey people and ask them to assess their state of health. The EU survey on Income and Living Conditions, 2005, indicated that based on this measure Ireland had the highest levels of self-perceived health of those countries in Europe which have conducted such a survey. Table 2.4 indicates that over 80% of both men and women in Ireland assess their health to be either “good” or “very good”. (Department of Health and Children, 2007: 8)

Table 2.4

Perceived Health Status, 2005

Aston University

Illustration removed for copyright restrictions

Source: Central Statistics Office
EU Survey on Income and Living Conditions, 2005

Very significant long term improvements in the mortality rates for the major causes of death are also evident, with the notable exception of cancer deaths, which in common with other countries, have shown only a minor decline.

Table 2.5

Principal Causes of Death: Rates per 100,000 - 1997 to 2005

Aston University

Illustration removed for copyright restrictions

In respect of ischaemic heart disease the mortality rate per 100,000 has reduced from 194 to 113 between 1997 and 2005. The stroke mortality rate per 100,000 has reduced from 69.2 to 42.8. The rate for all diseases of the circulatory system has reduced from 351.8 to 218.2 while the rate for all cancers has reduced from 209.1 to 180.9 in the same period.

It is too early to assess the impact of the smoking ban in bars and restaurants, which was introduced in Ireland on the 26th of January, 2004 but one would expect it to have a positive impact on the health of the population.

The rising numbers of elderly people in future will also have a major impact on the planning and delivery of health services.

Environmental constraints can affect the efficiency of hospitals and are drivers of change. These include the following factors (Jacobs et al. 2006):

- population mortality rates are heavily dependent on the demographic structure of the population under consideration
- surgical outcomes are often highly contingent on the severity of the disease of the patients
- hospital performance may be related to how care is organised in the local community
- the performance of emergency ambulance services may depend on geography and settlement patterns.

2.7 Acute Hospitals Efficiency Measures – Casemix Model

The Commission on Health Funding was established in 1989 to examine the financing of the health services in Ireland and to make recommendations on the extent and source of the future funding required to provide an equitable, comprehensive and cost effective public health service and on any changes in administration which seemed desirable for that purpose. In their conclusions they stated:

> *Each hospital should then be funded for the provision of an agreed level of service to public patients, based on the activity level implied by its role and catchment area, and the casemix based cost of meeting this. Techniques such as diagnosis related Groups (DRGs) should be used to determine the level of funding required for a specified level of service.* (Commission on Health Funding, 1989: 19)

> *It is fully accepted that the clinical workload of hospitals varies greatly. Casemix is an attempt to categorise and quantify this "mix" of cases by classifying patients into discrete classes or groups (DRGs) which share common attributes and similar patterns of resource use. The development of DRGs provided the first operational means of defining and measuring a*

> *hospital's casemix complexity, and comparing it with other hospitals.* (The Modernisation of the National Casemix Programme in Ireland, 2004: 6)

In 1993 the casemix model was introduced into 15 public acute hospitals. Acute hospitals are those where the patient's length of stay would be expected to be 30 days or less. Following this, the programme has expanded each year and there are now 37 public acute hospitals participating in the National Casemix Programme. Casemix works by coding hospital activity, using the hospital inpatient enquiry (HIPE) programme, and assessing hospital costs, using the specialty costs programme. The HIPE programme currently operates in the 62 biggest hospitals in the country. When a patient is discharged from hospital their age, gender, diagnosis, procedures performed and discharge status is coded using the WHO international classification of diseases (I.C.D.) which allows for 12,000 diagnoses and 8,000 individual procedures, each of which is allocated a separate code. The data is then grouped into over 6,000 DRGs. The basis of the entire system is to break down illnesses into 25 major diagnostic categories (M.D.C.s) based around body parts. In the specialty costs programme, cost data, based on information derived from the audited accounts of 37 of the 62 HIPE hospitals is broken down across 16 cost centres and apportioned to each specialty in the hospital. These costs are then allocated to the 600 or so DRGs, giving an average cost per case.
Casemix is the combining of the activity and cost data to give an average cost per case, length of stay and resource use relative to other activity in the hospital and elsewhere. In Ireland casemix is used for acute hospital activity only. Hospital outpatient care was included in the model for the first time in 2009. Presently the 37 acute hospitals involved in the casemix programme have a percentage of their annual budgets adjusted based on their casemix performance. The entire exercise is budget neutral in that the Department of Health and Children does not gain from the exercise. Any money deducted from hospitals below the mean is given to hospitals above the mean, in an effort to reward good management. There is one overall casemix adjustment made to each hospital's annual grant at the beginning of each year.

Casemix adjusted inpatient and day case procedures as set out in the casemix model are used in the DEA models included in this research.

The casemix model is the only comprehensive measure of relative efficiency for acute hospitals being used in the Irish hospitals' context. The 37 hospitals that participate in this process are the subject of this research. This group includes all of the large and medium sized public hospitals in Ireland, both statutory and voluntary. The list of these hospitals and their casemix adjustments for 2009 are set out in Table 2.4. These results are based on activity and expenditure in 2007.

Table 2.6

Hospital Casemix Adjustments 2009

Hospital	Adjustment €
Mullingar	1,977,061
Wexford	1,234,761
Letterkenny	953,515
St. Luke's	910,912
St. James'	869,542
Cork University	669,515
Kerry	592,521
Croom	585,569
Louth	490,207
Beaumont	468,534
Mater	331,635
Galway	323,154
Mallow	318,023
Mayo	310,467
Rotunda	199,263
Waterford	197,757
Connolly	83,122
Holles Street	27,113
Temple Street	5,839
South Tipperary	-1,377
Crumlin	-5,839
Portlaoise	-8,229
Portiuncula	-140,349
Coombe	-226,376
South Infirmary	-235,567
Mercy	-245,130
Cavan	-291,400
St. Mary's	-336,895
St. Vincent's	-442,023
Sligo	-674,748
Merlin Park	-773,450
Loughlinstown	-827,643
Limerick	-839,366
Navan	-939,516
Tullamore	-1,083,815
Drogheda	-1,173,308
Tallaght	-2,303,478

The figures in this table are the monetary adjustments made in respect of each hospital which are reflective of their relative efficiency. However they are not measures of efficiency. The efficiency measure for each hospital is dependent on the proportion that each monetary adjustment represents of the total revenue of the hospital.

It can be seen from this table that the casemix adjustment for 2009 varied from a positive adjustment in Mullingar hospital of €1,977,061 to a negative adjustment in Tallaght hospital of €2,303,478. This would appear to indicate that Mullingar is a highly efficient hospital and that Tallaght is a highly inefficient one. But is Mullingar producing a better performance than Tallaght? This may in fact be the case but this conclusion is open to question. The casemix measure has a number of weaknesses not least being the fact that it does not measure health outcomes nor does it take into account any qualitative factors. It concentrates solely on quantitative measures. It is therefore quite possible that a hospital appearing to be the most efficient using the casemix model could be the most dangerous from a patient perspective with poor health outcomes. We therefore need to look at what we are measuring when analysing a hospital's performance. This is a critical part of this research.

Other issues that could affect the results in the casemix model include the weightings used, lack of demarcation between treatment areas, hospital groupings and the urban/ rural differences. The weightings used in the model will have major implications for the results. Following a major review of casemix in 2004 by the Department of Health and Children it was decided that commencing on the 1st of January 2005 an Australian casemix system, ICD-10-AM, would be used (The Modernisation of the National Casemix Programme in Ireland, 2004, Department of Health and Children). This system which is applied consistently across all acute hospitals ensures comparability between hospitals. Casemix clearly defines groups of patients and their related costs. However, this can lead to conflicts with clinicians where they might see incorrectly the re-classifying of some of their work under casemix as a downgrading of the work. Similarly the re-classifying of certain clinical procedures between day case procedures and outpatient side-room procedures can lead to disagreements with hospital management because of the potential negative impact on the casemix adjustment. This arises because day case procedures have casemix weightings applied to them while outpatient attendances do not. Hospital groupings are another issue for debate. Should hospitals be grouped at all? Should teaching hospitals be treated differently to non-teaching hospitals? One argument would be that casemix does not reflect the level of teaching status and associated costs. However, the true meaning of teaching is ill defined and there is a lack of agreement on the implications and the resources required. In Ireland teaching hospitals are grouped separately to other acute hospitals and their casemix adjustments are confined to within their own group. This obviously affects the results of all hospitals participating in the casemix model programme. Finally the urban and rural differences can affect the casemix results. Internationally this issue relates to countries that have significant distances between hospitals. The Irish casemix model does not take into account any urban/rural divide, presumably on the basis that no such divide exists.

2.8 Healthstat Hospital Dashboard

Since 2008 the Health Service Executive have been developing and implementing a monthly healthstat hospital dashboard for twenty nine public acute hospitals in Ireland. These dashboards record the performance of each hospital across a number of metrics and allow each of them both to monitor their own performance and to compare their performance with that of other hospitals. The dashboard is centred around three key themes and within each theme there are a number of metrics. These are:

(a) Access - The waiting times experienced by people using hospitals.
(b) Integration - The patient journey once in the system.
(c) Resources - The right people in the right place, value for money, and the effectiveness of applied resources.

The aim of healthstat is to share best practice and address problem areas in specific hospitals in a positive way. What sets healthstat apart from previous individual hospitals' performance measurement systems is the specific focus on follow-up. The healthstat forum and the online publication of results encourage hospitals to work for consistent performance improvement.

Whilst the development of healthstat is a welcome improvement to the performance measurement system in Ireland it suffers from the major disadvantage that it provides only a partial measure of performance. This runs the risk of leading to tunnel vision induced by focusing on partial indicators of performance which could result in sub-optimal decision-making. It does not provide a comprehensive measure of organisational efficiency and therefore different indicators may provide conflicting messages.

2.9 Conclusion

It is clear that between 1997 and 2006 major changes have taken place in the acute hospital system in Ireland. There has been an obvious move towards the provision of more day procedures as distinct from inpatient admissions. The 13.1% increase in inpatient beds in the period is dwarfed by the 96.2% increase in day beds. When increases in population are factored in there is a 64.7% increase in the ratio of day beds per head of population while the ratio of inpatient beds to population actually reduced by 2%. These changes are reflected in the fact that hospital inpatient discharges increased by 12.6% and hospital day cases increased by 128.5% during this period. At the same time the average length of stay for an inpatient remained fairly constant, only reducing by 3.9% from 6.5 days in 1997 to 6.3 days in 2006. One reason for such a small change may be that the more straightforward medical procedures are now being carried out as day cases whilst the more complex cases still require overnight admission.

Expenditure on public health in Ireland has increased by 236% between 1997 and 2006. Expenditure on public general hospitals has increased from €1.8 billion in 1997 to €5.4 billion in 2006. This is an increase of €785, from €495 to €1,280, on general hospitals per head of population during this period. This huge increase in health expenditure with no apparent improvement in the service has led to many questions being raised and has resulted in a much closer examination of hospital performance. The radical changes in the health service structures and operations since the establishment of the Health Service Executive in 2005 have also focussed more attention on hospital efficiencies. There are now more demands being placed on hospitals to show that they are delivering a superior performance. The casemix model goes some way towards providing an indication of the relative efficiency of hospitals but it does not take all necessary factors into account. Similarly the healthstat dashboard system, whilst improving the availability of information on performance in specific areas, it does not provide a comprehensive measure of performance.

The information and graphs contained in this chapter highlight how dynamic the Irish health system has been during this period in terms of resources used and patients treated. This constant system change adds to the complexity in drawing overall conclusions on productivity efficiency within the acute hospital sector. Given this fact the results of the research will be based on a snapshot picture of the performance of each hospital during a specific period that may be affected by adjustments to changes in the health system.

Chapter 3

Literature Review

3.1 Introduction

The increasing cost of healthcare has become a major political issue in most Western countries. The World Health Organisation states: *"Better health is unquestionably the primary goal of a health system. But because health can be catastrophically costly and the need for it unprecedented, mechanisms for sharing risk and providing financial protection are important."* (WHO, 2000: 21)

Health spending has come more and more under the microscope and the level of performance of many healthcare organisations has been called into question. *"Spiralling health care costs are causing worldwide concern, and a key component of health sector reform efforts in many countries has to do with making the best use of existing resources."* (Parker and Newbrander, 1994: 107)

Good performing organisations have been rewarded whilst bad performing organisations have been penalised. But what does "good" or "bad" performance mean? To answer this question one must look at the various stakeholders with an interest in performance within the health service. These audiences can broadly be broken down into those in governance roles; managers and providers; health care professionals, and patients and their carers. Each group is likely to have a different perspective on what indicates a "good" performance. Those in governance roles represent the electorate and the taxpayers and would have a particular interest in the impact of government health care policies on performance. Managers are agents for the owners but may also have other agendas. Performance indicators are a means of expressing owners' interest and preferences and these may be linked to the remuneration of the managers, as in performance related pay. The role of clinicians is complex in that while they are acting in the patient's best interest they as managers may be expected to attain wider goals. Patients and their carers are more concerned with performance specifically in relation to the services they are seeking or are receiving. The elements of performance with which they are concerned are also likely to be different. The areas of particular interest to them would be access, effectiveness, patient-experienced quality and clinical outcomes.

Aaron and Ginsburg (2009) pose the question: is health spending excessive and if so what can be done about it? While they accept that the case that the United States spends more than is optimal on health care is overwhelming the challenge is how to lower spending without lowering net welfare. They state:

> *If spending is rising and if that is problematic the practical questions are as follows: what exactly is wrong with spending more on some good than one spent in the past? And what tools are available to control spending on something that is beneficial on average but not for each patient?*
> (Aaron and Ginsburg, 2009: 1260)

3.2 Why measure performance?

In the UK pressure to improve NHS efficiency stems from concerns about “unacceptable variations” in the standards of services provided across the health service. Evidence of variations includes differences in survival rates, rates of treatment and unit costs. Broadly speaking assessments of organisational efficiency can be drawn from two types of data. These would firstly be performance indicators, which measure specific data and secondly comprehensive measures, designed to provide an indication of overall organisational efficiency. One needs to be careful when interpreting performance indicators. Traditionally in the hospital sector partial measures of performance, such as the average length of stay or day case activity, have been used to make inferences about overall organisational efficiency. These measures however do not provide a comprehensive view of organisational efficiency and different indicators may in fact provide conflicting results.

3.3 Why does a hospital not perform to its optimal ability?

Debreau (1951) gave two principle reasons why deviations from optimal performance occur. Firstly, market failure and secondly non-profit maximising firm behaviour. Both of these reasons for failure to achieve optimal performance are pertinent in the health service. Market failure exists because individuals consume healthcare not for its own sake but to improve their health. Non-profit maximising firm behaviour arises as in many cases healthcare services are provided in public institutions where the principal aim of the doctor is neither to optimise profit nor to optimise resource utilisation but to maximise the welfare of the patients treated. The Department of Health and Children (2001) set out four principles that guided the development of the Irish health strategy. These principles were equity, people-centredness, quality and accountability. Equity, people-centredness and quality fit together well and complement each other. However, there is an obvious potential for conflict between them, and equity in particular, and the principle of accountability. The strategy, when dealing with accountability, states that better planning and evaluation models must demonstrate that available resources are used as efficiently and effectively as possible. This may not always tally with the aim of equity. In relation to equity the WHO (1999) states that everyone should have a fair opportunity to attain full health potential and, more pragmatically no one should be disadvantaged from achieving this potential, if it can be avoided. Inequity refers to differences in health which are not only unnecessary and avoidable but, in addition are considered unfair and unjust. It may not, for example, be efficient to provide an equitable health service to elderly patients who may be much more expensive to treat than younger patients. Similarly it may not be very efficient to provide a health service to remote areas or to marginalised communities such as immigrants or members of the traveller community. Justifying the provision of very expensive medicines or procedures in different situations may conflict with the principle of accountability. In these situations there often needs to be a trade-off between efficiency and equity.

For all of these reasons healthcare institutions are particularly suspect of inefficiency and low productivity.

3.4 How does one measure performance?

Various methods have been used to measure such performance. In the hospital sector, in particular, the use of performance indicators has become a fact of life. In many countries, including Ireland, league tables have been introduced that compare the performance of each hospital.

Traditionally both in Ireland and in the UK health agencies have relied on partial measures of performance, such as length of patient stay, day case activity and waiting times, to make inferences about organisational activity. There are two major drawbacks in using such partial measures. Firstly they do not provide a comprehensive view of organisational efficiency and secondly they may provide conflicting messages. An organisation may appear to be performing well according to one indicator but may appear to be performing less successfully according to another. It is therefore not straightforward to draw conclusions about an organisation’s overall performance from a narrow range of indicators.

Many of these performance measurements have concentrated on measuring efficiency rather than looking at the quality of service being provided and the outcome for the patient. This over-concentration on efficiency has the potential to result in short-term decision making at the expense of long-term sustainability of optimal performance.

The perceived weakness in using partial measures of performance has driven policy-makers to consider the possibility of devising comprehensive measures of performance. Two supposedly comprehensive measures that have previously been devised for the NHS were the labour productivity index (LPI) and the purchaser efficiency index (PEI). The LPI measured the ratio of cost-weighted activity to the number of employees whilst the PEI reported the percentage change over time of cost weighted activity by the percentage change over time in real funding. The LPI and PEI suffered from two problems. The selection of weights was likely to be controversial and the indices assumed that a simple relationship between outputs and inputs held at all levels of operation.

3.5 Overview of performance measurement

Two general approaches are available to measure overall efficiency. These are parametric (econometric) methods, such as Stochastic Frontier Analysis (SFA), and non-parametric methods, such as Data Envelopment Analysis (DEA). Both of these attempt to measure efficiency by estimating the optimal level of output conditional upon the amount and mix of inputs. There is no consensus on which of these is the most appropriate technique as each has its own strengths and limitations. Parametric techniques require more decisions to be made regarding functional form or the distribution of error term, but these decisions can be tested. In contrast, there are no standard procedures available to guide model construction in the non-parametric framework.

In 2000 in the UK, the Public Services Productivity Panel produced a report in which the efficiency of the police service was analysed (Spottiswoode, 2000). This study recommended the "joint use of two of the most advanced relative efficiency measuring techniques – Stochastic Frontier Analysis and Data Envelopment Analysis".

The parametric approach to efficiency measurement can be divided into two alternative estimation techniques: Corrected Ordinary Least Squares (COLS) and Stochastic Frontier Analysis (SFA). When there is only cross-sectional data available COLS and SFA are two classes of econometric technique available. Both follow the same general process:

- Identify a dependent variable (y)
- Specify a set of explanatory variables (x) that are thought to explain or predict differences in output or cost

- Interpret residual differences between observed and predicted output or cost as arising from either measurement error or inefficiency (c).

The dependent and independent variables are then related by specifying an econometric model of the general form:

y = a + bx + c

where y is a measure of output or cost , a is a constant, x is a vector of explanatory variables (e.g. labour, capital and materials), b captures the relationship between the dependent and explanatory variables and c is a residual representing the deviation between observed data and the relationship predicted by the explanatory variables in the model. Data on y and x observed at hospitals is used to estimate the parameters a and b. In most statistical or econometric models of this form the relationship between y and x are the primary focus. Generally, the residual c is not afforded attention in its own right with researchers interested only that it satisfies classical assumptions of having zero mean and constant variance (Cook and Weisberg, 1982).

> *In efficiency analyses, by contrast, the residual is often the only parameter of interest and it is from the residual that estimates of efficiency are derived. The difference between COLS and SFA rests upon the interpretation accorded to the residual. In COLS the entire residual is interpreted as arising from inefficiency. In SFA, the residual comprises a mixture of inefficiency and measurement error.*
>
> (Jacobs et al. 2006:41)

Jacobs et al. (2006) set out a number of considerations when estimating efficiency using the SFA model. These are:

- whether to estimate a production or cost function
- whether to transform variables
- whether to estimate a total or average function
- which explanatory variables to include
- how to model the residual
- how to extract efficiency estimates

DEA is a non-parametric linear programming approach that was first introduced by Charnes, Cooper and Rhodes in 1978 for constant returns to scale technologies and modified by Banker, Charnes and Cooper in 1984 for variable returns to scale technologies. This technique identifies best practice within a sample and measures efficiency based on differences between observed and best practice units and is typically used to measure technical efficiency. Using DEA the efficiency measure is related to best practice and not average practice. One of its main advantages is that it can readily incorporate multiple inputs and outputs and, to calculate technical efficiency, only requires information on output and input quantities and not prices.

Other advantages of the model are that firstly possible sources of inefficiency can be determined as well as identifying efficiency levels and that secondly by identifying the "peers" for organisations that are not observed to be efficient DEA provides a set of potential role models that an organisation can look to, in the first instance, for ways of improving its operations. Whilst the DEA model may appear to be more flexible than the parametric method it does have its disadvantages. Because DEA generates efficiency scores by comparing an organisation with its peers a result showing full efficiency will be generated if no peers exist. Similarly, when assigning an inefficiency score to an observation lying off the frontier, only its peers are considered, with information pertaining to the remainder of the sample discarded. In contrast, the parametric approach appeals to the full sample information when estimating relative efficiency. SFA is also to be preferred in situations where there is likely to be a high degree of measurement error as DEA does not recognise the possibility of measurement error. DEA scores are also sensitive to output and input specification and the size of the sample. DEA is based on the simple notion that an organisation that employs less input than another to produce the same amount of output can be considered more efficient. Among the most important considerations when undertaking DEA are:

- Choice of inputs and outputs
- Whether to assume constant or variable returns to scale.

If longitudinal data are available some of the strong assumptions required for the SFA model of efficiency can be relaxed. Jacobs et al. (2006) quote Schmidt and Lin (1984):

> *Repeated observations of the same organisation make it possible to control for unobservable organisation-specific attributes and, thereby, to extract more reliable parameter estimates, both of the explanatory variables and of the efficiency term. Specifically three shortcomings of cross-sectional analysis can be addressed.*
>
> (Jacobs et al. 2006: 69)

The first shortcoming that can be addressed is that when only a single observation is available per organisation, it is necessary, in order to partition the composite error term, to specify how inefficiency is distributed among organisations. However, there is no economic basis for selecting one distribution over another and the choice is somewhat arbitrary (Schmidt, 1985). Repeated observations of the same organisation can substitute for distributional assumptions if the fixed-effects panel data estimator is used. The second shortcoming that can be addressed is that under some formulations of the production model the inefficiency term and the explanatory variables are unlikely to be independent. For instance, it is quite likely that if an

organisation knows its level of technical efficiency this will affect its choice of input levels. Again the use of the fixed-effects estimator makes it possible to avoid the assumption of independence. The third shortcoming that can be addressed is that with cross-sections, only the entire residual can be estimated consistently, with the variance of the conditional distribution of the inefficiency term failing to become zero as the sample size approaches infinity. With panel data, adding more observations from the same organisation generates more information about each organisation so that the inefficiency term can be estimated consistently as the number of observations over time approaches infinity (Jacobs et al. 2006).
DEA can also be applied to panel data, by calculating what is known as the Malmquist index. Changes in productivity over time can be attributed to three separate explanations (Fare et al. 1994, Giuffrida 1999). First, the technical efficiency of an organisation may change, at a given scale of operation. Second, the efficiency of an organisation may change in response to a change in the scale of operation. Third, the underlying technology may change, inducing a shift in the production frontier, which will affect the efficiency of all organisations. The Malmquist index provides estimates of each of these effects by calculating separate distance functions in each period and by varying the assumptions about the available technology.
For these reasons time series analysis is always to be preferred to cross sectional analysis in both SFA and DEA efficiency models. In this research cross sectional data was used. This was because the objective of the research was to provide a snapshot of the technical and scale efficiency of public acute hospitals in Ireland using DEA and incorporating some qualitative measures for the first time. A recommendation of this research would be that a longitudinal study of these hospitals would be undertaken.

Both DEA and SFA have been widely used in health studies. Hollingsworth and Peacock (2008) provide a comprehensive review of 188 published papers covering efficiency measurement applications in healthcare institutions. Grosskopf and Valdmanis (1987) assessed, using DEA, the efficiency of 22 public and 60 private not-for-profit hospitals in California and found efficiency means of 0.96 and 0.94 for public and not-for-profit units respectively. Parkin and Hollingsworth (1997), using DEA, highlighted the dependency of the research results on model specification for 75 acute hospitals in Scotland and found a large amount of difference in efficiency results depending upon specification. Maniadakis and Thanassoulis (2000), using DEA, evaluated the performance of acute hospitals in the UK over the period after the introduction of the internal market in the NHS in 1991. The results indicated that there was productivity regress in the first year after the reforms but progress thereafter. Lee et al. (2008), using DEA, assessed the association between hospital ownership and technical efficiency in a managed care environment. Their results showed that non-profit hospitals were more efficient than for-profit hospitals for all four years examined in the study. Zuckerman et al. (1994) applied SFA to 1,600 US hospitals in 1986/87 found, for pooled data, an inefficiency of 0.132 for teaching, 0.135 for non-teaching, 0.141 for public, 0.144 for proprietary and 0.129 for private not-for-profit hospitals. Folland and Hofler (2001), using SFA, on a sample of 791 US hospitals in 1985 concluded that group mean inefficiencies were robust to variations in methods, and that individual hospital ranks were not highly correlated, however, not-for-profit hospitals were more efficient than for-profit. Li and Rosenman (2001), using SFA, on a panel of 90 US hospitals between 1988 and 1993 found average inefficiency of 33 per cent, with hospitals with a higher case mix index, or more beds, to be less efficient, while for-profit hospitals were more efficient.

Since 1993 acute hospitals in Ireland have utilised a casemix model for estimating relative efficiency. This model takes into account the relative complexities of each medical procedure and uses a weighting method to calculate the average cost of each procedure. Using this method the cost of treating each patient in each hospital is compared with the national average cost and a table of hospital performance is produced annually based on these measures. Funding for Irish public hospitals is partially based on this casemix model whereby resources are re-distributed annually from those hospitals deemed to be relatively inefficient to those hospitals deemed to be relatively efficient. The weakness of this process is that it concentrates solely on efficiency as a measure of performance whilst ignoring both the effectiveness of the organisation and the quality of service that it provides.

Another method of measuring performance is Kanji's Business Excellence Measurement System (KBEMS). This system is grounded on critical success factors that are defined as a limited number of areas in which results if satisfactory will ensure successful performance for the organisation. Two structured models were developed; Kanji's Business Excellence Model (KBEM) (Kanji, 1998) and Kanji's Business Scorecard (KBS) (Kanji & Sa, 2002). The first KBEM is dedicated to the measurement of performance from the internal stakeholder's perspective whereas the latter KBS assesses performance from the external stakeholders' point of view. Internal and external scores are then combined to calculate the final Organisation Performance Excellence Index (OPI). The final OPI which gives an aggregate measure of the excellence of the organisation in managing all of the critical success factors is simply the average between the scores of performance excellence based on the assessment of internal and external stakeholders. The limitation of this model is that it is based only on the measurement of non-financial performance measures. Neither does the model examine organisation activity levels or the efficiency or effectiveness of the organisation. The strength of the model, however, is that it does provide a plausible method for evaluating non-financial performance measures and particularly stakeholders' perception of the quality of the service being provided.

Another method of measuring performance is "The Balanced Scorecard" approach. Since Kaplan and Norton published their first paper on the subject in 1992 this approach has been widely adopted by many organisations. The balanced scorecard model includes both financial and operational measures. It complements the financial measures with operational measures on customer satisfaction, internal processes, and the organisation's innovation and improvement activities. It provides answers to four basic questions:

- How do customers see us? (customer perspective)
- What must we excel at? (internal perspective)
- Can we continue to improve and create value? (innovation and learning perspective)
- How do we look to shareholders? (financial perspective)

While giving senior managers information from four different perspectives, the balanced scorecard minimises information overload by limiting the number of measures used. It forces managers to focus on the handful of measures that are most critical.

Thus the balanced scorecard approach looks at both financial and non-financial measures that give managers *"a fast but comprehensive view of the business" (Kaplan and Norton 1992: 71).* They described the Balanced scorecard *"as the dial and indicators in an aeroplane cockpit" (Kaplan and Norton 1992: 72).*

Whilst many organisations have adopted Kaplan and Norton's four perspectives others have found it necessary to modify some of the perspectives in order to reflect the particular organisation's circumstances. This would be true of healthcare organisations.

The NHS has adopted a Performance Assessment Framework that aims to provide a broader view of performance. This framework is just one part of a wider NHS system of performance measurement and management that seeks, as a common goal, to improve performance. The framework adopts a multi-stakeholder approach reflecting stakeholder's interests across six dimensions:

- Health Improvement
- Fair Access
- Effective Delivery of appropriate health care
- Efficiency
- Patient/ carer experience of the NHS
- Health outcomes of NHS health care

In their paper Amaratunga et al. (2002) discussed both the application of the balanced scorecard concept as a widely used management framework for optimal measurement of organisational framework within NHS facilities directorates and the fundamental points to cover its implementation. They identified this framework as a strategic measurement and management system for facilities management.

Zelman et al. (2003) reviewed the use of the balanced scorecard in health care and reached the following conclusions.
The balanced scorecard:

- is relevant to healthcare, but modification to reflect industry and organisational realities is necessary
- is used by a wide range of healthcare organisations
- has been extended to applications beyond that of strategic management
- has been modified to include perspectives, such as quality of care, outcomes and access and
- has been used by two large-scale efforts across many health care organisations in the health care sector in the USA and Canada which differ namely in the units of analysis, purposes, audiences, methods, data and results.

Whatever productivity model was used it would have to be capable of aggregating efficiency and effectiveness measurements. Pritchard (1990) stated that "a productivity measurement system should produce an overall index of productivity". He described such a model in the productivity measurement and enhancement system (ProMes). This research focused on the Data Envelopment Analysis model. Table 3.1 sets out the strengths and weaknesses of the Data Envelopment model, the Balanced Scorecard model, the Stochastic Frontier model and the Productivity Measurement and Enhancement System and the reasons for selecting DEA as the model to be used.

Table 3.1

Performance measurement models – Strengths and weaknesses

Model	Strengths	Weaknesses	Aspects that I will use in my model and why
Balanced Scorecard	Multi-dimensional in nature Integrates both financial and non-financial performance measures Links performance measures to organisational strategy Gives leaders a fast but comprehensive view of the organisation Snapshot of overall performance that focuses attention on those things critical to success All four perspectives are linked together by the cause-and-effect relationships	Difficult and time consuming to implement in a large organisation Requires top-level support and commitment Too many measures Poor balance between objective and subjective measures Lacks either outcomes or performance drivers of outcomes Leads to game playing or dysfunctional behaviour	Whilst the multi-dimensional nature of this model is attractive the difficulty of implementing it in a large organisation is a major drawback The subjectivity involved when deciding on the balance between each of the measures weakens the model as an overall organisation measurement tool
Data Envelopment Analysis	Can readily incorporate multiple inputs and outputs Calculating technical efficiency only requires information on output and input quantities (not prices) Possible sources of inefficiency can be	Being a deterministic rather than a statistical technique it produces results that are particularly sensitive to measurement error It only measures efficiency relative to best practice within the particular sample	I used DEA in the model. Using DEA allowed me to incorporate multiple inputs and outputs in the model. This was critical in order to provide me with an overall organisational measure of productivity. As well as measuring

	determined as well as efficiency levels Inputs and outputs can have very different units Comparisons are directly against peers By identifying the peers for organisations which are not observed to be efficient, it provides a set of potential role models that an organisation can look to for ways of improving its operations	It does not measure absolute efficiency Its scores are sensitive to input and output specification and the size of the sample Large problems can be computationally intensive	efficiency it also allowed me to determine possible sources of inefficiency. The DEA model reduced subjectivity as it determined the weightings of each activity. The model only required information on output and input quantities and not prices. By identifying peers the model provided potential role models that an organisation could look to for ways of improving its operations.
Stochastic Frontier Analysis	It allows for the separation of the inefficiency effect from statistical noise due to data errors, omitted variables, random unobserved heterogeneity etc. It allows statistical inference on the significance of the variables used in the model, using standard statistical tests.	Vulnerable to errors in the specification of the functional form. It requires the specification of a production, cost, revenue or profit function as well as assumptions about the error terms. The specification of the decomposition of the error terms is imposed a priori.	This model is advantageous when there is likely to be a high degree of measurement error.

ProMes	The ability to provide a single index of productivity as well as sub-indices of the important indicators of productivity A flexible system The ability to aggregate across units Clear statement of organisation objectives Regular feedback to personnel Feedback on performance used to improve productivity Positive motivational properties Establishes contingencies that show the relationship between the amount of an indicator and the effectiveness of that amount Takes account of non-linearity	Dependant on quality of feedback Success dependant on the degree to which units prioritised their actions on the feedback Subjectivity involved when ranking indicators	The ability of this model to provide a single index of productivity makes it attractive. However the subjectivity involved when ranking the indicators seriously weakens it. The success of this model is highly reliant on the quality of feedback, which may not always be of the highest standard. The ability of this model to deal with non-linearity is advantageous.

3.6 Hospital efficiency literature

Both parametric and non-parametric methods have been used for measuring efficiency in many different types of healthcare institutions. These have included hospitals, nursing homes, health districts, primary care programmes and physician practices. While non-parametric methods have long been the main tool of measurement parametric methods have become more popular in recent years. This is due to new methodological developments which would include the ability to accommodate multiple outputs and inputs and the availability of software to facilitate analysis.

Hollingsworth and Peacock (2008) provide a comprehensive review of 188 published studies between 1983 and 2005 covering efficiency measurement applications in healthcare institutions. Over 50% of the reviewed studies are in the hospital sector reflecting its central role in the healthcare sector. Obviously the availability of hospital data also makes it a more attractive area for research purposes.

Hollingsworth (2008) provides a framework on how to conduct a hospital efficiency study. From a supplier's perspective he suggests some initial criteria as a starting point in both macro and micro terms. From a macro perspective he suggests getting end users involved early, having a balanced view from both the health authorities and their staff, and providing the end users with the information that was originally intended. This research fits well with these proposals. Users were involved at the initial stages both in the focus groups and in the groups that completed the questionnaires. A balanced view was sought from those in governance roles as well as staff and the general public. The DEA models provided the information that was sought in relation to the efficiency of the hospitals in the sample.
Hollingsworth sets out micro issues as including:

- Are you asking the right questions?
- What is your underlying economic theory of production?
- Is your model specified correctly?
- Are your data really good enough to answer the questions, particularly your output data?
- Have you any data on quality?
- If you have quality data, how will you weight it relative to quantity data?
- Is your sample inclusive enough, are you comparing like with like?
- If you are happy with your data and models, what techniques will you use?
- Are you undertaking two stages analysis?
- Do you need to generate confidence intervals?

This research again fits well into this framework. The use of focus groups at an early stage ensured that the right questions were asked and through the use of questionnaires the model specifications were clearly identified. The use of data on quality, even though weightings were not used, also added to the usefulness of the models. The DEA technique was selected having analysed other performance measurement techniques.

From a demander perspective Hollingsworth set out in a table a suggested checklist, based on the Drummond et al. (2005) list for assessing economic evaluations, for assessing if an efficiency analysis should be made use of. He states that the two assessment questions asked by Drummond et al. are also pertinent here: is the methodology appropriate and are the results valid? If the answer to this is yes – do the results apply in my setting? While the checklist is a starting point it does provide excellent guidance for assessing efficiency measurement studies. Again the current research fits well with this checklist. The research question is well defined, relevant inputs and outputs were included and accurately measured, quality and quantity of data was clear and comprehensive, the use of DEA was clearly justified, and sensitivity analysis was performed on the models. Finally the presentation and discussion of study results attempted to include all issues of concern to the users.

The focus of this research was on the DEA measurement method.

The major advantage of DEA as a management method is that it can readily incorporate multiple inputs and outputs, and calculate technical and scale efficiencies, whilst only requiring information on input and output quantities and not prices. As well as identifying efficiency levels the method also identifies "peers" for organisations that are not deemed to be efficient and thus provides potential role models for inefficient organisations to look at. While SFA is also widely used as a performance measurement method I felt that its advantages were outweighed by its disadvantages for this research. Jacobs et al. (2006) state that SFA appeals to economic theory but that: *"The theoretical underpinnings of SFA are derived mainly from an extension of the theory of the firm, and the suitability of this theory as a basis for efficiency analysis remains to be established."* (Jacobs et al., 2006: 152)

In particular, SFA is vulnerable to errors in the specification of a production, cost, revenue or profit function as well as its assumptions about error terms even though it is advantageous in situations where there is likely to be a high degree of error measurement. Jacobs et al. (2006) state that***:***

"Advocates of DEA would argue that the problem of having to provide a prior specification for the model can be avoided by applying the non-parametric technique." (Jacobs et al., 2006: 152)

With DEA the frontier is positioned and shaped by the data and not by theoretical considerations. DEA is therefore highly flexible with the frontier moulding itself to the data. This flexibility of functional form is an attractive feature of the technique. However, a drawback would be that the location of the DEA frontier is sensitive to observations that may have unusual types, levels or combinations of inputs or outputs that would have a scarcity of adjacent reference observations or peers. DEA assumes correct model specification while SFA allows for the possibility of modelling or sampling error. If measurement error is thought to be present then SFA may be the more appropriate technique. However, it may be possible to sustain the argument that there is no measurement error. Gannon (2005) refers to Banker et al. (1993) who show that DEA is favoured when measurement error is an unlikely threat and where the assumptions of neoclassical production theory are questionable, while SFA on the other hand deals with severe measurement error and where simple functional forms provide a close match to the properties of the underlying production technology.

DEA generates efficiency scores for each organisation by comparing it only to its peers. Therefore if no peer exists then the organisation is assigned full efficiency. Similarly when assigning an inefficiency score to an organisation lying away from the frontier only its peers are considered, discarding information pertaining to the remainder of the sample. SFA gets over these issues by utilising the full sample information when estimating relative efficiency.

There are other differences that influenced my choice of DEA. As already stated one of the key strengths of DEA over SFA is that it can readily model multiple input-output production processes. SFA is ill suited to the consideration of multiple outputs. Both methods are susceptible to the influence of outliers and small sample sizes, with DEA more susceptible to outliers. In relation to sample size Jacobs et al. (2006) quote Banker et al. (1993):

“Small sample sizes do not prevent the application of DEA, but as with all parametric estimation processes, SFA estimates are likely to be more imprecise the smaller the sample size.” (Jacobs et al., 2006: 155)

Given the relatively small sample in this research the DEA technique may have been more appropriate.

In designing the DEA model a number of critical decisions need to be made. These are the input/output choices, the aggregating of inputs/outputs, input/output orientation, and returns to scale. Each of these choices will now be discussed in turn.

3.7 Input/Output choices

In order to assess the performance of an organisation we need to measure how efficiently it is transforming inputs into outputs. Using the DEA model we can readily incorporate multiple inputs and outputs into the model. The choice of the unit of assessment and the identification of the inputs and outputs are critical to the DEA model.

> *Most reported assessments of performance simply state the input-output variables used rather than detail the process of their identification. Yet the input-output set used can be critical to the views ultimately derived with respect to the relative performance of the units being assessed.*
>
> (Casu et al., 2005: 1364)

The choice of inputs and outputs in the DEA model is critical to its results. Much research in the hospital and health sectors have focussed on ensuring the use of relevant and sufficient inputs and outputs that would capture the production process. The choices to be made may relate, for example, to the inclusion of quantitative or qualitative measures in the model.

It is important to ensure that all variables included in the model are relevant and also that relevant variables are not omitted from the model. It is suggested that the exclusion of relevant variables is likely to be more damaging to frontier models than the inclusion of irrelevant variables (Smith, 1997). Not including some outputs will also disadvantage those organisations that are relatively efficient at producing those outputs. At the same time it is important not to include too many inputs or outputs in the model as this may inflate the efficiency scores and result in some organisations appearing efficient by default.

> *If we do not delineate the unit of assessment properly, or if we omit some important inputs and outputs, the assessment will be biased.*
>
> (Thanassoulis, 2001: 4)

The results from the model may depend on the input/output mix. Dittman et al. (1991) looked at 105 acute units in the USA and found that efficiency was dependant on the input/output mix. Ozcan (1992) also found that efficiency measurements of 40 acute units were dependant on the variables used in the model.

In many cases the researcher does not have access to all preferred inputs and outputs. This is particularly the case in relation to qualitative data. Similarly the researcher may not have a choice in the quality of the input and output data that can be used. For these reasons the findings of some research should be treated with caution.

In measuring hospital efficiency specifying inputs is generally more clear-cut than specifying outputs. Inputs would include labour costs, medical consumable costs, non-medical consumable costs and capital costs. Labour categories include doctors; nurses; allied health professionals; care assistants; catering staff; cleaners, porters and administration staff. Non-labour costs would include medical, non-medical and utilities costs.

Capital is more difficult to measure, mainly due to the difficulty of both measuring existing capital stock and then attributing it to specific time periods. Because of this proxies are generally used for estimating capital. The majority of research studies of the health sector have use “number of beds” as a proxy for capital (Burgess and Wilson 1996, Kerr et al. 1999, Maniadakis et al. 1999, Chang et al. 2004, Kontodimopoulos et al. 2006, Friesner et al. 2008 and Kirigia et al. 2008) even though this would be far from an ideal measure of capital.

Output is a far more problematic variable in the DEA model. Whilst some outputs such as activity levels are readily available and relatively easy to measure others are not. These latter output measures would include qualitative measures. These measures can be broken down into those that assess the success or otherwise of the medical procedure carried out such as mortality rates, morbidity rates or the un-planned patient re-admission rates, and those that relate to the non-medical quality of service such as hospital hygiene, patient satisfaction and hospital acquired infections. One of the main difficulties with using qualitative measures is that the data may not be available. I found this to be the case when carrying out the research. Mortality rates were not publicly available either by hospital or by doctor. Likewise the rate of un-planned re-admission of patients was not available. I did succeed in getting the mortality rate by hospital but only with great difficulty and on a confidential basis. The situation was only slightly better with the non-medical qualitative measures. Hospital hygiene data and hospital acquired infection data were available by hospital but patient satisfaction data was not. It appears to be ridiculous that such an important measure as patient satisfaction was not being recorded on a national basis. Ad hoc patient satisfaction measures were however being carried out in some hospitals.

Non-qualitative output measures are, however, more readily available and are very useful to management and not just from a casemix perspective. Measures such as the average length of stay of a patient per specialty, overall patient throughput, the length of time patients are waiting to be treated, numbers of patients being treated as day cases, the number of patients that have their surgery on the day of admission and the overall length of both inpatient and outpatient waiting lists, are all important output measures that are useful to hospital management in the running of hospitals. One wonders just how robust measurement models are when relevant inputs or outputs are omitted because they are not available and whether the results of such models are biased in some way because of the omitted variables. Output measures that would have added to the models in this research but that were not available were patient satisfaction measures and measures of the numbers of patients who returned unexpectedly to each hospital a short time after having their medical procedure. These qualitative measures would have given a more accurate picture of the quality of service being provided in each of the hospitals.

The process of developing a DEA model may require testing various combinations of inputs and outputs before deciding on the final model. This also allows for the sensitivity of the model to different specifications to be tested. Casu et al. (2005) used a computer-supported group support system with an advisory board to enable the analysts to extract information pertaining to the boundaries of the unit of assessment and the corresponding input-output variables. Their approach ensured a more comprehensive and less biased approach to the choice of inputs and outputs for the DEA model.

3.8 Aggregating inputs/outputs

The next issue is that of aggregating inputs or outputs. The most common aggregates are those of labour inputs. For example all categories of doctor including consultants, registrars, senior house officers and interns are often aggregated to one heading of “doctors”. Similarly, all grades of nurse are generally aggregated to “nurses”. Other labour groups are also aggregated. Valdmanis (1992) used as inputs in her DEA model physicians, nurses and full-time equivalents for other staff. Burgess and Wilson (1996) just used the overall number of personnel as their labour input while Magnussen (1996) again used physicians and nurses as labour inputs. Maniadakis and Thanassoulis (2000) used as labour inputs doctors, nurses and other personnel while Puig-Junoy (2000) used full-time equivalent physicians, full-time equivalent nurses and equivalents and full-time equivalent non-salary personnel. More recently Kontodimopoulos et al. (2006) and Kirigia et al. (2008) used doctors and nurses as inputs whilst Hajialiafzali et al. (2007) used full-time equivalent medical doctors, full-time equivalent nurses and other personnel.

Outputs can be difficult to aggregate because of the wide variation in types of medical procedures. However, casemix systems have been introduced to address this problem. By adjusting outputs to take account of casemix ensures greater comparability between the outputs of each hospital. Interestingly when Grosskopf and Valdmanis (1993) compared the results of DEA models using weighted and unweighted outputs they found there to be no significant differences between the results. However, as they suggested their sample was quite homogeneous which may have influenced the results. With casemix systems a weighting is applied to each procedure where a complex case would carry a higher weighting than a simple case. Casemix systems have allowed greater comparability between outputs and support their aggregation. Dalmau-Matarrodona and Puig-Junoy (1998) used case mix adjusted discharged patients as output in their DEA model. Linna (1998) used diagnosis related groups (DRG) of inpatients and Linna and Hakkinen (1999) used DRG weighted total patient admissions. Chern and Wan (2000) used case mix adjusted patient discharges in their model while Chirikos and Sear (2000) used case-mix weighted patient admissions. Similarly Maniadakis and Thanassoulis (2000) used casemix adjusted outpatient attendances, day cases and inpatient discharges in their model. More recently Gannon (2005) used DRG adjusted inpatients as an output in her model while Lee et al. (2008) used the case mix adjusted number of patients discharged.

3.9 Input/Output orientation

Model orientation is the next choice that needs to be made. Should the model be input or output orientated? The choice of whether an input or output orientation would be used in the model is dependant on the objective of the production units and the constraints under which they operate. It also depends on what inputs/outputs are used and which of these are controllable or exogenously fixed. For example if the numbers treated were to be an output then output orientation would imply maximise numbers treated but that number depends on how many people fall ill, which is exogenously fixed. Alternatively if an organisation such as a hospital has to operate under tight budget constraints then its priority may be to minimise its inputs while producing a given output. In this situation an input orientation may be appropriate. Zere et al. (2006) examined the technical efficiency of district hospitals in Namibia using an input orientated DEA model. This study was driven by the need to alert policy makers to the potential resource gains by ensuring that those hospitals that absorb the majority of health resources are technically efficient. Similarly Maniadakis and Thanassoulis (2000) used an input orientated DEA model to evaluate the performance of acute hospitals in the U.K. over the period after the introduction of the internal market in the National Health Service in 1991. Lee et al. (2008) examined the association between hospital ownership and technical efficiency in a managed care environment using an input orientated DEA model. Likewise if a hospital aims to maximise its outputs while holding its inputs constant then an output orientation may be warranted. Al-Shammari (1999) used an output orientated DEA model when measuring the technical efficiency of hospitals in Jordan. The objective of the research was to identify the relatively efficient hospitals, the relatively inefficient hospitals, the efficiency reference set for the relatively inefficient hospitals, and the alternative actions that would make the relatively inefficient hospitals efficient. Similarly Valdmanis et al. (2004) examined the capacity of public hospitals in Thailand and the production of care for poor and non-poor patients using an output orientated DEA model.

If a hospital is technically inefficient using an input orientation then it will also be technically inefficient using an output orientation even though generally the two technical efficiency scores will differ. The peers for the technically inefficient hospital will also differ depending on whether an input or output orientation is used in the model. In the current difficult economic climate, where cost containment is the main priority, the input orientated model may be the more appropriate option to use. A large amount of demand for hospital services is determined by exogenous factors. These are the effects of the external environment which may include various characteristics of health care organisations, such as differences in ownership; location; the health needs of their patient populations, the local health economy and community and primary care, or institutional constraints such as access to capital resources. Just as a hospital cannot be completely sure what type of patient will arrive at the emergency department, hospitals in lower socio-economic areas may appear less efficient because of the health status of their population. The same issue could arise in an area with a high elderly population with a higher demand on hospital services. Inadequately accounting for the environment in which hospitals operate may lead to seriously faulty conclusions. However, there remains an active and unresolved debate about how to incorporate such environmental variables into DEA (Fried et al. 2002). Dyson et al. (2001) suggested that if exogenous factors are used as variables in a standard DEA model, then it may be possible to either include all the exogenous factors as inputs and use an output oriented DEA model, or include all the exogenous factors on the output side and use an input oriented model. Another suggested approach is to use DEA models that account explicitly for the existence of exogenous and/or constrained factors. While these recommendations have been made there is still no generally accepted method for dealing with the issue.

> *The complexity of these recommendations, and the fierce demands they make on data, are indicative of the complexity of the environmental variable problem. There is no generally accepted method for taking into account environmental variables in DEA models or for testing whether an environmental variable has a significant influence on the production process and the resultant efficiency estimation. For health care, the issue is often likely to be the single biggest source of technical and policy debate, and it must therefore be treated with great caution.*
>
> (Jacobs et al., 2006: 116)

While using input orientation may be appropriate for hospital services studies in certain circumstances, particularly when working under tight budgetary constraints, all hospital studies need not necessarily be input orientated. An output orientation may be used when maximising outputs and keeping inputs constant.

3.10 Economies of scale

The DEA model allows the estimation of whether a production unit has increasing, constant or decreasing economies of scale. Returns to scale describe whether or not a production unit is operating at optimal size. If a hospital, for example, is operating at constant returns to scale then size does not matter. This implies that there are no economies or diseconomies of scale present, and this generally means that doubling all inputs will lead to a doubling of all outputs. This assumption is inappropriate where economies or diseconomies of scale exist. For example if increasing returns to scale exists then a doubling of all inputs should result in a more than doubling of all outputs. Similarly it is inappropriate where diseconomies, decreasing returns to scale, exist and a doubling of all inputs should lead to less than doubling of all outputs. However, in most organisations because of financial or labour constraints, government regulations, social objectives or imperfect competitions they are not operating at optimal scale (Coelli et al., 2005). Imposing constant returns to scale in a model may lead to bias in estimating efficiency which may be more serious than estimating under variable returns to scale, in a situation where constant returns to scale would be more appropriate. The less restrictive variable returns to scale frontier allows the best practice level of outputs to inputs to vary with the size of the organisations in the sample.

Smith (1997) suggested that using an inappropriate return to scale assumption would be more inaccurate when the sample being tested is small.

Many studies have been carried out using DEA to estimate economies of scale. Ferrier and Valdmanis (1996) estimated using DEA the efficiency of 360 rural hospitals in the USA and found the scale efficiency to be 0.893. Mobley and Magnussen (1998) used DEA to compare 178 hospitals from the USA with 50 hospitals from Norway in 1991 and found the scale efficiency to be higher in the Norwegian sample. Dalmau-Matarrodona and Puig-Junoy (1998) estimated using DEA the efficiency of 94 Spanish acute hospitals in 1990 and found that scale efficiency to be influenced by size and severity of illness. Hollingsworth and Parkin (2001) used DEA to estimate the scale efficiency of 49 neonatal care units in the UK in 1990/91 and found varying economies of scale.

3.11 DEA Models

DEA models have been used when investigating many different aspects of hospital efficiency worldwide. They have been used in not just measuring technical and scale efficiency in hospitals and health centres at specific points but also when assessing the impact of structural, governance or managerial changes over time. These would include assessing the pre- and post- merger performance of hospitals, assessing the performance of hospitals pre- and post- the introduction of the internal market in the UK and assessing seasonal efficiency variations of hospitals. DEA models have also been used in comparing the efficiency of public and private hospitals, in assessing whether ownership types influence performance, in assessing whether teaching or non-teaching hospitals are more efficient, in assessing whether hospital locations affect their performance and in estimating performance targets. The use of DEA models has increased and their application has broadened over the last twenty five years. In particular their increased use in sub-Saharan Africa in measuring hospital and health centre efficiencies has been noticeable.
The literature covering the broad application of DEA models over the past twenty five years will now be discussed, assessing the model specifications used and critically evaluating the relevance, reliability and accuracy of each of the models used.

Ownership

DEA models have been used extensively in assessing the efficiency of different ownership types. In the U.S.A. for example many studies have been carried out that compared the efficiency levels of public, private and not-for-profit health institutions.

Grosskopf and Valdmanis (1987) examined 22 public and 60 private not-for-profit hospitals in California and found efficiency means of 0.96 and 0.94 for public and not-for-profit units respectively. Their DEA model was input orientated. The inputs in the model were physicians, full-time-equivalent non-physician labours, admissions and net plant assets. The outputs in their model were acute care, intensive care, surgeries, ambulatory and emergency care. Their results suggested that ownership affected efficiency. They showed that public hospitals and not-for-profit hospitals had different best practice frontiers, and that public hospitals appeared to use relatively fewer resources. The results could reflect differences in quality of care by ownership. The main criticisms that I would have of this model are that there were no qualitative output measures used and it is not clear how each of the specifications were chosen. Having said that, the inputs used in the model adequately covered labour and capital requirements.

Bannick and Ozcan (1995) used DEA to assess differences in performance efficiency among two branches of the federal hospital system. They looked at 284 Federal Units, finding Department of Defence hospitals to be more efficient than Veteran Administration (VA) Units. In their DEA model they included six input measures and two output measures. The input measures used were capital investment in operational beds, service mix intensity, supplies, medical providers, nurses and support staff. The output measures used were inpatient days and outpatient visits. The criticisms that I would have of this model are that there were no qualitative output measures used and that it is not clear how the inputs and outputs were chosen. However, the size of their sample enabled them to cover all input areas of labour, non-labour and capital.

Valdmanis (1992) looked at the efficiency of 41 public and not-for-profit hospitals in Michigan and found using an input orientated DEA model that the efficiency of public hospitals was higher than the not-for-profit hospitals. She also found that alterations in the input-output model brought differences in efficiency levels and ranks. The inputs used in the model were physicians, nurses, full-time-equivalent others, admissions and net plant assets. The outputs were adult, paediatric, elderly, acute and intensive care inpatient beds; number of surgeries; number of emergency and ambulatory visits, and total house staff. Again the absence of a qualitative measure would be my main criticism of this model. The rationale for their choice of inputs and outputs in the model is also not clear.

Ozcan et al. (1996a) examined the efficiency of 85 hospitals and found the not-for-profit units to be more efficient than the for-profit units with efficiency levels of 0.72 and 0.61 respectively.

Burgess and Wilson (1996) examined the efficiency of 2,246 hospitals and found the Veterans' Affairs units to be more efficient at 0.87 than non-federal for-profit and not-for-profit units whose efficiency levels ranged from 0.82 – 0.83. In their study using DEA they analysed the four types of ownership structure in the U.S.A. hospital industry. These were private non-profit, private for-profit, federal, and state and local government. Their sample of 2,246 hospitals was made up of 134 Veterans' Affairs (VA) hospitals, 319 Non-federal hospitals, 254 For-profit hospitals and 1,539 Not-for-profit hospitals. The model was run using both input and output orientations. It had seven input and six output measures. The inputs were the number of acute care hospital beds weighted by scope of service index, the number of long-term beds, registered nurse full-time equivalents, licensed practicing nurse full-time equivalents, other clinical labour full-time equivalents, non-clinical labour full-time equivalents and long term care labour full-time equivalents. The outputs were acute care inpatient days, case mix weighted acute care inpatient discharges, long-term care inpatient days, number of outpatient visits, ambulatory surgical procedures and inpatient surgical procedures. The results show empirical evidence of differences in the technical efficiency across the types of hospital, although the authors were unable to test for the source of these differences. The argument by Hansmann (1980) that third party payment systems may tend to homogenise hospital ownership types in terms of technical efficiency is not supported by the results. To the extent that differences found among the ownership types are due to different incentives and constraints faced by managers across different types, any sensible attempt at health care reform should pay particular attention to incentive effects of new regulation.

The main criticisms that I would have of this model are that no qualitative output measures were used and again it is not clear how the inputs and outputs were chosen. However, labour and capital were adequately covered in the model and utilising casemix adjusted acute patient discharges improved the reliability of the model.

Wei (2006) measured, using a DEA model, the efficiency and productivity change in Taiwan hospitals over the five year period 2000 to 2004. His sample of 110 hospitals included 43 Public, 29 Proprietary and 38 Private. His model had five inputs and three outputs. The inputs were the number of beds, the number of physicians, the number of paramedical personnel, the number of registered nurses and the number of staff. The outputs were the number of patient days, the number of patients for operations and outpatient services. The model was input orientated with constant returns to scale.

The results showed that the average technical efficiency was 0.69 and that 6% of hospitals had achieved an efficient performance. In terms of ownership, the average efficiency was between 0.67 and 0.71, while public hospitals had a higher efficiency with 6.98% being efficient. The average scale efficiency value for all hospitals was 0.92 of which 7.82% of hospitals were efficient. The returns to scale of the hospitals were overly large, and there would appear to be room for downscaling. The analysis showed that from 2003 to 2004 the productivity of all levels of hospitals had significant growth, due to improved technical efficiency. The research also showed that after the first year of implementing the National Health Insurance Global Budget System, the productivity of all hospitals showed deterioration.
The main criticisms that I would have of this model are that the data was not adjusted for casemix, that no qualitative output measures were used and again that it is not clear how the inputs and outputs were chosen. However, labour and capital inputs were adequately covered in the model.

Lee et al. (2008) in their study assessed, using DEA, the association between hospital ownership and technical efficiency in a managed care environment. The model used was input orientated with variable returns to scale. It had four inputs and three outputs. The inputs were service complexity, hospital bed numbers, amount of full-time equivalent labour used and medical supply expenses. The outputs were casemix adjusted medical discharges, number of outpatient visits and the number of full-time equivalent trainees. The data used was from the American Hospital Association Survey Data for acute general hospitals in Florida from 2001 to 2004. The results showed that non-profit hospitals were more efficient than for-profit hospitals for all four years examined in the study. Another finding was that teaching hospitals were more efficient than non-teaching hospitals in 2001 to 2003, but not in 2004. The main criticisms that I would have of this model are that there are no qualitative output measures used and it is not clear how the inputs and outputs were chosen. The model, however, adequately covers labour, non-labour and capital inputs and by using casemix adjusted patient discharges it improves its reliability.

Mobley and Magnussen (1998) used DEA to compare 178 U.S.A. hospitals with 50 Norwegian ones in 1991. Using variable returns to scale the average technical efficiency of the Norwegian hospitals was 0.937, while the figures for the U.S.A. hospitals were 0.884 in for-profit hospitals, 0.936 in not-for-profit hospitals and 0.917 and in non-urban was 0.917. Scale efficiency was also higher in Norwegian hospitals.

These results appeared to indicate a higher level of efficiency in public hospitals than in private ones. The research carried out by Hollingsworth and Peacock (2008) would appear to corroborate this finding. They found that public hospitals had a higher efficiency score at 0.90 than not-for-profit hospitals at 0.832, which were generally private, and the for-profit hospitals at 0.831. Their results also indicated that the sample of European hospitals examined had a higher mean efficiency of 0.876 than the U.S. sample which had a mean efficiency of 0.826.
Hollingsworth and Peacock (2008: 91) stated: *"The results, that public provision seems more efficient and that European hospitals have higher average efficiency would seem contrary to the perception that private market provision of services is more efficient than public provision of services."*

While the majority of studies would appear to indicate that public hospitals are more efficient than private ones, a number of studies have found the opposite to be the case. Chang et al. (2004) examined over 483 hospitals in Taiwan using an output orientated DEA model and found efficiency scores ranging from 0.58 to 0.93, with private hospitals being more efficient. The inputs in their model were patient beds, number of physicians, number of nurses and supporting medical personnel. The outputs were patient days, clinic or outpatient visits and the number of surgical patients. The main criticisms that I would have of this model are that there are no qualitative output measures used, outputs were not adjusted for casemix and it is not clear how the inputs and outputs were chosen.

Similarly Ferrier and Valdmanis (2002) when estimating efficiency scores for a sample of psychiatric hospitals found private not-for-profit provision to be most efficient.

Hospital size

Gruca and Nath (2001) investigated, using DEA, the impact of ownership, size and location on the relative technical efficiency of community hospitals in Ontario, Canada, where a single payer system was in operation. The inputs that they used in their model were fulltime equivalent nurses, fulltime equivalent ancillary staff, fulltime equivalent administration staff, purchased services and supplies, and the total number of staffed beds. The outputs that they used were casemix adjusted inpatients, casemix adjusted outpatients and the total days of long-term care. They examined the efficiency of 168 community hospitals in 1986 and found that there were no significant differences in efficiency across ownership type, size or location.

The detailed results showed that secular hospitals were more efficient than religious ones with efficiency scores of 0.75 and 0.67 respectively and with government hospitals averaging 0.70. They also found rural hospitals to be more efficient than urban with efficiency scores of 0.77 and 0.72 respectively, small hospitals to be more efficient than large with scores of 0.77 and 0.69 and those with long-term beds more efficient than those without with scores of 0.77 and 0.58. Their findings also suggest that model formulation and differences in payer mix across types of hospitals in the U.S. had a strong influence on the measurement hospital ownership – efficiency relationship. The main criticism that I would have with this model is that no qualitative measures were used.

Other researchers focused on the size of units. Chern and Wan (2000) examined 80 non-government hospitals in the U.S.A. using an input orientated DEA model and found that efficiency fell over the two years 1984 and 1993 from 0.80 to 0.76, with medium sized units being more efficient in 1984 and larger units more efficient in 1993. The inputs in the model were beds, service complexity, non-physicians full-time-equivalents and operating expenses. The outputs were casemix adjusted discharges and outpatient visits.

Kerr et al. (1999) examined 23 hospitals in Northern Ireland using an output orientated DEA model and found that larger units appeared to be more efficient. The inputs in the model were nurses, consultants, administration staff, ancillary staff and beds. The outputs were surgical, medical, obstetrics and gynaecology, and accident and emergency patients.

McCallion et al. (1999) examined a similar sample of 23 hospitals in Northern Ireland and found again larger hospitals to be more efficient than smaller ones.

DEA versus SFA

Gannon (2005) measured the technical efficiency of acute hospitals in the Republic of Ireland during the period 1995 to 2000, using DEA and Stochastic Frontier Analysis (SFA), and found that efficiency levels ranged from 0.93 and 0.97. The inputs used were in the form of capital and labour. In terms of capital the average number of beds in the year in each hospital was used and the labour inputs were measured by the number of people employed in each hospital as counted in December each year. The outputs consisted of inpatients, outpatients and day cases. Inpatient and day case figures were adjusted for casemix but unadjusted outpatient data was used. The results of the study showed that when comparing the DEA and SFA methods that there were lower efficiency scores under the SFA method. This suggested that DEA efficiency measures are not controlling for other factors such as the type of production process or other environmental factors that are not included in the model. However, before coming to any definite conclusions it is critical to be sure that both the DEA and SFA models were using the same input-output data and the same assumptions on returns to scale. A number of criticisms could be made of this model. It is not clear how the inputs and outputs in the model were chosen. Are they relevant measures? The choices may be influenced by the bias of the researcher. There were no quality measures included in the model. Therefore are the hospitals' performance adequately reflected in the results? The fact that outpatient data was unadjusted for casemix may have influenced the results.

Model Specification

Parkin and Hollingsworth (1997) using an input orientated DEA model highlighted the dependency of the research results on model specification for 75 acute units in Scotland and found a large amount of difference in efficiency results depending upon specification. They also found efficiency to be as low as 0.63. The inputs used in the model were the average number of staffed beds; the number of trained, learning and other nurses; the number of professional, technical, administration and clerical staff; junior and senior non-nursing, medical and dental staff; the cost of drug supply, and the hospital capital charge. The outputs were medical acute discharges, surgical acute discharges, A&E attendances, outpatient attendances, obstetric and gynaecology discharges, and other specialty discharges.

Impact of the internal market

Maniadakis and Thanassoulis (2000) used DEA to evaluate the performance of acute hospitals in the U.K. over the period after the introduction of the internal market in the National Health Service in 1991. The data set used covers a sample of 75 Scottish hospitals over the financial years 1991/2 to 1995/6 inclusive. The outputs used reflect accident and emergency attendances, outpatient attendances, day patients and inpatient discharges. The outpatient attendances, day cases and inpatient discharges were adjusted for casemix.

The inputs used reflect the number of doctors, nurses and other personnel, the number of hospital beds and the cubic metres of the hospital buildings. The price used for each of the three labour inputs was the mean annual salary for that professional group and as a proxy for the price of beds and hospital volume the capital charge was used per bed and per cubic metre respectively.

The results indicated that there was productivity regress in the first year after the reforms but progress thereafter. Hospitals became more cost efficient over time. An overall finding was that the magnitude of the changes in hospital performance diminished over time and that there were substantial differences between individual hospitals. While some hospitals showed substantial gains in productivity others showed substantial losses. It was suggested that the gains in productivity were not enough to argue that the internal market had made any significant impact on productivity.

All output measures were adjusted for casemix which should have improved the accuracy of the results.

Impact of mergers

Ferrier and Valdmanis (2004) explored whether mergers improve hospital productivity. They used DEA to generate both efficiency and productivity measures to ascertain whether hospital mergers, at least in the short term, result in performance gains. Using data over the period 1996 to 1998 they applied DEA, both pre-merger and post-merger, to a set of hospitals that were merged in 1997 as well as to a matching control group of non-merging hospitals over the same timeframe. The model had five inputs and three outputs. The inputs were staffed beds, the number of fulltime equivalent physicians, fulltime equivalent medical residents, fulltime equivalent registered nurses and fulltime equivalents for other personnel. The outputs were adjusted patient admissions, the total number of surgeries and the number of Emergency Department visits. Technical efficiency was calculated using an input orientated model. The sample comprised 76 hospitals in total. In 1996 there were 38 pre-merged hospitals and 38 matched control hospitals. In 1997 and 1998 there were 19 merged hospitals, 38 matched control hospitals and 19 pseudo merged hospitals (these were formed by combining the matched pairs of the 38 controlled hospitals). The results show that mergers did result in improvement in terms of efficiency and scale measures as compared to the control and pseudo merged hospitals in 1997. However, merged hospitals did not appear to have sustained improvements in productive performance as evidenced in the 1998 cross sectional study. Overall the results showed that merged hospitals did not do better relative to the control and pseudo merged hospitals. Therefore, they were unable to conclude that mergers led to an unambiguous improvement in either efficiency or productivity.

The main criticisms that I would have of this model are that no qualitative output measures were used and again it is not clear how the inputs and outputs were chosen.

Harris et al. (2000) compared the pre-merger and post merger efficiency of 20 U.S. hospitals using an input orientated DEA model and found that the mergers appeared to increase efficiency with efficiency levels increasing from 0.81 to 0.85. The inputs in the model were service mix, size, employees and operational expenses. The outputs were adjusted discharges, and outpatient visits.

Seasonal inefficiencies

Friesner et al. (2008), using DEA, looked for evidence of seasonal inefficiency in 80 hospitals in Washington State. The sample was made up of 33 small rural hospitals, 28 mid-sized urban hospitals and 19 large urban hospitals. The inputs in the model were licensed hospital beds, the number of square feet in the hospital and paid labour hours. The outputs were total outpatient visits, Medicare inpatient days, Medicaid inpatient days, all other inpatient days, casemix indices for each of the three inpatient groups. The model was run under both constant and variable returns to scale.

The results suggest that technical and cost efficiency vary by quarter. Allocative and scale efficiency also vary on a quarterly basis, but only if the data are jointly disaggregated by quarter and another firm-specific factor such as size or operating status. Thus, future research, corporate decisions and government policies designed to improve the efficiency of hospital care need to account for seasonal trends in hospital efficiency. The greatest mean technical efficiency is in Quarter 1, followed by Quarters 2, 4 and 3.

The main criticisms that I would have of this model are that no qualitative output measures were used and again it is not clear how the inputs and outputs were chosen.

Teaching versus non-teaching hospitals

Grosskopf et al. (2001) compared teaching and non-teaching hospitals in terms of their provision of patient services. They compared the frontiers of each type of hospital using a DEA approach which they applied to a sample of 236 teaching and 556 non-teaching hospitals operating in the U.S.A. in 1994. The data used in the study was taken from the 1994 American Hospital Association (AHA) Survey of Hospitals. The inputs used in the model were the number of licensed and staffed beds, the number of physicians, the number of fulltime equivalent interns and residents, the number of fulltime equivalent registered nurses, the number of fulltime equivalent licensed practical nurses, and the number of fulltime equivalents for other labour. The outputs used in the model were the number of inpatients, the number of non-surgical patients treated, the number of inpatient surgeries, the number of outpatient surgeries, the number of Emergency Department visits, and the number of outpatient visits.

The results indicated that only 10% of teaching hospitals in the sample could compete with non-teaching hospitals. Almost 90% of the teaching hospitals did worse than non-teaching best practice even after eliminating inefficiencies relative to their own frontier. These were the hospitals that were at greater risk for takeover or merger, in which case their teaching function could well be eliminated.

Estimating performance targets

Thanassoulis et al. (1995) explored the use of DEA to assess units providing peri-natal care in England and to estimate performance targets for them. The paper proposed a plausible set of inputs and outputs for peri-natal care in which the output set incorporated both activity levels and quality measures. The inputs were fulltime equivalent obstetricians, fulltime equivalent Paediatricians, General Practitioners' fees, fulltime equivalent midwifes, fulltime equivalent nurses and the number of babies at risk. The outputs were the total number of birth episodes performed by the District Health Authority, the total number of deliveries to mothers resident in the District Health Authority, the number of special care consultant episodes, the number of intensive care consultant episodes, satisfied mothers, very satisfied mothers, the number of abortions, and the number of babies at risk surviving.

The results showed how the incorporation of quality measures into a DEA model could lead to implausible weights and hence spurious efficiencies for some units. Thus the use of extended DEA models for assessing efficiency and exploring potential performance targets was illustrated. The model using weight restrictions offered measures of efficiency but the targets it yielded were not consistently in line with desired improvements to input-output levels. The weights based targets DEA model and the DEA model based on user specified ideal input-output levels did not yield measures of relative efficiency. However, they generally yielded good targets. Therefore the choice of model in practice will to some extent depend on whether an efficiency measure or targets for improvement are more desired.

Evaluating hospital efficiencies

One major area of DEA research has been to look at systematic differences in efficiency across hospitals and to identify the factors causing those differences. These studies, which have been carried out across a wide range of countries, have utilised many different specifications in their models.

Magnussen (1996) examined 46 Norwegian non-teaching hospitals and found them to have mean efficiencies ranging from 0.93 to 0.94. The inputs used in this model were physicians and nursing personnel full-time-equivalents, other personnel full-time-equivalents and the number of beds. The output choices were between patients and patient days and between medical/ surgical outputs versus simple/complex outputs as the aggregate criterion. The study examined the sensitivity of hospital efficiency measures to different output specifications and in order to take account of each hospital's casemix DRGs were used. The principal findings of this study were that the distribution of efficiency was found to be unaffected by changes in the specification of hospital output. However the ranking of hospitals and the scale properties of the technology were found to depend on the choice of output specifications.

Linna (1998) examined 43 hospitals in Finland from 1988 to 1994 and found efficiency scores ranging from 0.81 to 0.93. He investigated the development of hospital efficiency and productivity in Finland using DEA and SFA. His DEA model, which was input orientated, used as inputs net operating costs, total number of beds, average wage rate of labour, and the annual price index for local government and health care expenditure. The outputs used were emergency visits, outpatient visits, DRG inpatients, bed days (applied for inpatient episodes exceeding a certain cut off point), residents trained, on the job training nurses and research. The findings from the study indicated that the choice of model used did not affect the results. The results revealed a 3-5% annual average increase in productivity, half of which was due to improvements in cost efficiency and half due to technological change.

Prior (1996) looked at the technical efficiency of 50 general hospitals in Spain and found overall inefficiency of 3%. The inputs in the DEA model were doctors and upper grade staff, nursing assistants and other middle grade staff, and the number of beds. The outputs were discharges, stays (bed days), visits, and activities.

Chang (1998) estimated the efficiency of six public hospitals in Taiwan from 1990 to 1994 using an input orientated DEA model. He found that the average efficiency score ranged from 0.88 to 0.987 and concluded that the occupancy rate had a positive impact on efficiency and that the proportion of retired patients had a negative effect on efficiency. He also concluded that measuring performance in non-profit organisations could not simply look at the efficiency performance measure itself. It should also include the identification and evaluation of relevant operating characteristics because they are all important factors associated with efficiency performance. The inputs in the DEA model were full-time-equivalent physicians, full-time-equivalent nurses and medical supporting personnel, and full-time-equivalent general and administration personnel. The outputs were clinic visits, weighted patient days including general care patient days, acute care and intensive care patient days, and chronic care patient days.

Al-Shammari (1999) sought to measure and evaluate the productive efficiency of a sample of fifteen hospitals in Jordan using a multi-criteria DEA model over the period 1991 to 1993. Three inputs and three outputs were used in the model. The inputs used were the number of hospital beds, the number of physicians and the number of personnel. The outputs used were the number of patient days, the number of minor operations and the number of major operations. The model was output orientated with constant returns to scale.
The results show potential for reductions of 46.8% in bed-days and 79.1% in physicians, while there is no potential to reduce health personnel. There is also potential for increases of 0.015% in patient days, 27.6% in minor operations and 8.8% in major operations.
The main criticisms that I would have of this model are that the data was not adjusted for casemix, that no qualitative output measures were used and again it is not clear how the inputs and outputs were chosen.

Athanassopoulos and Gounaris (2001) examined 98 Greek hospitals using an input orientated DEA model and found overall efficiency had a mean of 0.81, with rural hospitals more efficient than urban and small hospitals generally less efficient. The input in the DEA model was total cost and the outputs were medical patients, surgical patients, medical examinations, and laboratory tests.

Kirigia et al. (2002) measured, using DEA, the technical and scale efficiency of 54 district level public hospitals in Kenya with a view to identifying the inefficient ones and the magnitude of input reductions needed to make them efficient. He found an average technical efficiency score of 0.96 in the hospitals. A second objective of the study was to make the policy implications of the results explicit for policy makers and hospital managers. The input and output data was only available for 54 district level hospitals. They constituted 55% of all district level public hospitals in Kenya. The DEA model included eleven inputs and eight outputs. The inputs were medical

officers/pharmacists/dentists; clinical officers; nurses; administrative staff; technicians/technologists; other staff; subordinate staff; pharmaceuticals; non-pharmaceutical supplies; maintenance of equipment, vehicles and buildings, and food and rations. The outputs were outpatient department casualty visits; special clinic visits; maternity/family planning visits; dental care visits; general medical admissions; paediatric admissions; maternity admissions, and amenity ward admissions. The model used an input orientated approach with constant returns to scale. The model used variable returns to scale when calculating the scale efficiency score.
The results indicated that forty (74%) of the hospitals were technically efficient.
Of the remaining fourteen inefficient hospitals, two had a technical efficiency score between 50% and 60%, two between 61% and 70%, two between 71% and 80%, two between 81% and 90% and six between 91% and 99%. The inefficient hospitals had an average technical efficiency score 84%, implying that on average they could reduce their utilisation of all inputs by about 16% without reducing output.
The results also indicated that 38 (70.5%) of the hospitals were scale efficient. The average scale efficiency score for all of the sample was 90%, implying that there was room to increase total outputs by 10%. The DEA model had indicated that 26% of the hospitals were run inefficiently and in order to become efficient they must either reduce their inputs or increase their outputs. The authors tabulated the input reductions and/or output increases needed to make each individual inefficient public hospital efficient.
The criticisms of this model would include the non-use of either casemix adjusted figures or qualitative measures.

Harrison et al. (2004) in a sample of over 200 U.S. federal hospitals found efficiency ranging from 0.68 to 0.79 using an input orientated DEA model. The inputs in the model were operating expenses, full-time-equivalent staff, services, and beds. The outputs were admissions and outpatient visits.

Chen et al. (2005) examined 89 U.S. hospitals using an input orientated DEA model and found technical efficiency ranging from 0.75 to 0.80 (0.81 to 0.85 using variable returns models). The inputs in the model were the general service cost, routine and special case cost, cumulative capital investment and ancillary service costs. The outputs used were routine care bed days and special care bed days.

Liu and Mills (2005) examined six Chinese hospitals using an output orientated DEA model between 1978 and 1997 and found that efficiency had decreased during that period. The inputs used in the model were doctors, nurses, fixed asset value, hospital beds and supplies value. The outputs used were admissions, outpatient visits and surgical operations.

Zere et al. (2006) examined the technical efficiency of district hospitals in Namibia for the four financial years from 1978/1979 to 2000/2001 using DEA. Their objective was to quantify the level of technical inefficiency in the country's hospitals so as to alert policy makers of the potential resource gains to the health system if the hospitals that absorb the lion's share of the available resources are technically efficient.
All 30 public sector hospitals were included in this study. The inputs used in the model were total recurrent expenditure, number of hospital beds and number of nursing staff. The outputs used were total outpatient visits and inpatient days. The model was input orientated and both variable and constant returns to scale were used.
The results of the study indicated that a substantial degree of pure technical and scale inefficiency existed in the hospitals. Average technical efficiency was less than 75%, less than half of the hospitals were on the technically efficient frontier and increasing returns to scale predominated. The results indicated that if the inefficient hospitals were to operate as efficiently as their peers on the best-practice frontier that the health system could reap efficiency gains amounting to 26-37% of the total resources used in running the hospitals.
There are a number of criticisms that could be made of the model which could limit its reliability and generalisability. These would include the choice of inputs and outputs; the non-use of casemix adjusted figures, the non-use of any quality measure and the use of questionnaires to gather quantitative data.

Akazili, J et al. (2008) measured, using DEA, the extent of technical efficiency in public health centres in Ghana. Their study calculated the technical efficiency of 89 randomly chosen health centres in Ghana. The inputs in their model were the number of non-clinical staff, number of clinical staff, the number of beds and expenditure on drugs and supplies. The outputs were general outpatient visits, number of antenatal care visits, number of deliveries, number of children immunised and number of family planning visits. The model used an input orientated approach with variable returns to scale. The results indicated that 65% of health centres were technically inefficient and 79% were scale inefficient. Inefficient centres had an average scale score of 86%, implying the potential for increasing total outputs by 14% using existing size.

The criticisms of this model would include the non-use of casemix adjusted figures, the non-use of qualitative measures, the fact that the choice of inputs and outputs was based on previous DEA health care studies in the African region and the availability of data.

Masiye, F (2007) investigated the performance of Zambian hospitals using DEA. The objectives of the study were to estimate the productive efficiency of the hospitals, examine the sources of inefficiency and explore policy options for improving performance. A sample size of 32 hospitals was chosen based on their budgets. However two hospitals were removed from the sample due to incomplete data, leaving a sample of 30. The inputs used in the model were total non-labour costs, number of doctors, number of nursing and other clinical staff, and the number of non-clinical staff. The outputs used the number of visits, the number of beds, the number of deliveries, and the number of tests or operations performed. The model was input orientated with variable returns to scale.
The results show that 11 (40%) of the hospitals are efficient and that the average relative efficiency level is 67%, indicating that collectively the hospitals could produce their current output levels while reducing their inputs by 33%. The average technical efficiency score for the 18 inefficient hospitals is 42%, again indicating great potential to reduce costs. Scale efficiency results show that only four (13%) hospitals were operating at optimal plant size, even though many others were very close to their optimal size. Hospital size is a major source of inefficiency in Zambian hospitals. The results also indicated that input congestion was also a source of hospital inefficiency. There are a number of criticisms of this model. The small sample size limits the generalisability of the results. The model does not use either casemix adjusted figures or qualitative measures.

Hatam et al. (2010) applied DEA to measure the technical, scale and economic efficiency of the general public hospitals in Fars Province in Iran. The twenty one general public hospitals affiliated to Shiraz University of Medical Sciences in Fars Province were the sample used in this study. The hospitals were measured and compared over each half year in 2005 and 2006. The inputs were the number of fixed hospital beds; the number of full-time-equivalent physicians, and the number of full-time-equivalent nurses and other personnel. The outputs were the bed occupancy rate; patient day admissions; occupied bed days; average length of stay, and rate of bed turnover.

The results show that 15 (71.4%) hospitals were technically efficient in the first half of 2005 and that 14 (66.7%) hospitals were technically efficient in each of the next three half years. On average over the four periods 67.8% of hospitals were technically efficient and the average technical efficiency score was 93%.
The results show that 7 (33.3%) hospitals were scale efficient in the first half of 2005 and 5 (23.8%) were scale efficient in the second half year. In each the half years in 2006 4 (19%) hospitals were scale efficient. On average over the four periods 23.8% of hospitals were scale efficient and the average efficiency score was 67%. Four hospitals were recognised as being economically efficient in each of the four time periods.
Criticisms of the model would be that the output data was not adjusted for casemix; that there were no qualitative measures included, and that the relevance or otherwise of the inputs and outputs used were not discussed.

3.12 Conclusion

In summary, DEA has been used widely in measuring hospital efficiency. As well as measuring the efficiency of hospitals within a particular country or region in a particular year or over a number of time periods DEA models have been used in addressing a variety of objectives. These include when comparing the efficiency levels of different ownership types (Bannick and Ozcan, 1995; Gruca and Nath, 2001; Chang et al., 2004; Wei, 2006, and Lee et al., 2008); different size hospitals (Kerr et al., 1999; McCallion et al. 1999; Chern and Wan, 2000, and Gruca and Nath, 2001), profit versus non-profit institutions (Grosskopf and Valdmanis, 1987; Valdmanis, 1992; Burgess and Wilson 1996; Ozcan et al., 1996a; Hollingsworth and Peacock, 2008, and Lee et al., 2008), pre- and post- merger situations (Harris et al. 2000, and Ferrier and Valdmanis, 2004), hospital locations (Gruca and Nath, 2001), different measurement methods (Gannon, 2005), pre- and post- introduction of the internal market in the NHS (Maniadakis and Thanassoulis, 2000), seasonal efficiency variations (Friesner et al., 2008), teaching versus non-teaching hospitals (Grosskopf et al., 2001, and Lee et al., 2008), inter country efficiency variations (Mobley and Magnussen, 1998), and estimating performance targets (Thanassoulis, 1995).
Whilst DEA has been a popular measurement model in the USA and Europe for many years it is now being used more widely in other parts of the world. In particular it is noticeable that it has become more popular in sub-Saharan Africa. It has been used in Namibia (Zere et al., 2006), Kenya (Kirigia et al., 2002), Ghana (Akazili, 2008), and Zambia (Masiye, 2007) in measuring the efficiency levels of hospitals in

these countries. As well as measuring efficiency DEA has allowed the examination of sources of inefficiency and the exploration of policy options to improve hospital performance in these countries.

It is clear from current literature that the application of DEA as a measurement model of hospital efficiency has developed considerably over the past twenty five years. However, there are still a number of areas that require further study. Little research has been done on the methodology for choosing the inputs and outputs in the DEA model and this needs further investigation. Likewise the infrequent use of quality measures in DEA models also needs to be addressed. Finally, there has been very little research using DEA models to measure hospital efficiency in the Republic of Ireland. This research aims to address these issues.

In this chapter the various stakeholders and their different perspectives on what constitutes “good” performance in a hospital were discussed. Why a hospital does not perform to its optimal ability and how the level of performance is measured were then considered. A number of papers in which different methods for measuring performance were reviewed and the advantages and disadvantages of each method were discussed. The chapter concluded by addressing the choices that need to be made when developing a DEA model and discussed the current literature on research using DEA models. The main features that the literature has highlighted as being important are:

- The need to make adjustments for casemix in the models. By adjusting outputs to take account of casemix ensures greater comparability between the outputs of each hospital.
- The ability to aggregate inputs or outputs in the DEA models. The most common aggregates are labour inputs. Outputs can be more difficult to aggregate because of the wide variation in types of medical procedures but casemix systems have been introduced to address this problem.
- The choice of inputs and outputs in the DEA model is critical to its results. Relevant and sufficient inputs and outputs that would capture the production process need to be used in the model.
- The choice of whether an input or output orientation should be used in the model is dependent of the production units and the constraints under which they operate, as well as what inputs and outputs are used and which of these are controllable or exogenously fixed.

- The choice of whether to use constant or variable returns to scale in the DEA model. If a hospital is operating at constant returns to scale then size does not matter. The less restrictive variable returns to scale frontier allows the best practice level of outputs and inputs to vary with the size of the organisations in the sample.

In the next chapter the proposed methodology for carrying out this research will be set out.

Chapter 4

Methods

4.1 Introduction

This chapter discusses the Data Envelopment Analysis technique, sets out its history and explains the mathematical basis behind the model. The methodology used in this research to identify the most relevant input and output measures to be included in the model is then discussed and the results from the focus groups and questionnaires are analysed. The chapter concludes with the specification of the performance measurement models that were used in the research.

4.2 Data Envelopment Analysis (DEA)

Data envelopment analysis is a non-parametric method that uses linear programming techniques to derive estimates of efficiency. DEA was initiated by Charnes, Cooper and Rhodes in 1978 in their seminal paper Charnes et al. (1978) and is based on relative efficiency concepts proposed by Farrell (1957). The technique was developed further by authors such as Fare et al. (1983) and Banker et al. (1984). The DEA model, introduced in Charnes et al. (1978), utilises a sequence of linear programmes, one for each decision making unit (DMU), to construct a piecewise linear production frontier, and to compute an efficiency index relative to the frontier. This original CCR model which assumed constant returns to scale was modified by Banker et al. (1984) for variable returns to scale in their BCC model.

Using linear programming DEA calculates the efficiency of an organisation within a group relative to observed best practice within that group. The most common efficiency concept is technical efficiency. This is the conversion of physical inputs into outputs relative to best practice. An organisation operating at best practice is said to be 100% technically efficient. If however an organisation is operating below best practice it is technically inefficient and its technical efficiency is expressed as a percentage of best practice. Allocative efficiency refers to whether inputs for a given level of output and set of input prices, are chosen to minimise the cost of production, assuming that the organisation being examined is already fully technically efficient. Cost efficiency refers to the combination of technical and allocative efficiency. An organisation will only be cost efficient if it is both technically and allocatively efficient.

As well as calculating efficiency scores DEA also identifies “peers” or role models for inefficient organisations. Using a combination or weighted average of the organisation’s peers a hypothetical best practice organisation is derived that can provide targets for the inefficient organisation.

DEA assesses efficiency in two stages. First, a frontier is identified based on either those organisations using the lowest input mix to produce their outputs or those achieving the highest output mix given their inputs. Second, each organisation is assigned an efficiency score by comparing its output/input ratio to that of efficient organisations that form a piecewise linear envelope of surfaces in multidimensional space. If there are M inputs and R outputs, then the production frontier becomes a surface in (M+R) dimensional space. The efficiency of a DMU is the distance it lies from this surface (Jacobs et al. 2006).
Efficiency in DEA is defined as the ratio of the weighted sum of outputs of a DMU to its weighted sum of inputs (Hollingsworth and Parkin, 1998; Smith, 1998). The following model introduced by Charnes et al. (1978) assumes constant returns to scale. The technical efficiency is computed by solving for each DMU the following mathematical programme. Given n outputs and m inputs, efficiency (h_0) for hospital 0 is defined as follows (Jacobs, 2000).

Maximise: $$h_0 = \frac{\sum_{r=1}^{p} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}}$$

subject to:

$$\frac{\sum_{r=1}^{p} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \qquad j = 1, \ldots\ldots, n$$

where:

y_{r0} = quantity of output r for hospital 0

u_r = weight attached to output r,u_r > e, r = 1,……,p

x_{i0} = quantity of input i for hospital 0

v_i = weight attached to input i, v_i > e, i = 1,……,m

e is a non-Archimedean infinitesimal.

This mathematical programme seeks out for hospital 0 the set of output weights u_r and input weights v_i that maximises the efficiency of hospital 0 subject to the important constraint that when they are applied to all other hospitals that none can have an efficiency score that is greater than 1.
The weights are specific to each unit. Clearly the model implies that $0 \leq h_0 \leq 1$ and a value of unity implies complete technical efficiency relative to the sample of units under scrutiny. Since the weights are not known a priori, they are calculated from the efficiency frontier by comparing a particular hospital with other ones producing similar outputs and using similar inputs, known as the hospital's peers. DEA computes all possible sets of weights which satisfy all constraints and chooses those which give the most favourable view of the hospital, that is, the highest efficiency score (Jacobs, 2000).

This can be stated as a mathematical linear programming problem by constraining either the numerator or the denominator of the efficiency ratio to be equal to one. The problem then becomes one of either maximising weighted output with weighted input equal to one or minimising weighted input with weighted output equal to one (Parkin and Hollingsworth, 1997).

The input minimising programme for hospital 0 in a sample of n hospitals is:

Minimise: $h_0 = Z$

subject to:

$\sum_{j=1}^{n} x_{ij} \lambda_j \leq x_{i0} Z$ j = 1,……,n

$\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{r0}$ j = 1,……,n

$\lambda_j \geq 0$, j = 1,……,n

where:

x_{i0} denotes the observed amount of input i for hospital 0.

y_{r0} denotes the observed amount of output r for hospital 0.

λ_j are weights applied across the n hospitals. When the nth linear programme is solved, these weights allow the most efficient method of producing hospital n's outputs to be determined.

Z is the efficiency score.

It is important to note that the two models quoted are dual to each other and yield the same value for the efficiency of a given model 0.

The model is solved giving each hospital an efficiency score. The model computes the factor Z needed to reduce the input of hospital 0 to a frontier formed by its peers, or convex combinations of them, which produces no less output than hospital 0 and uses a fraction Z of input of hospital 0. The hospital will be efficient if Z equals one and slacks are zero. It is important to note that the optimal solution can include what are termed slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier.

If Z is smaller than one, the hospital will be inefficient. The composite unit provides targets for the inefficient unit and Z represents the maximum inputs a hospital should be using to attain at least its current output (Hollingsworth and Parkin, 1998).

The second model above determines a production frontier. The hospitals that lie on this frontier will have an efficiency score of one.

The DEA model can be run with either constant or variable returns to scale. Returns to scale describe whether or not a production unit is operating at optimal size. If a hospital, for example, is operating at constant returns to scale then size does not matter. This generally implies that doubling all inputs will lead to a doubling of all outputs. This assumption is inappropriate where increasing returns to scale exist and a doubling of all inputs should result in a more than doubling of all outputs. Similarly it is inappropriate where decreasing returns to scale exist and a doubling of all inputs should lead to less than doubling of all outputs. The CCR model as set out above is consistent with a constant return to scale production frontier. Banker et al. (1984) extended this model to accommodate a more flexible return to scale model which may be more appropriate when not all DMUs can be considered to be operating at optimal scale. The model is given a further constraint in order to calculate the variable returns to scale frontier:

$$\sum_{j=1}^{n} \lambda_j = 1$$

The variable returns to scale approach produces technical efficiency scores which are greater than or equal to those obtained using constant returns to scale and is therefore probably the more flexible assumption of the underlying production technology (Coelli, 1996a).

4.3 Methodology

The measurement of efficiency in healthcare is a complex exercise. Carter (1991) stated that the NHS had just about all the factors liable to make the definition and measurement of its performance difficult. He stated:

> *It is characterised by heterogeneity, complexity and uncertainty. That it is a multi-product organisation, which has to mobilise a large cast with a high degree of interdependence between the different actors and where the relationships between the activity and impact is often uncertain. It is not always clear who "owns" the performance; the activities of the NHS are only one of many factors influencing the health of the population.*
>
> (Carter, 1991: 96)

Both efficiency and effectiveness measures are included in the performance measurement model. Both of these measures have disadvantages and can be dysfunctional when used alone. *"We speak of relative efficiency because its* *measurement by DEA is with reference to some set of units we are comparing with each other"* (Thanassoulis, 2001: 21).

Effectiveness is a broad concept that includes factors such as standards; objectives of the organisation, expectations of stakeholders and viability of the organisation relative to its competition. Quality can also be readily included in effectiveness measurements. Effectiveness measures in hospitals include achieving activity targets, reducing the average length of stay (ALOS) for patients, increasing day case procedures, minimising and eliminating waiting lists for inpatients, outpatients and day cases, and reducing waiting times for patients in Emergency Departments. These measures also impact on quality of service within a hospital. Maintaining or increasing activity levels will ensure that a greater number of patients are treated, thus improving access for patients and improving their quality of life; reducing the average lengthy of stay of patients while at the same time ensuring that the rate of unplanned re-admissions does not increase should increase patient satisfaction levels; providing surgical procedures as day cases obviates the necessity for a patient to be admitted overnight; reducing waiting lists and thus improving access to services will improve the chance of a better health outcome for the patient, and reducing waiting times in the Emergency Department also improves the chance of a better health outcome for the patient as well as increasing patient satisfaction levels. Other qualitative measures that are publicly available for all hospitals include hygiene measures and measures of hospital acquired infection. Mortality figures for each hospital are also recorded but are not available to the general public. Other measures such as the level of complaints and un-planned re-admission rates are not available. The mortality rate for each surgical procedure either by hospital or by surgeon is not available. These measures should be publicly available and subject to scrutiny if we are to ensure that a quality service is delivered. When comparing the DEA results for the best performing hospitals in this research with their measures of effectiveness, the results are varied. For example, the average length of stay varies between 4.5 and 12.6 for the top six best performing hospitals. This measure is hugely dependent on the specialties within each hospital and their casemix. Day case activity measures are showing an increase in five of the six hospitals on the previous year. However, most hospitals in the sample are showing this increase. Similarly the ED waiting times are no better in the best performing hospitals in the DEA models. Therefore, the DEA measures of relative efficiency are not reflected in better effectiveness or qualitative results in the hospitals.

The assumption in this research is that the organisational model is the goal orientated model, given that the organisation is run by rational decision-makers with a manageable set of goals. Organisational effectiveness can be thought of as the degree to which these goals are met.

The methodology used in this research involved a number of steps. The first step was to measure the key inputs and outputs for all of the acute hospitals in Ireland in 2007, for analysis. It was crucial that all important functions were measured and that the most relevant input and output measures were used in the model. To do this four focus groups of relevant stakeholders were established, a sample of the population of stakeholders were then circulated with a questionnaire asking them to indicate what they believed the most relevant input and output measures should be and finally an academic expert group was established to validate the model. The process was also complemented through the use of audio-taping at the focus group meetings. These measures removed any personal bias from the choice of input and output measures and thus supported the validity of the research. Having decided on the input and output measures the DEA model was run using PIM DEAsoft- V1 software.

4.4 Key Stakeholders

Before deciding on any measurement method it was important to identify the key stakeholders in the system. In the case of a public acute hospital in Ireland these would include:

Minister for Health and Children
Hospital Board of Directors
Patients
Staff
Hospital management

Whilst the above stakeholders may support the overall objectives of a hospital in being efficient, effective and providing a high quality service they may each have different priorities within these objectives. The Department of Health and Children sets out its vision for the Irish health system as:

> *A health system that supports and empowers you, your family and community to achieve your full health potential.*
> *A health system that is there when you need it, that is fair, and that you can trust.*
> *A health system that encourages you to have your say, listens to you, and ensures that your views are taken into account.*
>
> (Department of Health and Children 2001: 10)

The four principles that guide the Irish health strategy are equity, people-centredness, quality and accountability. The health strategy has four goals:

- better health for everyone
- fair access
- responsive and appropriate care delivery
- high performance

Better planning and evaluation models must demonstrate that available resources are used as efficiently and effectively as possible (Department of Health and Children 2001).

The Department of Health and Children may regard a hospital as performing well if waiting lists and waiting times are minimised for patients and that a quality service is provided efficiently and effectively. Good results ensure less political pressure on the Minister and his officials. The hospital Board of Directors may have a more intense focus on ensuring a hospital's survival. They may thus have more interest in financial issues and efficiencies as well as ensuring that a high quality of service is provided. Patients may be interested in ensuring that high quality medical care is available to them as speedily as possible and in a suitable environment. Hospital staff may be interested in protecting their job security, ensuring a safe working environment and providing a quality service. Hospital management may be interested in ensuring that the organisation remains financially viable; operates efficiently; is effective, provides a high quality of patient care and a safe work environment for their staff. The measurement method that was used in this research was that which met the requirements of hospital management. The very broad requirements of hospital management encompass all of the other stakeholders' needs.

4.5 Research Population

All of the population of public acute hospitals in Ireland were surveyed for this research. Whilst the total number of public acute hospitals in Ireland in 2007 was fifty, in order to ensure greater comparability between hospitals the following hospitals were eliminated from the sample:

(a) Those that did not have an Accident and Emergency Department
(b) Those that operated only as a single specialty hospital such as orthopaedic and maternity and
(c) Those that operated as paediatric hospitals

This reduced the sample size to twenty eight hospitals and eliminated many of the smaller hospitals and potential outliers. This is a small sample and as DEA scores are sensitive to sample sizes as well as the number of inputs and outputs in the model it should be borne in mind that a high level of efficiency could be as a result of using too many variables compared to the number of hospitals in the sample. The more variables used the less discriminating the model becomes. The larger the number of input and output variables used in relation to the number of hospitals in the model, the more hospitals will be assigned as fully efficient and hence the less discriminating the DEA model will be. Banker et al. (1989) suggest as a rule of thumb that the number of DMUs should be at least three times the number of inputs plus outputs in any DEA application, although there is no analytic support for this rule (Pedraja-Chaparro et al., 1999). The Central Limit Theorem states tells us that a sampling distribution always has significantly less variability, as measured by standard deviation, than the population it’s drawn from. Additionally, the sampling distribution will look more and more like normal distribution as the sample size is increased, even when the population is not normally distributed. If data follows a normal distribution we can be more confident that we can predict how data will behave. As a general rule a sample size of 30 or more is considered to be large enough for the Central Limit Theorem to take effect.

The hospitals in this sample ranged in size from 118 beds to 842 beds. Eight of these hospitals would be regarded as providing a tertiary service whilst the remainder would be regarded as providing a secondary care service. The reality is however that the difference in the level of complexity of services being provided by each type of hospital is less significant and so it can be difficult to distinguish between both types of hospital. One can find so called tertiary services being provided by secondary care hospitals and at the same time find secondary services being provided by tertiary care hospitals.

In order to address the issue of casemix differences between the hospitals in the sample the inpatient and day case data has been adjusted for casemix complexity. Casemix data was not available for outpatient attendances. McKillop et al. (1990) and Gregan and Bruce (1997) argued that the complexity of the casemix measured for example by employing diagnosis related groups (DRGs) should be taken into account when measuring hospital efficiency. In this model casemix adjusted inpatient and day-case figures as provided by the Department of Health and Children in their Specialty Costing Model for 2007 were utilised.

4.6 Validity and reliability

The research was largely based around observable data that could be empirically recorded. All of the empirical results could be verified. The model of performance measurement was based on reality and was scientifically based. The input and output data was based on actual results that could be verified. For the measurements to be valid they had to measure the right things and be free from bias. In order to overcome bias in choosing the inputs and outputs in the model four focus groups were established to recommend what inputs and outputs should be included in the performance measurement model. A questionnaire was then circulated to a sample population of stakeholders which allowed them to indicate the most relevant inputs and outputs that they deemed should be included in the model.

The reliability of the measurements relates to the consistency of the research findings. This will depend on how the research is executed for data gathering, recording and interpretation. With quantitative research reliability is about reducing random error in the statistical processes whilst with qualitative research it is more about ensuring that the protocol that governs the research is clear and consistent.

4.7 Validating the model

In order to validate the performance measurement model an expert group was established comprising of six academic experts from Aston University in Birmingham. A questionnaire was circulated to this group asking them to rate each of the associated measures of performance in terms of whether they were relevant, informative and/or necessary as a contributor to the overall model of performance. Their ratings were used to provide evidence in support of the model. This approach was adopted by Casu et al. (2005) when identifying the unit of assessment and the corresponding input-output variables for a DEA model. They organised a workshop to explore the knowledge of a group of experts comprising of senior academics and administrators from a number of different English Universities in order to more effectively identify the widest possible range of input-output variables in their model.

4.8 Ethics

It was important to be cognisant of ethical issues when carrying out the research. Reeves and Harper (1981) consider that there are four minimum requirements for any code of practice governing survey research within an organisation:

1. The researcher should consult with all interested parties before undertaking fieldwork and should proceed only by consent and agreement. This will probably require free access to employee representatives including representatives of the trade unions.
2. Agreement needs to be reached with all interested parties as early as possible over the dissemination of results before too great an investment of time is made in an inquiry which will lead nowhere.
3. The purposes of an employee survey and most types of survey research should not be concealed, as this prevents any judgement by respondents as to whether their participation may adversely affect them.
4. Any special circumstance that might affect the interpretation of the results should be clearly reported.

I carried out my research in accordance with these requirements. Approval from the Aston Business School Research Ethics Committee was also received for this research.

4.9 Specification of inputs and outputs

In order to ensure that the most relevant input and output measures were used in the model the process of specifying them was carried out in three steps. The first step was to establish four focus groups, the second step was to circulate questionnaires to the main stakeholders and the third step was to set up an academic expert group to validate the input and output measures that were to be included in the model.

Step 1

The first step was to establish four stakeholder focus groups and to organise a workshop for each one. This process was adopted by Casu et al. (2005) when they used a group support system to aid input-output identification in DEA. In their paper they stated that:

> *the rationale for using a group of experts is that groups have the advantage over individuals in having access to a wider range of expertise, and individual group members have the potential of being stimulated to consider additional aspects of the problem by other members of the group.*
>
> *Casu et al. (2005: 1364)*

The groups were drawn from the following:

Group 1 – Hospital Chief Executives

Group 2 – Management Staff of South Infirmary- Victoria University Hospital (SIVUH)

Group 3 – Board of Directors of South Infirmary – Victoria University Hospital

Group 4 – South Infirmary- Victoria University Hospital Service Users

The focus groups met between the 4th March and the 10th June, 2008. Each group discussed what inputs and outputs should be included in any model measuring hospital performance. In order to add rigour to this process audio-taping was used at these meetings.

The results of the four workshops are set out in the following tables where an "x" indicates a measure that was deemed relevant by the various groups:

Table 4.1

Results of Focus Groups - Inputs

Measures:	Focus Groups			
	1	**2**	**3**	**4**
Skill mix	*			
Nurses/Bed	*			
Radiographers /Bed	*			
Support staff	*			
Capital equipment	*			
Drug costs	*			
Staffing ratios		*		
Patient dependency		*		
Patient complexity		*		
Number of consultants		*	*	

Table 4.2

Results of Focus Groups - Outputs

Measure:	Focus Groups			
	1	2	3	4
Hygiene	*	*		*
Accreditation	*		*	*
Catering	*			
Waiting times	*	*	*	
Day cases	*			
Risk	*		*	
Health and Safety	*			
Waiting lists	*		*	
Inappropriate referrals	*			
Benchmarking	*		*	
Timeliness	*			
Length of stay	*			
Case mix	*			
Patient satisfaction	*	*	*	*
Car parking	*			*
Nursing care	*			
Waiting times OPD	*			
Patient pathway	*	*		
Theatre audit		*		
Infection rate		*		
Patient throughput		*		
Communications				*
Unreported x-rays		*		
Staff friendliness				*

Audits on nausea		*		
Staff courtesy				*
Accessibility		*		
New outpatient attendances		*		
Unplanned patient re-admissions		*		
Mortality rates				*
Morbidity rates		*		
Time taken to answer phone		*		
Adherence to budget		*		
Case mix adjustment		*		
A&E waiting times		*		
Time waiting for x-ray reports		*		
Time waiting for laboratory reports		*		
National treatment purchase fund referral rate		*		
Clinical pathway		*		
Access to beds		*		
Return outpatient attendances		*		

The only input mentioned more than once by the focus groups was the number of consultants and this was only mentioned by two of the groups. The hospital service users' group did not mention any input measure. It was quite clear that they were entirely focused on hospital outputs.

In contrast the number of output measures mentioned was large and varied. Patient satisfaction was the only output measure mentioned by all four focus groups while hygiene, accreditation and waiting times were mentioned by three groups. Risk, waiting lists, benchmarking, patient pathway and mortality rates were mentioned by two focus groups.

Step 2

The second step in the process was to design a questionnaire based on the results of the focus group discussions and to circulate this to four groups of stakeholders. This questionnaire is included in appendix 2. The objective of the questionnaire was to ascertain the opinions of different stakeholders as to what they perceived to be relevant performance measures. The questionnaire was a simple self completion form in which each stakeholder was given a list of hospital performance measures, which were deemed to be relevant at the focus groups workshops, and to tick those that they agreed were relevant. The questionnaires were accompanied by a letter from me explaining the purpose of the survey, how it was proposed to use the information gathered, how the particular respondent came to be selected and why it was important that they should take part in the survey. This letter is included in appendix 3. The questionnaires were sent to the following groups:
Group 1 - Former patients of the SIVUH.
Group 2 – SIVUH staff
Group 3 – Senior Health Officials
Group 4 – Members of Hospital Boards of Directors from other hospitals

These groups were chosen because it was felt that they provided a good cross-section of the general public, which was important given that there are times when patients may have a different opinion with other members of the public when it comes to health choices. The quality adjusted life year (QALY) has been created as a measure to combine the quantity and quality of life. QALYs can provide an indication of the benefits to be gained from a variety of medical procedures in terms of quality of life and survival for the patient. The use of QALYs in resource allocation decisions means that choices between medical groups competing for medical care are made explicit. By using QALYs there is an implication that some patients will be refused or not offered treatment for the sake of other patients. However, these choices have to be made and the patients concerned may have a different opinion to members of the general public. The National Institute for Health and Clinical Excellence (NICE) defines the QALY as a *“measure of a person’s length of life weighted by a valuation of their health-related quality of life over that period”.* (National Institute for Health and Clinical Excellence, 2008: 38)

Group 1

Four hundred questionnaires were issued to former patients. These were patients, whose age ranged nine to ninety five, who were discharged from the SIVUH during August 2008. One hundred and twenty four of these questionnaires were returned, i.e. a 31% return rate.

Group 2

Two hundred and forty questionnaires were sent to staff of the SIVUH across all disciplines. Eighty one of these questionnaires were returned, i.e. a 33.75% return rate.

Group 3

Fifty questionnaires were sent to senior Health Service officials. Eleven of these questionnaires were returned, i.e. a 22% return rate.

Group 4

Fifty one questionnaires were sent to members of Boards of Directors of four Voluntary hospitals, not including the SIVUH. Nineteen of these questionnaires were returned, i.e. a 37.25% return rate.

4.10 Questionnaire Results

The results of each questionnaire are set out in tables 4.3 and 4.4. Table 4.3 sets out the input results and table 4.4 sets out the output results in order of priority for each of the groups. The results from each of the groups are then combined in tables 4.5 and 4.6.

Table 4.3

Inputs deemed relevant by each group in order of priority and percentages received

Group/ Priority	Former Patients	%	Staff of the SIVUH	%	Senior Health Service Officials	%	Hospital Directors	%
1	Number of nurses	92.7	Modern equipment	76.5	Pay costs	90.9	Number of doctors	73.7
2	Number of doctors	90.3	Total number of staff	69.1	Non-pay costs	90.9	Number of nurses	68.4
3	Modern equipm-ent	88.7	Number of doctors	64.2	Total costs	72.7	Number of beds	68.4
4	Number of beds	83.1	Number of beds	64.2	Number of beds	72.7	Modern equipment	68.4
5	Number of support staff	72.6	Number of nurses	63.0	Total number of staff	63.6	Total number of staff	63.2
6	Number of radiogra-phers	63.7	Total costs	63.0	Number of doctors	63.6	Non-pay costs	57.9
7	Total number of staff	52.4	Number of support staff	53.1	Number of nurses	54.5	Pay costs	47.4
8	Total costs	35.5	Number of radiogra-phers	45.7	Drug costs	54.5	Total costs	47.4
9	Drug costs	29.0	Pay costs	44.4	Number of support staff	45.5	Number of support staff	42.1
10	Pay costs	26.6	Non-pay costs	39.5	Number of radiographers	36.4	Number of radiogra-phers	36.8
11	Non-pay costs	23.4	Drug costs	38.2	Modern equipment	36.4	Drug costs	36.8

Amongst former patients the input deemed to be most relevant was the number of nurses at 92.7%, closely followed by the number of doctors at 90.3%, modern equipment at 88.7% and the number of beds at 83.1%. Other inputs that would have achieved a 50% or higher score would have been support staff at 72.6%, number of radiographers at 63.7% and total number of staff at 52.4%. The input deemed to be least relevant was non-pay costs at 23.4%.

Amongst the staff of the SIVUH the input deemed most relevant was modern equipment at 76.5%. This was followed by the total number of staff at 69.1%, number of doctors at 64.2%, number of beds at 64.2%, number of nurses at 63.0%, total costs at 63.0% and number of support staff at 53.1%. All other inputs scored less than 50%. The input deemed to be least relevant was drug costs at 38.2%. Interestingly the range of results at between 38.2% and 76.5% was much narrower than the range of results obtained from the patients' questionnaires, which ranged from 23.4% to 92.7%. This would appear to indicate more of a consensus between staff as to relevant inputs.

The results from the senior health service officials showed that the input measures that they deemed to be most relevant were pay costs and non-pay costs, each at 90.9%. The next inputs deemed most relevant were number of beds and total costs, again both scoring the same at 72.7%. The number of doctors and total number of staff were deemed the next most relevant at 63.6%. These were followed by the number of nurses and drug costs at 54.5%, and the number of support staff at 45.5%. The inputs deemed to be least relevant were the number of radiographers and modern equipment at 36.4%.

The results from the directors showed that they deemed the number of doctors to be the most relevant input measure at 73.7%. This was followed by the number of nurses, the number of beds and modern equipment, all at 68.4%. The total number of staff came next at 63.2% and this was followed by non-pay costs at 57.9%, pay costs and total costs at 47.4%, and the number of support staff at 42.1%. The inputs deemed to be least relevant were the number of radiographers and drug costs at 36.8%.

Interestingly looking at the top six choices of all groups there was only one input common to all groups. That was the number of doctors. There were four inputs common to the top six choices of three of the groups. These were the number of nurses, the total number of staff, the number of beds and modern equipment.

Aside from the preferences of senior health service officials, the use of various cost measures as inputs were not deemed to be as relevant by the other groups and did not appear as their top choices.

Both patients and hospital staff had four inputs in common in their top six choices. These were the number of doctors, the number of nurses, the number of beds and modern equipment. The only differences occurred where patients included the number of radiographers and the number of support staff in their top six choices whilst the hospital staff included the total number of staff and total costs. However both groups had the same inputs in common in their top eight choices albeit in a different order. Following on from this they obviously had the same bottom three choices. These were drug costs, pay costs and non-pay costs.

Patients and senior health service officials had only two inputs in common in their top six choices. These were the number of doctors and the number of beds. Both groups also differed in their bottom three choices of inputs. Senior officials deemed the number of radiographers, the number of support staff and modern equipment as the least relevant inputs. Unlike patients they also rated monetary measures higher by including pay costs, non-pay costs and total costs as their most relevant inputs. This is surely a reflection of their roles in the health service and the high priority given to financial control.

Hospital staff and senior health service officials had four inputs in common in their top six choices. This was probably not surprising given their roles within the health service. Although given that the questionnaire would have been circulated to a much more diverse group of staff within the hospital I might have expected more differentiation. The four inputs in common were the number of doctors, the total number of staff, the number of beds and total costs. Surprisingly unlike the senior health service officials, hospital staff deemed the three monetary measures of drug costs, pay costs and non-pay costs to be the least relevant measures. Perhaps this reflects the closeness of hospital staff to the delivery of services and their lack of interest in the wider financial context. Maybe it also reflects the perceived remoteness of senior health service officials from the delivery of services at the coal-face. It is interesting to note the similarities between the choices of patients and hospital staff.

Patients and hospital directors had four inputs in common in their top six choices. These were the number of doctors, the number of nurses, the number of beds and modern equipment. The only differences occurred where patients included the number of radiographers and the number of support staff in their top six choices whilst the directors included the total number of staff and non-pay costs. Both groups choice of their least relevant inputs was also different. Whilst they both included drug costs the directors unlike the patients also included the number of radiographers and the number of support staff in this category. It is interesting that the directors deem monetary input measures as being more relevant than do the patients. Again this is probably a reflection of the directors' role and their focus on financial management issues.

Hospital staff and directors had the most similar top six choices of inputs with five inputs in common. These were the number of doctors, the number of nurses, the total number of staff, the number of beds and modern equipment. The only differences occurred where hospital staff included total costs in their top six choices whilst directors included non-pay costs. Both groups' choice of their least relevant inputs was also different. Whilst they both included drug costs the directors unlike the hospital staff also included the number of radiographers and the number of support staff in this category.

Senior officials and directors had four inputs in common in their top six choices. These were the number of doctors, the number of staff, the number of beds and non-pay costs. The only differences occurred where senior officials included pay costs and total costs in their top six choices whilst directors included the number of nurses and modern equipment. Both groups also had two inputs in common in their choice of the three least relevant inputs. These were the number of radiographers and the number of support staff. Their other least relevant input was modern equipment in the case of senior officials and drug costs in the case of the directors.

Table 4.4

Outputs deemed relevant by each group in order of priority and percentages received

	Former Patients	%	Staff of the SIVUH	%	Senior Health Service Officials	%	Hospital Directors	%
1	Hygiene	98.4	Hygiene	91.4	Hygiene	100	Hygiene	94.7
2	Health & Safety	86.3	How quickly patients are treated	88.9	Inpatient waiting times	100	Approval by professional bodies	94.7
3	Nursing care	85.5	Patient satisfaction	88.9	Numbers treated without overnight stay	100	Patient satisfaction	89.5
4	How quickly patients are treated	83.1	Time waiting to be seen by a doctor in A&E	88.9	Length of waiting lists	100	Staff communicat-ions	89.5
5	Infection levels	81.5	Staff commun-ications	88.9	Ability of hospital to operate within budget	100	Infection levels	89.4
6	Staff communi cations	80.6	Length of waiting lists	86.4	Infection levels	90.9	Staff courtesy	84.2
7	Patient satisfact-ion	77.4	Infection levels	85.2	Staff communicat-ions	90.9	Nursing care	84.2
8	Staff friendl-iness	77.4	Inpatient waiting times	85.2	Health & Safety	81.8	Time waiting to be seen by a doctor in A&E	78.9
9	Staff courtesy	76.6	Nursing care	84.0	Level of complaints	81.8	Health & safety	78.9
10	Time waiting to be seen by a doctor in A&E	71.0	Health & Safety	82.7	Outpatient waiting times	81.8	Level of complaints	73.7
11	Inpatient waiting times	66.1	Approval by professiona l bodies	80.2	How quickly patients are treated	81.8	How quickly patients are treated	68.4
12	Length of waiting lists	65.3	Level of complaints	79.0	Approval by professional bodies	72.7	Length of waiting lists	63.2

13	Out-patient waiting times	64.5	Outpatient waiting times	77.8	How quickly patients are released from hospital after their treatment	72.7	Staff friendliness	63.2
14	Car parking facilities	63.7	Ability of hospital to operate within its budget	77.8	Patient satisfaction	72.7	Outpatient waiting times	57.9
15	Approval by professio nal bodies	60.5	Staff courtesy	74.1	Number of new patients attending the outpatient's department	72.7	Number of patients having to return to hospital unexpectedly	57.9
16	Food	59.7	Time that operating theatres are available	72.8	Time waiting to be seen by a doctor in A&E	72.7	Time taken by hospital staff to answer phone calls	57.9
17	Time that operating theatres are available	57.3	Numbers treated without overnight stay	71.6	Staff courtesy	72.7	Ability of hospital to operate within its budget	57.9
18	How quickly patients are released from hospital after their treatment	52.4	How quickly patients are released from hospital after their treatment	69.1	Food	63.6	Food	57.9
19	Numbers treated without overnight stay	50.0	Staff friendliness	69.1	Total patient numbers	63.6	Inpatient waiting times	57.9
20	Ability of hospital to operate within its budget	46.8	Number of new patients attending the out-patient's department	63.0	Number of patients returning for further outpatient appointments	63.6	Numbers treated without overnight stay	57.9
21	How easy it is to get to the hospital	43.5	Total patient numbers	60.5	Number of patients having to return to hospital unexpectedly	63.6	How quickly patients are released from hospital after their treatment	47.4
22	Time taken by hospital staff to answer phone calls	42.7	Number of patients having to return to hospital unexpect-edly	60.5	Number of patients who die at the hospital following treatment	63.6	Time that operating theatres are available	47.4
23	Number of patients returning for further out-patient appointm ents	41.9	Number of patients who die at the hospital following treatment	60.5	Nursing care	63.6	How easy it is to get to the hospital	42.1

24	Patients refer red to the hospital when they do not need hospital treatment	41.9	Food	56.8	Time that operating theatres are available	54.5	Number of patients who die at the hospital following treatment	42.1
25	Number of patients having to return to hospital unexpec-tedly	37.9	Car parking facilities	55.6	Patients referred to the hospital when they do not need hospital treatment	45.5	Total patient numbers	36.8
26	Number of new patients attending the outp-atient's depart-ment	37.1	Number of patients returning for further outpatient appoint-ments	51.9	Staff friendliness	45.5	Number of new patients attending the outpatient's department	36.8
27	Number of patients who die at the hospital follow-ing treatment	37.1	Patients referred to the hospital when they do not need hospital treatment	49.4	Car parking facilities	36.4	Number of patients returning for further outpatient appointments	36.8
28	Total patient numbers	33.1	Time taken by hospital staff to answer phone calls	46.9	Time taken by hospital staff to answer phone calls	27.3	Car parking facilities	31.6
29	Level of complain ts	31.5	How easy it is to get to the hospital	37.0	How easy it is to get to the hospital	9.1	Patients referred to the hospital when they do not need hospital treatment	21.1

Hygiene occupied the position of top choice with all of the groups. This is not surprising given the high level of media attention on this subject. Hygiene has been highlighted in the media as the main reason for hospital acquired infections. Whilst this may not always be the case, it is undoubtedly a contributory factor in some cases. There are four outputs that appear in the top ten choices of all four groups. These are hygiene, health and safety within the hospital, infection levels at the hospital and staff communications with patients and their families. Again like hygiene, hospital acquired infections have received much media attention and it is not unexpected to see it being highlighted by the groups. There would be an expectation that health and safety in a hospital would be important but the highlighting of staff communications with patients and their families is interesting. Whilst this should not be surprising, it is an area that can be neglected. In a busy hospital the priority is looking after the health of the patient but it is clear from these results that staff needs to be more cognisant of the communication needs of patients and their families. There are two outputs that appear in the top ten choices of three of the groups. These are the time waiting to be seen by a doctor in the accident and emergency department and nursing care. Waiting times in the accident and emergency department is a key performance measure in Irish hospitals that is regularly highlighted by patient representative groups as a problem and I am surprised that it did not feature more highly in the survey. Nursing care is also an important measure which not surprisingly was rated higher by the former patients' group. There are six outputs that appear in the top ten choices of two of the groups. These are how quickly patients are treated, patient satisfaction, level of complaints, length of waiting lists, inpatient waiting times and staff courtesy. How quickly patients are treated is critical to a hospital, as it is linked to the key performance indicators of the length of waiting lists and waiting times. These are hugely political issues that are always at the top of the health agenda. Patient satisfaction is a very important measure that is unfortunately not measured in Ireland in any comprehensive way. The level of complaints is also an important performance measure but caution needs to be exercised when using this measure as the number of complaints recorded may be dependent on the complaints process and the culture within the hospital. Staff courtesy is another measure that can be neglected but it is clear that it is important from a stakeholder's point of view. There are five outputs that appear in the top ten choices of at least one of the groups. These are the numbers of people who are treated without having to stay in hospital overnight, outpatient waiting times, the ability of the hospital to operate within its financial budget, staff friendliness and approval by professional bodies. Treating more patients as day patients without them having to stay in hospital overnight is an important aim of the health service and more patients are being treated as day patients through advances in medicine and hospital efficiencies. Outpatient waiting times is also a key performance indicator. A higher priority was given in the past to inpatient waiting times but this is changing now with more of an emphasis on outpatients. The ability of a hospital to operate within its budget is critical to its ability to treat patients. Staff friendliness, clearly like staff courtesy should not be neglected. Approval by professional bodies is necessary to ensure that medicine is safely practiced in a hospital. No other measure appears in the top ten choices of any of the groups.

The output measure deemed most relevant by former patients was hygiene at 98.4%. This was followed by health and safety at 86.3%, nursing care at 85.5%, how quickly patients are treated at 83.1%, infection levels at the hospital at 81.5%, staff communications with patients and their families at 80.6%, patient satisfaction at

77.4%, staff courtesy at 76.6% and time waiting to be seen by a doctor in the Accident and Emergency department at 71.0%. All other measures scored less than 70% with the level of complaints deemed the least relevant output measure at 31.5%. Amongst staff of the SIVUH the results showed that the output measure deemed most relevant was hygiene at 91.4%. This was followed by how quickly patients are treated at 88.9%, patient satisfaction at 88.9%, time waiting to be seen by a doctor in the Accident and Emergency department at 88.9%, staff communications with patients and their families at 88.9%, length of waiting lists at 86.4%, inpatient waiting times at 85.2%, infection levels at the hospital at 85.2%, nursing care at 84.0%, health and safety within the hospital at 82.7% and approval by professional bodies of the standards of the hospital at 80.2%. All other measures scored less than 80% with how easy it is to get to the hospital deemed to be the least relevant measure at 37%. The results from the senior health service officials showed that the outputs that they deemed to be most relevant were hygiene, inpatient waiting times, numbers of people who are treated without having to stay in hospital overnight, length of waiting lists and the ability of the hospital to operate within its financial budget, all at 100%. These were followed by infection levels at the hospital and staff communications with patients and their families, at 90.9%. Health and safety within the hospital, how quickly patients are treated, outpatient waiting times and the level of complaints all scored 81.8%. All other measures were deemed relevant by less than 80% of the group. The measure deemed to be the least relevant was how easy it is to get to the hospital at 9.1%.

The results from the directors showed that hygiene and approval by professional bodies of the standards of the hospital were deemed to the most relevant inputs, both at 94.7%. These were followed by patient satisfaction, infection levels at the hospital, and staff communications with patients and their families, all at 89.5%. Next came staff courtesy and nursing care at 84.2%. All other measures were deemed relevant by less than 80% of the group. The measure deemed to be the least relevant was patients referred to the hospital when they do not need hospital treatment at 21.1%.

Table 4.5

Combined results from all groups for inputs

Choice/Group	Combined Stakeholders	%
1	Number of doctors	73.0
2	Number of beds	72.1
3	Number of nurses	69.7
4	Modern equipment	67.5
5	Total number of staff	62.1
6	Total costs	54.7
7	Number of support staff	53.3
8	Non-pay costs	52.9
9	Pay costs	52.3
10	Number of radiographers	45.7
11	Drug costs	39.6

The combined results for inputs showed that overall the number of doctors was deemed to be the most relevant input at 73%. This is not surprising given that hospital consultants were the only input mentioned by more than one group in the first stage of the process. Hospital directors rated the number of doctors to be the most relevant input whilst former patients rated it to be the second most relevant input measure. Hospital staff rated it third and senior health officials rated it in joint fifth place. The deemed importance of this input measure ties in with Jarman et al. (1999) whose results indicated that the ratio of hospital doctors to beds and general practitioners to head of population was the second best predictor of the variation in mortality between hospitals in England. The percentage of cases that were emergency hospital admissions was deemed to be the best predictor of this variation in mortality. When analyses were restricted to emergency admissions only the number of doctors per bed was the best predictor.

Overall the number of beds was deemed to be the second most relevant input measure at 72.1%. Hospital directors deemed it to be the second most important measure whilst hospital staff and senior health officials deemed it to be the third most relevant and former patients deemed it to be fourth. This is an important measure as obviously not having enough beds will reduce the number of medical procedures carried out and thus impact on waiting lists. A shortage of beds could result in patients waiting in the emergency department for an unacceptable length of time and result in those waiting to be admitted to hospital having to wait on trolleys in corridors. At the same time the efficient use of beds has to be monitored. Reducing the average length of stay for patients will result in greater efficiencies and thus improve bed availability.

Overall the number of nurses was deemed to be the third most relevant input measure at 69.7%. Former patients deemed it to be the most relevant measure whilst directors deemed it to be the second most relevant, hospital staff deemed it to be fifth most relevant and senior health officials deemed it to be seventh. It is not surprising that former patients would deem this to be the most relevant input measure given that most of their interactions during their hospital stay would be with nurses. Their satisfaction with their hospital stay would depend greatly on the quality of the nursing service that they would have received.

Overall modern equipment was deemed to be the fourth most relevant input measure at 67.5%. Hospital staff deemed it to be the most relevant measure whilst directors deemed it to be second most relevant, patients deemed it to be third and senior officials deemed it to be tenth. Reliable and up to date equipment would be particularly important for hospital staff and not surprisingly they deemed this to be the most relevant input measure. Hospital directors deemed it to be as important as the number of nurses and the number of beds. Whilst this is an important input it would be difficult to measure accurately across all of the hospitals.

Overall the total number of staff was deemed to be the fifth most relevant input measure at 62.1%. Not surprisingly hospital staff deemed it to be the second most relevant measure whilst both senior health officials and directors deemed it to be fifth and patients deemed it to be in seventh place. The total number of staff is an important input but it may also be broken down into the different disciplines working in the hospitals. Disciplines such as doctors and nurses are deemed more relevant by the groups, even though other disciplines are critical to the efficient running of a hospital.

Total costs were deemed to be the sixth most relevant input measure at 54.7%. Senior officials deemed it to be the third most relevant measure whilst hospital staff deemed it to be the fifth, hospital directors deemed it to be seventh and patients deemed it to be eighth. The importance of this measure not surprisingly was rated highest by those in management roles. Controlling costs are a critical part of the efficient management of a hospital and whilst this might not have featured too highly in the priorities of former patients it is surprising that hospital directors only rated it in seventh place.

The number of support staff was deemed to be the seventh most relevant input measure at 53.3%. Patients deemed it to be the fifth most relevant measure, hospital staff deemed it to be the seventh most relevant measure, and senior officials and hospital directors deemed it to be the ninth most relevant. This measure was not rated highly by any of the groups. However, the importance of staff in this group, which would include catering, portering and housekeeping staff, should not be underestimated. Such staff is critical to the smooth running of a hospital and contributes hugely to patient satisfaction levels when in hospital.

Non-pay costs were deemed to be the eighth most relevant input measure at 52.9%. Senior health officials deemed it to be jointly the most relevant input measure whilst hospital directors deemed it to be sixth, hospital staff deemed it to be tenth and patients deemed it to be in eleventh place. Similar to total costs the importance of this measure not surprisingly was rated highest by those in management roles. Controlling non-pay costs are a critical part of the efficient management of a hospital and whilst this might not have featured too highly in the priorities of former patients it is surprising that hospital staff only rated it in tenth place.

Pay costs were deemed to be the ninth most relevant input measure at 52.3%. Senior officials deemed it to be jointly the most relevant input measure whilst hospital directors deemed it to be seventh, hospital staff deemed it to be ninth and patients deemed it to be tenth. Again like non-pay costs this measure was rated highest by senior health officials. It is surprising given the high pay bill of hospitals that hospital directors did not rate this measure higher.

The number of radiographers was deemed to be the tenth most relevant input measure at 45.7%. Former patients deemed it to be the sixth most relevant measure, hospital staff deemed it to be the eighth most relevant input measure whilst hospital directors and senior officials deemed it to be in tenth place. Whilst radiographers are important staff in a hospital, each of the groups obviously did not deem them to be as important as doctors or nurses, or indeed support staff. Hospital staff only rated them in eighth place.

Overall drug costs were deemed to be the eleventh and least relevant input measure at 39.6%. Senior officials deemed it to be the seventh most relevant input measure whilst former patients deemed it to be the ninth most relevant measure and both hospital directors and hospital staff deemed it to be the eleventh most relevant measure. Drug costs are a major issue for most hospitals and are difficult to control. Whilst the use of more generic drugs is helping to reduce costs, these cannot be produced until the patents expire on the respective drugs, and the introduction of new drugs is constantly pushing up costs. It is surprising that senior health officials or hospital directors did not rate this measure higher.

Table 4.6

Combined results from all groups for outputs

Choice/Group	Combined Stakeholders	%
1	Hygiene	96.1
2	Staff communications with patients and their families	87.5
3	Infection levels	86.8
4	Health & Safety	82.4
5	Patient satisfaction	82.1
6	How quickly patients are treated	80.6
7	Nursing care	79.3
8	Length of waiting lists	78.7
9	Time waiting to be seen by a doctor in A&E	77.9
10	Inpatient waiting times	77.3
11	Approval by professional bodies	77.0
12	Staff courtesy	76.9
13	Ability of hospital to operate within budget	70.6
14	Outpatient waiting times	70.5
15	Numbers treated without staying in hospital overnight	69.9

16	Staff friendliness	63.8
17	Level of complaints	63.8
18	How quickly patients are released from hospital following treatment	60.4
19	Food	59.5
20	The amount of time that operating theatres are available for operations	58.0
21	The number of patients having to return to hospital unexpectedly	55.0
22	The number of new patients attending the outpatients dept.	52.4
23	The hospital mortality rate	50.8
24	The number of patients returning for further outpatient appointments	48.6
25	Total patient numbers	48.5
26	Car parking facilities	46.8
27	Time taken by hospital staff to answer phone calls	43.7
28	Patients referred to the hospital when they do not need hospital treatment	39.5
29	How easy it is to get to the hospital	32.9

The combined results showed that hygiene was deemed to be the most relevant output measure at 96.1%. It was deemed to be the most relevant measure by all of the groups. A high media profile no doubt contributed to this result. This was followed by staff communications with patients and their families at 87.5%. In a busy hospital the priority is looking after the health of the patient but it is clear from these results that staff needs to be more cognisant of the communication needs of patients and their families. Infection levels at the hospital were next at 86.8%. Again like hygiene, hospital acquired infections have received much media attention and it is not unexpected to see it being highlighted by the groups. Patients and hospital directors rated this measure as the fifth most relevant measure whilst senior health officials and hospital staff rated it in sixth and seventh place respectively. Health & safety within the hospital was rated at 82.4%. It is not surprising, given that here would be an expectation that health and safety in a hospital would be important given the type of organisation it is, that this measure would be rated highly. Patient satisfaction was rated at 82.1%. Patient satisfaction is a very important measure that is unfortunately not measured in Ireland in any comprehensive way despite the obvious relevancy of the measure. Patient satisfaction is an excellent qualitative measure of how a hospital is providing its services. How quickly patients are treated is rated at 80.1%. How quickly patients are treated is critical to a hospital, as it is linked to the key performance indicators of the length of waiting lists, rated at 78.7%, and inpatient waiting times, rated at 77.3% . These are hugely political issues that are always at the top of the health agenda. These issues on their own can at times overshadow all other measures. Nursing care is rated at 79.3%. Nursing care is also an important measure which not surprisingly was rated higher by the former patients' group, who deemed it to be the third most relevant output measure. The time waiting to be seen by a doctor in the accident & emergency department is rated at 77.9%. Waiting times in the accident and emergency department is a key performance measure in Irish hospitals that is regularly highlighted by patient representative groups as a problem and I am surprised that it did not feature more highly in the survey. The approval by professional bodies of the standards of the hospital is rated at 77%. Approval by professional bodies is necessary to ensure that medicine is safely practiced in a hospital and is a critical requirement for all hospitals. Staff courtesy is rated at 76.9%. Staff courtesy is another measure that can be neglected but it is clear that it is important from a stakeholder's point of view. This is a measure that could be reflected in a patient satisfaction survey. The ability of the hospital to operate within its financial budget is rated at 70.6%. The ability of a hospital to operate within its budget is critical to its ability to treat patients. If a hospital cannot maintain its viability then it will not be able to provide the required service and quality may suffer. Outpatient waiting times is rated at 70.5%. Outpatient waiting times is also a key performance indicator. A higher priority was given in the past to inpatient waiting times but this is changing now with more of an emphasis on outpatients. Keeping people on the outpatient list and thereby keeping inpatient list waiting list short is no longer acceptable. The numbers treated without having to stay in hospital overnight is rated at 69.9%. Treating more patients as day patients without them having to stay in hospital overnight is an important aim of the health service and the number of patients that are being treated as day patients through advances in medicine and hospital efficiencies is increasing every year. This is also a key performance indicator for all hospitals in the Irish health service. Staff friendliness is rated at 63.8%. Similar to staff courtesy this is an important measure that influences patient satisfaction and that needs to be borne in mind by all staff. The level of complaints is rated at the same level as staff friendliness, which is interesting, at 63.8%. The level of complaints is also an important performance measure but caution needs to be exercised when using this measure as the number of complaints recorded may be dependant on the complaints process and the culture within the hospital. How quickly patients are released from hospital following treatment is rated at 60.4%. Whilst this may appear to be a reasonable measure care needs to be taken that patients are not released prematurely and need to be re-admitted within a short time. The number of patients having to return to hospital unexpectedly is a key performance indicator and is rated by the combined groups at 55%. Food is rated at 59.5%. It is surprising that this measure was not rated more highly given anecdotal comments from time to time in the media, appearing as the sixteenth most relevant output measure with former patients and hospital directors, the eighteenth most relevant measure with senior health service officials and the twenty fourth most relevant measure with hospital staff. The amount of time that operating theatres are available for operations is rated at 58%. This would be an excellent measure of operating theatre efficiencies which would also reflect on inpatient waiting lists for surgery. The number of new patients attending the outpatients' department is rated at 52.4%. This is an excellent measure of the efficiency of consultant led clinics, as generally consultants would see new patients at their first outpatient visit. Patients may be seen by other members of the consultant's team at subsequent visits. The hospital mortality rate is deemed relevant by 50.8% of the combined groups. Care should be taken when looking at mortality rates across hospitals. The hospital mortality rates however may not be comparable across all hospitals due to some hospitals treating a different type of patient. Some medical specialties may also have different mortality rates. Likewise some doctors may be slow to accept patients with a bad medical prognosis and only accept those people with a better chance of survival. The obvious limitations of this measure should be taken into account when interpreting the results. The number of patients returning for further outpatient appointments is rated at 48.6%. This measure will give an overall measure of activity in respect of return outpatient appointments. Total patient numbers is rated at 48.5%. This is an important measure that provides details on total patient activity within each hospital. However, as medical procedures carried out can differ between hospitals this figure needs to be adjusted for casemix to ensure comparability between hospitals. Car parking facilities is rated at 46.8%. This is an issue that can be reflected in a patient's satisfaction with the hospital's services. Time taken by hospital staff to answer phone calls is rated at 43.7%. This is an issue that can cause frustration for patients and their families and staff needs to be cognisant of this. Again this is an issue that can be reflected in a patient's satisfaction with the services provided. Patients referred to the hospital when they do not need hospital treatment is rated at 39.5%. Caution needs to be exercised with this measure as patients may be referred to hospitals by other medical practitioners for precautionary reasons. While it may not benefit the efficiency of the hospital it may be necessary to ensure patient safety. How easy it is to get to the hospital is rated last at 32.9%. This is a wider issue that may not be within the control of the management of the hospital and may relate to geographical location or transport systems.

4.11 Level of disagreement between stakeholder groups

The level of disagreement between each of the stakeholders regarding the choice of inputs and outputs raises a number of important issues. The most obvious of which is the difficulties that have to be overcome in designing a performance measurement model that satisfies all stakeholders' requirements.

While the overall variation between most of the stakeholder groups in their choice of inputs was not significant there were still areas of difference between them. In particular, the choices of senior health service officials were at variance with the other groups. Their choices indicated that they were more focused on costs than any of the other groups. While also focusing on the total number of staff as an input the staff group that they deemed most relevant to include separately in the model were doctors but only in sixth place in order of relevance. In common with all of the groups they also deemed the number of beds to be a relevant input. It was clear that this group saw performance measured predominantly in terms of returns on financial input. This should not have been unexpected given their role in a difficult financial climate where the focus was increasingly aimed at achieving financial efficiencies and delivering value for money. One might have expected a similar result from the hospital directors' group given their governance and financial roles but this was not the case. They deemed non-pay costs to be relevant but only in sixth place in order of relevance. However, they had more in common with both the former patients group and the staff of the SIVUH group. All three groups had four inputs in common in their top six choices. These inputs were the number of doctors, the number of nurses, the number of beds and modern equipment. There was an interesting level of consistency reflected across these groups. The only difference in inputs was that the staff of SIVUH and hospital directors deemed costs to be relevant whilst the former patients did not. It is understandable that the former patients did not mention costs. Their exposure to the hospital system would have been largely based on direct interaction with clinicians and other support staff and not to hospital running costs. When the results of the questionnaires were combined as set out in table 4.5 the top four inputs were the number of doctors, the number of beds, the number of nurses and modern equipment with scores of 73%, 72.1%, 69.7% and 67.5% respectively. The total number of staff came next at 62.1% and this was followed by total costs at 54.7%. It was clear from the results that these were deemed to be the main inputs that should have been included in the performance measurement model.

The combined results indicated a much more varied list of output measures deemed to be relevant by each of the groups. Interestingly, when the outputs of all of the groups were combined in table 4.6, the results of the senior health officials' group again stood out as being at variance with the other groups. The staff at SIVUH group had seven of their top ten outputs in common with both the former patients' group and the hospital directors' group. Likewise, the hospital directors' group and the former patients' group had eight output measures out of their top ten outputs in common. However, the senior health service officials' group had only four outputs in common in their top ten choices with the former patients' group, five outputs in common with the hospital directors and six outputs in common with the SIVUH staff group. This again appeared to indicate a different attitude amongst the senior health service officials group to performance measurement. All of the groups deemed hygiene to be the most relevant output measure with a score of 96.1%. This was not surprising given the high profile given by the media to this area. What was surprising was that staff communications with patients and their families was deemed to be the second most relevant output measure with a score of 87.5%. All groups included this measure as one of their most relevant output measures. The senior health services officials' group rated this measure highest at 90.9%, the hospital directors' group rated this measure at 89.5%, followed by the staff of SIVUH group at 88.9% and the former patients' group last at 80.6%. It was interesting to note how highly this measure was regarded by those in governance roles and not quite as highly by former patients. Next came infection levels at 86.8%. This is not surprising given the high public awareness of hospital acquired infections such as MRSA and Clostridium difficile. Health and safety, patient satisfaction and how quickly patients are treated came next at 82.4%, 82.1% and 80.6% respectively.

Whilst overall the output results were reasonably consistent across all groups a number of results were surprising. One such result was that food was only rated at

59.5% overall, with the highest rating being received from the senior health service managers' group at 63.6%. Similarly, a low overall rating of 50.8% was given to the measure showing the number of patients who die following their treatment. More surprising still was that the group who gave the lowest rating to this measure at 37.1% was the former patients' group.

What was not surprising was the number of qualitative measures that were deemed to be relevant for inclusion in the performance measurement model. Measures such as hygiene levels, staff communication with patients and their families, infection levels, patient satisfaction, staff courtesy and staff friendliness all scored highly with each of the groups. This should make health service policy makers sit up and take notice of what people are saying is relevant to them. There are so many of these areas that are important to patients that need to be properly measured.

Step 3

The third step in the process was to establish an expert group to validate the inputs and outputs deemed relevant in the first two stages of the process. This group consisted of six academic experts from Aston University, Birmingham who were not involved in any way in the research. They were chosen because of their familiarity with the area being assessed and their ability to provide information regarding the outcomes being pursued and the resources needed in pursuit of these outcomes. A questionnaire containing the same questions that were used in the second step in the process was circulated to this group. On this occasion however two additional columns were included in the questionnaire. Each of the respondents was now being asked to indicate by ticking each column whether they deemed each input or output measure to be relevant, informative and/or necessary. In addition, space was provided for each expert to include any input or output that they felt should have been included, but that had not.

4.12 Results from the Experts' questionnaire

A response was received from five of the six experts circulated. Three responses were in the format requested in the questionnaire; one respondent replied using a different format and one respondent stated that he was not willing to take part in the survey. However, whilst refusing to take part in the survey he suggested that the appropriate approach would be the necessary and careful consideration of what were the relevant sets of inputs and outputs and very significantly in health, what he would consider a separate category, quality indicators. He suggested that it would need a set that adequately captured these, without excessive duplication, and that this could only be constructed by systematic consideration including statistical analysis of the relationship between variables. The other respondent whilst not replying in the format requested set out several possible sets of inputs and outputs to assess different aspects of performance. He felt that inputs needed to be considered jointly with outputs as it was only possible to judge if an input was relevant when it was known what the outputs were and the other way around. He set out some input output sets as follows:

1. Aspect of performance being assessed: Volume and clinical quality of care delivered relative to operating and capital expenditure.
   Inputs: Operating expenditure
   Non staff operating expenditure excluding medicine
   Expenditures on medicine
   Number of beds as a proxy for capital
   Outputs: Total patient numbers
   Number of new patients attending the Outpatients' department
   Number of patients who survive following treatment
   Number of patients not having to return to hospital unexpectedly
2. Aspect of performance being assessed: Speed of delivery of care set against staff, equipment and bed availability
   Inputs: Staff operating expenditure
   Non staff operating expenditure excluding medicine
   Number of beds as a proxy for capital
   Number of operating theatres
   Outputs: Number of patients admitted within x days of referral
   Number of patients seen in A&E within x minutes
   Number of outpatients admitted within x days of referral
3. Aspect of performance being assessed: Non clinical quality of care received by patients
   Inputs: Number of inpatients treated
   Number of outpatients treated
   Outputs: Number of patients' infection free during treatment
   Number not complaining

The responses from the other three experts are set out in tables 4.7 and 4.8. In each of the tables the numbers opposite each of the inputs and outputs indicate the number of experts who deemed them to be relevant, informative and/or necessary. These are then added together to give a total score for each input and output.

Table 4.7

Results from Expert Group – Inputs

Input	Relevant	Informative	Necessary	Total Score
Number of doctors	2	2	3	7
Number of beds	2	3	2	7
Pay costs	3	2	2	7
Non-pay costs	3	2	2	7
Total costs	2	2	3	7
Number of nurses	2	2	2	6
Total number of staff	2	3	1	6
Drug costs	2	2	1	5
Number of radiographers	2	1	1	4
Number of support staff	1	3	0	4
Modern equipment	1	0	0	1

Based on these responses all three of the experts deemed both the number of doctors and total costs to be necessary inputs. Two of the experts deemed the number of nurses, the number of beds, pay costs and non-pay costs to be necessary inputs. Only one of the experts deemed the number of radiographers, the total number of staff and drug costs to be necessary inputs. None of the experts deemed the number of support staff or modern equipment to be necessary inputs. Whilst at least one of the experts deemed each input as relevant only pay costs and non-pay costs were deemed to be relevant by all three experts.

Combining the scores under each heading would indicate that the number of doctors, the number of beds, pay costs, non-pay costs and total costs are the most relevant, informative and necessary inputs to include in the performance measurement model. These inputs are closely followed by the number of nurses and the total number of staff.

Table 4.8

Results from Expert Group - Outputs

Output	Relevant	Informative	Necessary	Total Score
Patient satisfaction	2	3	2	7
Total patient numbers	3	2	2	7
Infection levels	2	3	2	7
Mortality rate	2	2	3	7
Level of complaints	2	2	3	7
How quickly patients are treated	2	2	2	6
How quickly patients are released from hospital following treatment	2	3	1	6
Length of waiting lists	2	2	2	6
Outpatient waiting times	2	3	1	6
Inpatient waiting times	1	3	1	5
Hygiene	2	2	1	5
Numbers treated without staying in hospital overnight	3	2	0	5
Ability of hospital to operate within budget	1	2	2	5
The amount of time that operating theatres are available for operations	1	2	2	5
The number of patients having to return to hospital unexpectedly	1	2	2	5
Health & Safety	1	2	1	4
Patients referred to the hospital when they do not need hospital treatment	2	1	1	4
The number of patients returning for further outpatient appointments	2	2	0	4

Time waiting to be seen by a doctor in A&E	1	2	1	4
Staff communications with patients and their families	2	1	1	4
Approval by professional bodies	2	1	0	3
Staff courtesy	2	1	0	3
Staff friendliness	1	1	1	3
Nursing care	1	1	1	3
Food	2	0	0	2
The number of new patients attending the outpatients department	2	0	0	2
Time taken by hospital staff to answer phone calls	1	1	0	2
Car parking facilities	0	0	0	0
How easy it is to get to the hospital	0	0	0	0

Based on these responses, only the number of patients who die at the hospital following treatment and the level of complaints are deemed to be necessary outputs by all three of the experts. The following output measures are deemed necessary by two of the experts:

- Infection levels at the hospital
- Patient satisfaction
- How quickly patients are treated
- Length of waiting lists
- The ability of the hospital to operate within its financial budget
- The amount of time for which operating theatres are available for operations
- The number of patients having to return to hospital unexpectedly
- Total patient numbers

4.13 Level of disagreement between the stakeholders and the expert group

The results from the combined questionnaires of the four stakeholder groups are compared with those of the academic expert group in tables 4.9 and 4.10. Table 4.9 compares those inputs deemed relevant by the combined group of stakeholders with those deemed relevant, informative and necessary by the academic expert group in order of priority. The order of priority for the choices of inputs of the academic expert group was determined by adding the ratings that each achieved under the headings of relevant, informative and necessary. Table 4.10 provides the same information in respect of outputs.

Table 4.9

Inputs deemed relevant by stakeholders and deemed relevant, informative and necessary by academic expert group

Choice/ Group	Combined Stakeholders	%	Academic Experts	Academic Experts' Score
1	Number of doctors	73.0	Number of doctors	7
2	Number of beds	72.1	Number of beds	7
3	Number of nurses	69.7	Pay costs	7
4	Modern equipment	67.5	Non-pay costs	7
5	Total number of staff	62.1	Total costs	7
6	Total costs	54.7	Number of nurses	6
7	Number of support staff	53.3	Total number of staff	6
8	Non-pay costs	52.9	Drug costs	5
9	Pay costs	52.3	Number of radiographers	4
10	Number of radiographers	45.7	Number of support staff	4
11	Drug costs	39.6	Modern equipment	1

The comparison between both groups indicated a clear agreement that the number of doctors and the number of beds should be included in the performance measurement model. Total costs and the number of nurses were also included in the top six choices of both groups and were considered for inclusion in the model. Pay costs and non-pay costs were rated higher than the number of nurses by the academic experts but were not deemed to be as relevant by the stakeholders' groups. Likewise the stakeholders' groups deemed modern equipment and the total number of staff to be more relevant than total costs but the academic experts did not rate them as highly. Information on modern equipment in each hospital was not available and could not be included in the model. The number of beds was used instead as a proxy for capital consumed. Pay costs, non-pay costs and the total number of staff were therefore considered as inputs in the model

Table 4.10

Outputs deemed relevant by stakeholders and deemed relevant, informative and necessary by academic expert group

Choice/Group	Combined Stakeholders	%	Academic Experts	Academic Experts' Score
1	Hygiene	96.1	Patient satisfaction	7
2	Staff Communications with patients and their families	87.5	Total patient numbers	7
3	Infection Levels	86.8	Infection levels	7
4	Health & Safety	82.4	Mortality rate	7
5	Patient Satisfaction	82.1	Level of complaints	7
6	How quickly patients are treated	80.6	How quickly patients are treated	6
7	Nursing Care	79.3	How quickly patients are released from hospital after treatment	6
8	Length of Waiting Lists	78.7	Length of Waiting Lists	6
9	Time waiting to be seen by a doctor in A&E	77.9	Outpatient waiting times	6
10	Inpatient Waiting Times	77.3	Inpatient waiting times	5
11	Approval by Professional Bodies	77.0	Hygiene	5
12	Staff Courtesy	76.9	Numbers treated without having to	5

			stay overnight	
13	Ability to operate within Budget	70.6	Ability to operate within Budget	5
14	Outpatient Waiting Times	70.5	Amount of time that theatres are available	5
15	Numbers treated without having to stay overnight	69.9	Number of patients having to return unexpectedly	5
16	Staff Friendliness	63.8	Health & Safety	4
17	Level of Complaints	63.8	Patients referred to hospital when they do not need treatment	4
18	How quickly patients are released from hospital after treatment	60.4	Number of patients returning for outpatient appointments	4
19	Food	59.5	Time waiting to be seen by a doctor in A&E	4
20	Amount of time that operating theatres are available for operations	58.0	Staff Communications with patients and their families	4
21	Number of patients having to return unexpectedly	55.0	Approval by Professional Bodies	3
22	Number of new patients attending the outpatients department	52.4	Staff courtesy	3
23	Number of patients who die following treatment	52.4	Staff friendliness	3
24	Number of patients returning for outpatient appointments	48.6	Nursing care	3
25	Total patient numbers	48.5	Food	2
26	Car parking facilities	46.8	Number of new patients attending the outpatients department	2
27	Time taken by hospital staff to answer phone calls	43.7	Time taken by hospital staff to answer phone calls	2
28	Patients referred to hospital when they do not need treatment	39.5	Car parking facilities	0
29	How easy it is to get to the hospital	32.9	How easy it is to get to the hospital	0

Comparison between the top 50% outputs, i.e. 15 outputs, of both groups indicated that they had nine outputs in common. These were hygiene; infection levels; patient satisfaction; how quickly patients are treated; length of waiting lists; inpatient waiting time; ability of hospital to operate within budget; outpatient waiting times, and numbers treated without staying in hospital overnight. Each of these measures will have to be considered when specifying the outputs in the model. The measures that were deemed relevant by the combined stakeholders' groups and that were not included in the academic experts' top fifteen choices were staff communications with patients and their families, health & safety, nursing care, time waiting to be seen by a doctor in the emergency department, approval by professional bodies, and staff courtesy. While these are all important factors that affect the perception of a hospital in the minds of the public they were not all readily measurable. In fact the only one that was easily measured was the time waiting to be seen by a doctor in the emergency department and this was deemed to be more informative than relevant or necessary by the academic expert group. The measures that were included in the top fifteen choices of the academic expert group and that were not regarded as such by the combined stakeholders' groups were total patient numbers, mortality rate, level of complaints, how quickly patients are released from hospital following treatment, the amount of time that operating theatres are available for operations, and the number of patients having to return to hospital unexpectedly. Of these, the mortality rate and the level of complaints were regarded as being necessary by all of the academic experts whilst being deemed relevant by 58.8% and 63.8% respectively of the combined stakeholders' group. As both of these measures were deemed necessary by all of the academic experts they were considered as output measures when designing the model. Total patient numbers, the amount of time for which operating theatres are available for operations and the number of patients having to return to hospital unexpectedly were deemed necessary by two of the academic experts even though they were only deemed relevant by 48.5%, 58% and 55% respectively of the combined stakeholders' group. For this reason they were considered when deciding on the outputs to be included in the final model. How quickly patients are released from hospital after their treatment was only deemed necessary by one academic expert and was not considered when deciding on the output specifications for the model.

4.14 Model specification

It was important that the measures of inputs and outputs that were included in the model were as comprehensive as possible. It was also important to ensure that hospitals were not disadvantaged by excluding an output that they were relatively efficient at producing. At the same time including too many different inputs and outputs may have resulted in inflated efficiency scores, thus allowing more scope for each hospital to be relatively unique. The objective was to include the smallest number of input and output measures that adequately captured all essential aspects of the hospitals' operations.

The process of developing the final model involved testing different combinations of inputs and outputs. As well as ensuring that the most appropriate measures of input and output were used, this process also allowed the sensitivity of the model to different specifications to be tested.

Inputs

Based on the above the inputs that were considered for inclusion in the performance measurement model were:

- Number of full-time equivalent doctors
- Number of beds
- Number of full-time equivalent nurses
- Total costs
- Non pay costs
- Pay costs
- Total number of staff

In considering which of these inputs to include in the model it was necessary to ascertain both the availability and the accuracy of the data. All of the suggested inputs were available for each of the acute hospitals in Ireland from the Irish Health Service Executive. The accuracy of the data was dependant on the reliability of the Health Service Executive's published reports.

> *The production function is central to the economic theory of production because it provides some of the information needed to calculate the costs of output. Without it we would not know the amounts of resources required to provide different levels of output. Adding market prices for factor inputs allows the calculation of costs based on the production function…. Combining factor prices with the production function yields the cost function.*
>
> (Hollingsworth and Peacock, 2008: 13)

The model is measuring the technical efficiency of the hospitals. Technical efficiency means transferring physical inputs such as labour and capital into outputs at the best level of performance. Given a choice of either using full-time equivalent staff numbers or staff salary payments to measure labour input, I would be more in favour of using the latter. I say this because salary scales and allowances are standard across all hospitals, regardless of size or location, whilst using full-time equivalent staff numbers may not pick up all relevant costs. These costs would include overtime, various allowances and premium payments. This is important because, a hospital that pays overtime instead of increasing staff numbers will appear more efficient from a staffing perspective but may be less efficient from a costing perspective. One drawback with using salary costs instead of full-time equivalents is that visiting doctors may not be recognised if they are being paid by their main hospital. However having given my reasons for preferring cost measures over physical measures I used both types to test whether the use of either measure led to any significant differences in the results. If the results are different it may mean that salary costs per full-time equivalent are varying significantly between hospitals or hospitals are treating other relevant costs inconsistently.

Managers can change both salary costs and full-time equivalent staff numbers by using more efficient work practices. The basic salary scale may be fixed under national pay agreements but the number of staff required or their overtime hours or other premium payments may be more efficiently managed. Policy makers are interested in improving efficiency and this is done by either maximising output whilst keeping costs, including labour costs, constant or minimising costs whilst keeping output constant. The choice of approach is dictated by the organisation's priorities and the economic environment pertaining at the time. The input orientated approach may be appropriate if there is a shortage of finance or if demand for output is reducing. Similarly the output orientated approach may be more appropriate if demand for output is increasing.

The choice of nurses and doctors in the model reflected the traditional division of labour used in hospitals. The majority of non-medical staff are nurses whilst medical staff comprise physicians, surgeons and non-consultant hospital doctors (trainee doctors).

Overall pay costs or total full-time equivalents were also used in the model. Both of these were used as inputs to test whether the use of either measure led to any significant differences in the results.

Inputs other than labour are important for providing acute hospital services. Ideally any model of performance measurement should account for all inputs used by each hospital. Whilst this is not always possible non pay costs which were made up of all materials and services used including food, pharmaceuticals, medical consumables, gas and electricity were included in the model. Similarly total costs were also used in the model.

The number of beds in each hospital was used as a proxy for capital inputs. This was used because measuring capital inputs is extremely difficult and subject to considerable variation. The difficulty in measuring capital inputs is that a capital item provides a flow of services over a number of years and it is difficult to determine how much of the purchase price should be charged to each period along with how interest and depreciation costs should be allocated. As well as this it is extremely difficult to get accurate and reliable figures for capital inputs in all of the acute hospitals in Ireland. Whilst clearly using the number of beds is not the most accurate proxy it is a simple measure that can be used subject to its obvious limitations.

Outputs

Based on the results above, the outputs that needed to be considered in the performance measurement model were:

- Hygiene
- Infection levels at the hospital
- Patient satisfaction
- How quickly patients are treated
- Length of waiting lists
- Level of complaints
- Number of patients who die at the hospital following treatment
- Total patient numbers
- The ability of the hospital to operate within its financial budget
- The amount of time for which operating theatres are available for operations
- The number of patients having to return to hospital unexpectedly
- Inpatient waiting times
- Outpatient waiting times
- Number of patients treated without having to stay in hospital overnight

Dyson et al. (2001) state that there are four key assumptions with respect to the input/output set selected:

- it covers the full range of resources used
- captures all activity levels and performance levels
- the set of factors are common to all units
- environmental variation has been assessed and captured if necessary.

They identified three pitfalls that need to be addressed. These are the number of inputs and outputs, correlated factors, and mixing indices and volume measures. The first pitfall identified is the number of inputs and outputs. As DEA allows flexibility in the choice of weights on the inputs and outputs, the greater the number of factors included the lower the level of discrimination. In order to achieve a reasonable level of discrimination a suggested rule of thumb is that the number of inputs and outputs included in the model should be at least twice the product of the number of inputs and outputs. Discrimination can be increased, therefore, by being parsimonious in the number of factors used. On the input side if there are inputs that can be priced, then the flexible weights can be replaced by fixed prices, thus reducing the number of inputs. Similarly on the output side, discrimination can be enhanced by eliminating any performance measures that are not strongly related to the objectives of the organisation. The second pitfall identified is correlated factors. Given that subsets of the inputs and outputs are often correlated, it is tempting to omit such factors in order to increase discrimination. While this is unlikely to have a major impact the omission of a highly correlated variable can on occasion lead to significant changes in efficiencies. Omission of variables purely on grounds of correlation should therefore be avoided. The third pitfall identified is to mix indices often associated with performance measures, with activity levels, which are volume measures. One approach would be to use some surrogate measure rather than an index. An example given by Dyson et al. (2001) is that of a local authority that could use, rather than an index for social deprivation, the number of summonses and distress warrants issued to recalcitrant payers of local taxes, which would be replacing an index with a volume measure.

Dyson et al. (20001) identified four pitfalls when measuring factors. These relate to percentages and normalised data, qualitative data, undesirable inputs and outputs, and exogenous and constrained factors. The first pitfall occurs with the desire to incorporate indices, ratios or percentages into the input/output set. This may be acceptable if all the inputs and outputs are of this kind, but the danger occurs when attempts are made to mix them with volume measures. One way of dealing with this pitfall is to use a proxy measure. Another approach would be to scale the index percentage by a volume measure to make it compatible with any other volume measures. A third approach would be to separate the numerator and denominator, and to include the numerator as an input and the denominator as an output. Dyson et al. give as an example the assessment of peri-natal care units (Thanassoulis et al., 1995) where the infant mortality rate was perceived as a key performance measure.

The approach adopted was to include babies at risk in the denominator and survivors in the same category in the numerator. The second pitfall identified the challenge of incorporating qualitative variables into an analysis. Attempts to measure factors such as customer perception of service quality are identified as being problematic for a DEA evaluation in two distinct ways. The first issue is that such measures are often highly subjective and secondly the same satisfaction rating among customers in different branches of the same organisation may correspond to different levels of service quality. This may arise for example where customers of branches located in affluent areas may often have higher expectations regarding the quality of service provided to them than customers in other areas. One way suggested of dealing with this pitfall is to use surveys with care, and to try to cover a large number of respondents with an instrument designed to reduce the effect of subjectivity on the measurement process. An alternative approach suggested would be to use categorical or ordinal variables for which a number of model extensions are available (Cook et al., 1993; Banker and Morey, 1986). The third pitfall identified was the handling of undesirable inputs and outputs in the DEA model. These anti-isotonic data include undesirable outputs such as the emission of pollutants or inhibiting inputs such as the number of competitors impacting on a business unit. The approaches suggested to address this pitfall are to invert the anti-isotonic factor; to subtract the value of the undesirable factor from a large number, the result being isotonic, or thirdly to move the variable from the output to the input side of the model, or vice versa. The final pitfall identified is dealing with exogenous and constrained factors. If the exogenous factors are used as variables in a standard DEA model, then it may be possible to either include all the exogenous factors as inputs and use an output oriented DEA model, or include all the exogenous factors on the output side and use an input oriented model. Another suggested approach is to use DEA models that account explicitly for the existence of exogenous and/or constrained factors.

All of these pitfalls need to be borne in mind, as well as the suggested protocols for avoiding them, when designing the DEA model.

Hygiene

Hygiene is deemed to be the most relevant output measure by all of the stakeholder groups as well as being the first priority of the combined stakeholders' group. The academic expert group did not give it as high a priority with nine other output measures rated higher. The percentage of stakeholders that deemed this measure to

be relevant ranged from 100% in the case of senior officials to 91.4% in the case of hospital staff. These high ratings are not surprising given the wide public interest in the incidence of hospital acquired infections. Rightly or wrongly infections such as MRSA, Clostridium difficile and others are often blamed on inadequate hospital hygiene. Whilst this is not necessarily the main reason for the increasing incidence of such infections it is seen by the general public as such. It is also the case that even if there are other reasons for acquiring an infection, inadequate hygiene factors can contribute to the spread of the infection and in a hospital environment this can be catastrophic. The Health Information and Quality Authority (HIQA) carry out annual hygiene audits on all public acute hospitals in Ireland and their findings are publicised. Hospitals that achieve a poor hygiene rating in this audit are severely criticised not just by the Department of Health and Children but also in the media and by the general public. This has certainly focussed peoples' attention on maintaining hygiene standards in hospitals.
These figures were used in the performance measurement model.

Infection levels

Infection levels at the hospital were included by all groups as one of their top ten output measures. The percentage that deemed this measure to be relevant ranged from 90.9% in the case of senior health service officials to 81.5% in the case of patients. As in the case of hygiene there is a wide public interest in this measure. The only surprise with this result is that patients deemed this measure to be less relevant than did the other groups.
Infection levels were regarded as a relevant and necessary performance indicator by the combined stakeholders' group and the academic expert group. This measure may influence people when choosing the hospital where they would wish to have their medical procedure carried out. Obviously this may not be as important an issue for emergency patients as they may not be in a position to choose their hospital. The infection levels in a hospital may be assumed by the public to be indicators of cleanliness, even though this may not always be the case. However it could be regarded as a quality measure of a hospital. Data on infection levels in hospitals in Ireland has not always been readily available and it is only in recent years that hospitals have been providing this information. Providing such information has been a very sensitive issue for hospitals as high infection levels have the potential to seriously impact on their perceived status and their ability to attract patients. Negative publicity such as this may also adversely affect staff morale and would need to be

sensitively handled. Infection levels for MRSA are now regularly provided by hospitals and it is this measure that was used in the model. The figure that was used was the MRSA rate per 1,000 bed days used in each hospital for 2007, produced by the Irish European Antimicrobial Resistance Surveillance System (EARSS).

Patient satisfaction

Patient satisfaction was included by hospital staff and hospital directors as one of their top ten output measures with percentages of 88.9% and 89.5% respectively. It is surprising that patients only rated this measure at 77.4%.

Patient satisfaction measures have traditionally been the main method for gauging patients' views on healthcare. However as this information is not routinely recorded in Ireland it is only through the use of specially designed patient surveys, issued to representative samples of patients, that this data is collected. Surveys of public views on health system performance require careful consideration and may be misleading. Patients with a more positive health status are likely to be more satisfied than those with a less positive health status as a result of their experience of health care. Because of the lack of information little is known about Irish patients' views of the Irish Health Service. This was also the case in the UK. A report prepared by the Commission for Health Improvement (CHI) recommended in 2002 that:

- Investment in the further development of robust survey instruments for use in different types of Trusts and with specific groups of patients should be a priority.
- An ongoing programme of research would be useful as so little is known at present.
- Lay people and patients should be involved in identifying key indicators and designing report formats for performance information.
- It seems likely that many patients would prefer to obtain information on health care performance via their GPs.

Because of this lack of information on patient satisfaction levels in Ireland it was not possible to use patient satisfaction as an output measure in the performance measurement model.

How quickly patients are treated

How quickly patients are treated was included by two groups as one of their top ten output measures. It was rated most highly by staff at 88.9% and least highly by directors at 68.4%. This is an important measure as the speed at which a patient is

treated can have major implications for their chances of survival or their long-term quality of life. It could thus be regarded as a qualitative measure.
How quickly patients are treated is also a measure of efficiency. The numbers of patients treated is indicative of the efficiency with which patients are treated. The numbers of inpatient, day patient and new outpatients treated is an appropriate measure of this efficiency and these measures were used in the model.

Length of waiting lists

The length of waiting lists was included by two groups as one of their top ten output measures. The percentage that deemed this measure to be relevant ranged from 100% in the case of senior health officials to 63.2% in the case of hospital directors. Patients only rated this measure at 65.3%. Again for such a much publicised measure it is surprising that it did not score higher, particularly amongst patients.
The length of waiting lists is also a measure of efficiency. The numbers of patients on a waiting list is probably indicative of the efficiency with which patients are treated. The numbers of inpatient, day patient and new outpatients treated is an appropriate measure of this efficiency and these measures were used in the model. The DEA results, however, do not indicate shorter waiting lists for those hospitals deemed to be efficient. Efficiency scores do not appear to impact on the length of waiting lists. Possibly the fact that a hospital is seen to be efficient may attract more patients, thus lengthening the waiting lists.

Level of complaints

The level of complaints was included by senior health service officials and hospital directors as one of their top ten output measures with percentages of 81.8% and 73.7% respectively. The level of complaints can be a reflection of the quality of services being provided by the hospital. There are a number of issues however that militate against using it as an accurate measure. The first is that as it is largely dependant on the complaints systems used in each hospital, those hospitals with a good system of recording complaints may appear to have a higher level of complaints than those with a bad system. Secondly some hospitals may have a culture where complaints are welcomed and accepted as a means of improving the quality of service whilst other hospitals may frown upon any complaints and even perhaps under-record complaints as they arise. Therefore the use of complaints received as a measure of quality may not be very accurate. In any event national statistics for hospital complaints are currently not recorded in Ireland and thus not available to be included in the performance measurement model.

Number of patients who die at the hospital following treatment

Another output measure would be to look at mortality rates across hospitals. These rates however may not be comparable across all hospitals due to some hospitals treating a different type of patient. Some medical specialties may have different mortality rates. Likewise some doctors may be slow to accept patients with a bad medical prognosis and only accept those people with a better chance of survival. Jarman et al. (1999) looked at mortality rates in English hospitals using routinely collected data and concluded, having adjusted for casemix, that the percentage of total admissions classified as emergencies is the most powerful predictor of variations in hospital mortality. They also found that the ratios of hospital doctors to beds and general practitioners to head of population served, seem to be critical determinants of standardised hospital death rates; the higher these rates, the lower the death rates in both cases. Similarly the socio-economic aspect of the region in which the hospital is located may influence the medical outcome for the patient. To ensure comparability between hospitals those hospitals that do not have an Emergency Department have been excluded from the research. Survival rates have been used as an output in the model. However, the obvious limitations of this measure should be taken into account when interpreting the results.

Total patient numbers

Total inpatient numbers treated is an important output measure within each hospital. However, as medical procedures carried out can differ between hospitals this figure was adjusted for casemix to ensure comparability between hospitals. These figures were readily available for each hospital each year from the National Casemix Programme and have been included in the performance measurement model. There are now 37 public acute hospitals participating in the National Casemix Programme. Casemix works by coding hospital activity, using the hospital inpatient enquiry (HIPE) programme, and assessing hospital costs, using the specialty costs programme. The HIPE programme currently operates in the 62 biggest hospitals in the country. When a patient is discharged from hospital their age, gender, diagnosis, procedures performed and discharge status is coded using the WHO international classification of diseases (I.C.D.) which allows for 12,000 diagnoses and 8,000 individual procedures, each of which is allocated a separate code. Weightings are then applied to each procedure and diagnosis using an Australian casemix system, ICD-10-AM. The data is then grouped into over 6,000 DRGs. The basis of the entire system is to break down illnesses into 25 major diagnostic categories (M.D.C.s) based around body parts.

Ability to remain within budget

The ability of a hospital to remain within its budget is critical for the survival of that hospital. Of course whether a hospital does or does not remain within its budget can be influenced by many factors both internal and external. The ability of hospital management to ensure that the hospital manages its budget is an important factor but there are also external factors outside of management's control that can impact on achieving this objective. These could relate to situations where patients who require high cost drugs, expensive blood products or costly medical devices impact negatively on expenditure. This is particularly an issue in hospitals with emergency departments where it can often be difficult to predict the type of patient that will need to be treated. The mix of public or private patients that are treated in a hospital can also influence the budgetary outcome. In public hospitals in Ireland the amount of money charged to a public patient for treatment is tightly controlled and does not equate to the economic cost of treating that patient. Whilst the charge levied on a private patient attending a public hospital is also controlled it is much closer to the economic cost of treating that patient. Therefore a higher than expected proportion of public patients attending the hospital could have an adverse effect on the hospital's budgetary situation. For these reasons I do not believe that the ability of a hospital to remain within its budget is necessarily an indication of the efficiency or otherwise of that hospital. Having said that, it could also be argued that all hospitals are exposed to the same external factors and that a hospital that cannot remain within its budget is probably inefficient. However, that also assumes that the funding provided by the HSE to the hospital to run its services is adequate in the first place. For the foregoing reasons this output measure has not been included in the model.

Amount of time that operating theatres are available

The amount of time that operating theatres are available for operations would be an important measure of the utilisation of theatre time. This information would be difficult to acquire for all of the acute hospitals in Ireland. However the number of patients treated in each hospital would be an indication of the overall hospital's efficiency which in turn would reflect the efficiency or otherwise of each sector of the hospital including the operating theatres. In any event as the information on operating theatres was not available it could not be included in the model.

The number of patients who return to hospital unexpectedly

The number of patients having to return to hospital unexpectedly is an important measure as it may be an indicator of the level of the service quality being provided by a hospital. If a hospital has a high unplanned re-admission rate then this may be an indication that the service that it provides is of poor quality. The unplanned re-admission rate could be used as a proxy for the overall quality of the service being provided by the hospital and could be used as such in the performance measurement model. However, the unplanned re-admission rate for hospitals in Ireland is not publicly available and thus could not be used in the model.

Inpatient waiting times

Inpatient waiting times was included by two groups as one of their top ten output measures. It was not deemed to be as relevant by former patients or hospital directors. The percentage that deemed this measure to be relevant ranged from 100% in the case of senior health service officials to 57.9% in the case of Hospital directors. It is surprising that this measure did not score more highly as any delays in receiving inpatient care can impact negatively on medical outcomes.

The combined score for all of the stakeholders was 77.3% but inpatient waiting times were rated as being more informative than relevant or necessary by the academic expert group. As well as that, the number of inpatients treated in a hospital should be indicative of its efficiency and should correlate with inpatient waiting times. As the number of inpatients treated was used as an output measure I did not use this measure in the model.

Outpatient waiting times

The combined results for all stakeholders' groups indicated that 70.5% of all groups deemed outpatient times to be a relevant measure. The academic experts generally concurred with that result but also deemed outpatient waiting times to be more relevant and informative than necessary to be included in the model. The numbers of patients treated is indicative of the efficiency of the hospital and thus the length of time that people are waiting. There should be a direct correlation between the length of time that people are waiting for their outpatient appointment and the numbers of new patients seen in the outpatients' clinic. I used the number of new outpatient attendances in the model because outpatient waiting lists generally include only those patients who have not yet been seen by a clinician. The new outpatient attendances were not adjusted for casemix in the model, as this information was not available.

The number of patients treated without having to stay in hospital overnight

Patients having procedures without having to stay in hospital overnight are called day cases. There has been an enormous increase in the number of patients being treated as day cases in Irish hospitals. The number of patients treated as day cases has increased by 128.5% between 1997 and 2006, as set out in table 2.6. This is obviously becoming more important as an output measure. It was deemed relevant by 69.9% of the combined stakeholders' groups but while it was deemed relevant by all members of the academic expert group it was not deemed to be necessary by any of them. However, given that only nine other output measures was scored higher by the academic expert group the number of day cases, as adjusted for casemix, was included in the model.

4.15 Performance Measurement Model

Taking into account all of the above, sensitivity analysis was carried out on ten performance measurement models. The maximum number of input and output measures that were used was six. Including more than six measures with a sample size of twenty eight hospitals resulted in most hospitals appearing to be uniquely efficient and thus on the efficient frontier. The maximum number of inputs and outputs in the model was therefore restricted to six. It is important to note that the sample is still relatively small at 28 hospitals, with 6 variables in nine of the models and four variables in one of the models. Given that the models were run under variable returns to scale, scale size comes in as a restriction resulting in a loss of some degrees of freedom and thus a loss of some discrimination on efficiency. The larger the number of input and output variables used in relation to the number of hospitals in the model, the more hospitals will be assigned as fully efficient and hence the less discriminating the DEA model will be. Banker et al. (1989) suggest as a rule of thumb that the number of DMUs should be at least three times the number of inputs plus outputs in any DEA application, although there is no analytic support for this rule (Pedraja-Chaparro et al., 1999). The model specifications are set out in table 4.11.

Casemix adjusted procedures have been used for both inpatients and day cases in the performance measurement models. Studies have shown that using only raw data with unweighted aggregation runs the risk of giving biased results to the research. While there is a wide usage of case mix adjusted data it generally only relates to inpatient admissions or discharges. Attendances at outpatient, day procedure or emergency departments are less often adjusted for case mix. This may be due to the non-availability of the data but even so it may lead to inaccurate results.

Model 1 focused on quantitative measures only. The inputs used were medical and nursing full-time equivalent staff numbers and total bed numbers. The outputs used were inpatient discharges (as adjusted for casemix), day case attendances (as adjusted for casemix) and new outpatient attendances. Casemix adjusted outpatient attendances were not available. The model was input orientated with variable returns to scale. It was run with the assumption that the objective was to minimise inputs for a given level of output. The model was also run under constant returns to scale in order to measure scale efficiencies. The aim of the model was to minimise inputs for a given level of output.

Model 2 again focused on quantitative measures only. The inputs used were medical salaries and nurse salaries instead of using full-time-equivalent staff numbers. All other input and output measures were the same as in Model 1. This model was run to test the robustness of the use of the labour inputs in Model 1. I wanted to ensure that there were no significant differences between using staff numbers instead of staff salaries in the model. The model was again input orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 4.11

Model Specifications

Model: Inputs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Doctor FTEs	*		*				*			*
Nurse FTEs	*		*							
Beds	*	*	*	*	*	*	*		*	*
Doctor salaries		*		*						
Nurse salaries		*		*						
Pay costs					*	*			*	
Non-pay costs					*	*			*	
Other staff FTEs							*			*
Model: Outputs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Inpatient discharges	*	*	*	*	*	*	*	*	*	*
New outpatient attendances	*	*	*	*	*	*	*			
Total outpatient attendances								*		
Day case attendances	*	*				*				
Survival rates			*	*	*		*		*	*
Infection free rate								*		
Hygiene rating								*	*	*

Model 3 included a qualitative measure of output for the first time in these performance measurement models. Of the qualitative measures that were deemed to be relevant only the mortality figures were available. These figures are not available to the general public and it was only through perseverance that they were acquired. Mortality data is publicly available by region but not by hospital. Other output measures, such as patient satisfaction measures and unplanned re-admission data were not available. Therefore medical full-time equivalent staff numbers, nurse full-time equivalent staff numbers and total bed numbers were included as inputs and inpatient discharges (as adjusted for casemix), new outpatient attendances and survival rates in each hospital were included as outputs. The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. The model was also run under constant returns to scale in order to measure scale efficiencies.

Model 4 included medical staff salaries, nursing staff salaries and total bed numbers as inputs and inpatient discharges (as adjusted for casemix), new outpatient attendances and survival rates in each hospital as outputs. Again this model was run to test the robustness of the labour inputs used in model 3. The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. The model was also run under constant returns to scale in order to measure scale efficiencies.

Model 5 looked at overall costs for each of the hospitals. However to ensure comparability between hospitals both pay and non-pay costs were adjusted in a number of areas. Firstly the pay costs in the Voluntary Hospitals were reduced in respect of superannuation expenditure incurred by them, as these costs were incurred centrally in respect of the statutory hospitals and were not reflected in their published expenditure. Similarly as much of the administration support for the statutory hospitals was provided centrally administration pay costs were omitted from all hospitals in the sample. Insurance costs incurred by the statutory hospitals were also paid centrally and as such were not reflected in their expenditure. The non-pay expenditure of the non-statutory hospitals was therefore reduced in respect of any insurance costs incurred.
The inputs were then pay costs (as adjusted for administration and superannuation costs), non-pay costs (as adjusted for insurance costs) and total bed numbers. The outputs were inpatient discharges (as adjusted for casemix), new outpatient attendances and survival rates in each hospital. The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. The model was also run under constant returns to scale in order to measure scale efficiencies.

Model 6 again looked at overall costs for each of the hospitals. However, on this occasion only quantitative measures were used. To ensure comparability between hospitals both pay and non-pay costs were adjusted as in Model 5. The inputs were therefore pay costs (as adjusted for administration and superannuation costs), non-pay costs (as adjusted for insurance costs) and total bed numbers. The outputs were inpatient discharges (as adjusted for casemix), new outpatient attendances and day patient attendances (as adjusted for casemix) in each hospital. The model was input orientated with variable returns to scale. It was run with the assumption that the objective was to minimise inputs for a given level of output. The model was also run under constant returns to scale in order to measure scale efficiencies.

Model 7 looked at overall staff numbers instead of overall staff salaries which were used in Model 6. To ensure comparability between hospitals administrative staff were not included in the numbers. A similar adjustment was made in models 5 and 6 in relation to administration pay costs. As much of the administrative support for the statutory hospitals is provided centrally and is not reflected in the hospitals' numbers it would be inaccurate to include administrative staff as an input in the model. The inputs were therefore medical staff full-time equivalent numbers, other full-time equivalent staff numbers (less administration staff) and total bed numbers. The outputs were inpatient discharges (as adjusted for casemix), new outpatient attendances and survival rates in each hospital. The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. The model was also run under constant returns to scale in order to measure scale efficiencies.

Model 8 looked at the non-clinical quality of care being offered to the patients attending the hospitals. As inputs the total number of inpatients treated and subsequently discharged, and the total number of outpatient attendances was used. The outputs used were the infection free rate and the hygiene rating for each of the hospitals. The infection free rate was derived from the HSE healthcare associated infection report on staphylococcus aureus bloodstream isolates in Ireland by acute public hospital and the hygiene rating for each hospital was available from the HSE national hygiene audit report. The model was output orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies. To ensure comparability between hospitals both pay and non-pay costs were adjusted as in Models 5 and 6.

Model 9 brought together the overall costs for each hospital as inputs along with quantitative, clinical and non-clinical qualitative outcomes as outputs. The inputs were therefore pay costs (as adjusted for administration and superannuation costs), non-pay costs (as adjusted for insurance costs) and total bed numbers. The outputs were inpatient discharges (as adjusted for casemix), survival rates and hygiene ratings in each hospital. The model was output orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies.

Model 10 looked at overall staff numbers as an input instead of overall staff salaries which was used in Model 9. To ensure comparability between hospitals administrative staff were not included in the numbers. A similar adjustment was made in model 7. As much of the administrative support for the statutory hospitals was provided centrally and was not reflected in the hospitals' numbers it was inaccurate to include administrative staff as an input in the model. The inputs were therefore medical staff full-time equivalent numbers, other full-time equivalent staff numbers (less administration staff) and total bed numbers. The outputs were inpatient discharges (as adjusted for casemix), survival rates and hygiene ratings in each hospital. The model was output orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies.

4.16 Conclusion

This chapter discussed the DEA technique and the methodology used to identify the most relevant input and output measures to be included in the performance measurement model. It concluded with the specification of the model that was used in the research.

In the next chapter the results from using the DEA model to measure the technical efficiency of each of the public acute hospitals in Ireland in 2007 are set out. In this process different combinations of inputs and outputs were used in order to test the sensitivity of the model to different specifications.

Chapter 5

Results and Discussion

5.1 Introduction

In chapter 4 the methodology used in the research was set out. In this chapter the performance measurement model will be developed and its sensitivity tested to include different combinations of inputs and outputs. For ease of use each hospital has been given its own code number as follows:

U1: Waterford Regional Hospital
U2: St. Luke's Hospital, Kilkenny
U3: Wexford General Hospital
U4: South Tipperary General Hospital
U5: Cork University Hospital
U6 : Kerry General Hospital
U7 : Mercy University Hospital
U8 : South Infirmary-Victoria University Hospital
U9 : Sligo General Hospital
U10 : Letterkenny General Hospital
U11 : Galway University Hospital
U12 : Mayo General Hospital
U13 : Portiuncula Hospital
U14 : Mid Western Regional Hospital
U15 : Our Lady of Lourdes Hospital Drogheda
U16 : Louth County Hospital
U17 : Cavan General Hospital
U18 : Our Lady's Hospital, Navan
U19 : Mater Hospital, Dublin
U20 : Beaumont Hospital, Dublin
U21 : Connolly Hospital, Blanchardstown
U22 : Adelaide & Meath National Children's Hospital
U23 : Longford Westmeath General Hospital, Mullingar
U24 : Tullamore General Hospital
U25 : Portlaoise General Hospital
U26 : St. James's Hospital Dublin
U27 : St. Vincent's University Hospital Dublin
U28 : St. Columcille's Loughlinstown

5.2 Models explained

The inputs and outputs used and the results from each of the ten DEA models have been set out in tables 5.1 to 5.40. Tables 5.41 and 5.42 provide a summary of the results from all of the models. The inputs and outputs of each model are firstly set out in a table showing the arithmetic mean and the standard deviation of each. These statistics give a general description of the span of resources and output sets of the hospital sample. This table is then followed by three further tables setting out the results for each of the models. The first of these tables has been divided into four columns. The first column indicates the code for the hospital, the second column sets out the technical efficiency of each hospital, the third column indicates the peer group to which each hospital belongs and the fourth column indicates the peer count for each of the hospitals.

If the technical efficiency of a hospital is equal to 1, then the hospital is deemed to be on the best practice frontier and thus relatively efficient. The lower the technical efficiency score the less efficient a hospital is deemed to be, for example, a score of 0.8967 would indicate an 89.67% level of technical efficiency while a score of 0.999 would indicate a 99.9% level of technical efficiency.

The peer group to which a hospital belongs highlights those hospitals that are potential role models for that hospital. These are hospitals that it could look to for ways of improving its operations.

The peer count indicates the number of times that each efficient hospital has been used as a reference hospital, i.e. a peer for others with a similar input/output mix. The more times that a hospital is a peer for other hospitals is an indicative measure of how truly efficient it is.

The second table that is used in each of the models highlights the potential savings or additional output that could be generated if inefficiency could be eliminated in each of the hospitals based on the technical efficiency scores generated in the first table.

The third table used shows the calculation of the scale efficiency for each of the hospitals. Scale efficiency indicates whether a hospital is operating at an optimal size or is either too big or too small. Scale efficiency is calculated by dividing the technical efficiency score for each hospital under constant returns to scale by its technical efficiency score under variable returns to scale. If a hospital's scale efficiency score is equal to 1, then it is deemed to be scale efficient, otherwise it is operating under increasing returns to scale which indicates that it is too small or under decreasing returns to scale which indicates that it is too big. If a hospital is operating under

increasing returns to scale, then a doubling of its inputs will lead to a more than doubling of its outputs. Similarly if a hospital is operating under decreasing returns to scale, then a doubling of its inputs will result in a less than doubling of its outputs.
It is important to note that given a relatively small sample of 28 hospitals, with 6 variables in nine of the models and four variables in one of the models, and given that the models were run under variable returns to scale that scale size comes in as a restriction resulting in a loss of some degrees of freedom and thus a loss of some discrimination on efficiency. The larger the number of input and output variables used in relation to the number of hospitals in the model, the more hospitals will be assigned as fully efficient and hence the less discriminating the DEA model will be. Banker et al. (1989) suggest as a rule of thumb that the number of DMUs should be at least three times the number of inputs plus outputs in any DEA application, although there is no analytic support for this rule (Pedraja-Chaparro et al., 1999). The Central Limit Theorem states tells us that a sampling distribution always has significantly less variability, as measured by standard deviation, than the population it's drawn from. Additionally, the sampling distribution will look more and more like normal distribution as the sample size is increased, even when the population is not normally distributed. If data follows a normal distribution we can be more confident that we can predict how data will behave. As a general rule a sample size of 30 or more is considered to be large enough for the Central Limit Theorem to take effect.

5.3 Model 1

Table 5.1 sets out the inputs and outputs in model 1 and the descriptive statistics of each of the variables

Table 5.1

Model 1 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Medical staff full-time-equivalents	178.31	121.64
Nursing staff full-time equivalents	606.71	380.36
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
Day patient attendances as adjusted for casemix	20239.54	15275.98
New outpatient attendances	21301.04	15762.35

These statistics give a general description of the span of resources and output sets of the hospital sample. The high standard deviations reflect the wide variety of hospital sizes included in the sample.

The results for model 1 are set out in Tables 5.2, 5.3 and 5.4. Table 5.2 shows the efficiency score for each hospital as well as their peer groups and peer counts.

Table 5.2

Model 1 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.969	7,16,19,20,24	0
U2	1	2,3,18	1
U3	1	3,16	4
U4	0.9451	2,16,18,19	0
U5	1	5,19	2
U6	0.9737	3,16,19	0
U7	1	7,16	2
U8	1	8,24	0
U9	1	9,19,24	0
U10	0.986	7,16,19,24	0
U11	0.9665	16,19,24	0
U12	1	12,16,24	0
U13	0.9748	8,16,19,28	0
U14	0.9743	16,19,24	0
U15	0.9085	5,19,28	0
U16	1	16	14
U17	0.9432	16,19,20,24	0
U18	1	18	5
U19	1	19,24	16
U20	1	20,24	3
U21	0.8967	18,19,28	0
U22	1	19,24	0
U23	0.9299	3,18,19	0
U24	1	24	12
U25	0.9605	16,18,19,28	0
U26	1	3,16,19	0
U27	0.9971	5,20,24	0
U28	1	28	4
Average	**0.979**		
%	**97.9%**		

This model was input orientated with variable returns-to-scale and focused totally on quantitative data. Table 5.2 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each hospital. It also gives the peer count for each hospital which indicates the number of times that each hospital appeared in the peer group of other hospitals. The results showed that fifteen of the twenty eight hospitals (53.6%) were technically efficient and were thus on the best practice frontier. It is important to note that the optimal solution can include what are termed slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier.
“For constraints with non-zero slacks, the performance of the peer group suggests that the decision making unit under scrutiny can improve beyond the level implied by the overall efficiency estimate.” Jacobs et al. (2006: 109)
The slacks for U2 are small at 0.3710 of nurses. This indicates that nursing numbers can be reduced by 0.3710 whilst maintaining the same level of output. Hospital U20 which is also deemed to be technically efficient has very small slacks of 0.0181 for doctors.
Similarly, U2 is deemed to be technically efficient while at the same time it has as peers U3 and U18. Each of these hospitals contributes to the construction of the virtual hospital for U2. Lambda is a vector describing the percentages other hospitals used to produce the virtual hospital. In this case the lambdas for the percentages U3 and U18 are respectively 0.4% and 0.2%. Similarly while U3 is also deemed to be on the efficient frontier it has U16 as its peer hospital with a lambda score of 0.1%. All of the hospitals deemed to be on the efficient frontiers have peers except for U16, U18, U24 and U28.

The average relative efficiency of all of the hospitals was 97.9%. This would appear to indicate a high level of technical efficiency overall. As already highlighted, it is important to note that given a relatively small sample of 28 hospitals with 6 variables and given that the model was run under variable returns to scale that scale size comes in as a restriction resulting in a loss of some degrees of freedom and thus a loss of some discrimination on efficiency. The larger the number of input and output variables used in relation to the number of hospitals in the model, the more hospitals will be assigned as fully efficient and hence the less discriminating the DEA model will be. The average relative efficiency of 97.9% also indicates that collectively all of the hospitals in the sample could produce their current output levels with a 2.1% reduction in the inputs included in the model. Whilst this is a relatively low percentage it would equate to a not inconsiderable saving in medical, nursing and bed resources. Table 5.3 sets out the potential savings to be made in each of the inefficient hospitals if they were to operate on the efficient production frontier. The total savings per annum would be 84.04 medical staff, 289.72 nurses and 175 beds.

The technical efficiency scores vary between the most inefficient hospital, U21, with a score of 89.67% and the efficient hospitals at 100%. This indicates that hospital U21 could produce its current level of output with a 10.33 % reduction in its inputs. This would result in an annual saving in staffing levels of 13.95 doctors and 53.41 nurses and a saving in beds of 26.

Table 5.3

Model 1 – potential savings

Hospital	Medical Staff Savings	Nursing Staff Savings	Bed Savings
U1	7.21	23.28	16
U2	0.00	0.00	0
U3	0.00	0.00	0
U4	4.71	18.89	11
U5	0.00	0.00	0
U6	2.78	12.71	8
U7	0.00	0.00	0
U8	0.00	0.00	0
U9	0.00	0.00	0
U10	2.17	7.90	6
U11	11.37	37.79	22
U12	0.00	0.00	0
U13	1.98	7.09	6
U14	5.67	19.88	12
U15	18.23	55.89	31
U16	0.00	0.00	0
U17	5.13	18.67	14
U18	0.00	0.00	0
U19	0.00	0.00	0
U20	0.00	0.00	0
U21	13.95	53.41	26
U22	0.00	0.00	0
U23	7.16	21.46	15
U24	0.00	0.00	0
U25	2.80	9.96	6
U26	0.00	0.00	0
U27	0.89	2.78	2
U28	0.00	0.00	0
Totals	**84.04**	**289.72**	**175**

Table 5.2 also indicates the number of times that each efficient hospital has been used as a reference hospital, i.e. a peer, for itself as well as for others with a similar input-output mix. This facilitates comparisons to be made between those hospitals in relation to their characteristics, operating procedures and other attributes.
For example hospital U13 would have as peers, hospitals U8, U16, U19 and U28. Each of these hospitals contributes to the construction of the virtual hospital for U13. Lambda is a vector describing the percentages other producers used to construct the virtual producer. In the case of U13 Lambda for the percentages of U8, U16, U19 and U28 used are respectively 36.50%, 50.45%, 7.12% and 5.93%.
Overall fifteen hospitals achieved an efficiency score of 100%. However, many of these hospitals did not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U16, U19 and U24 are truly efficient because they are respectively peers for fourteen, sixteen and twelve other hospitals in the sample. Hospitals U2, U3, U5, U7,U18, U20 and U28 whilst each achieving an efficiency score of 100% are peers for between one and five other hospitals in the sample and thus may also have scope to improve their efficiency.

Scale efficiency measures the impact of scale size on the productivity of a decision making unit. Scale efficiency tests indicate whether a hospital is operating at activity levels that are either contributing to higher than minimum average costs or at its most productive scale size.
Table 5.4 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. In order to calculate the technical efficiency of the hospitals in the sample, model 1 was run using constant returns to scale and with an input orientation. It was clear from this table that some hospitals appeared to be operating at too large a scale to maximise the productivity of their inputs, i.e. with decreasing returns to scale (DRS), and that others appeared to be too small and were exhibiting higher than average costs, i.e. with increasing returns to scale (IRS). The PIM DEAsoft-V1 software used indicated whether the hospitals were showing decreasing or increasing returns to scale. However:

> *A production correspondence is said to exhibit increasing returns to scale (IRS) if a radial increase in input levels (i.e. keeping input mix constant) leads under Pareto-efficiency to a more than proportionate radial increase in output levels; if the radial increase in output levels is less than proportionate we have decreasing returns to scale (DRS) and otherwise we have constant returns to scale (CRS).*
>
> (Thanassoulis, 2001: 124)

With an output orientation a decision making unit is Pareto-efficient if it is not possible to raise any one of its output levels without lowering at least another one of its output levels and/or without increasing at least one of its input levels. With an input orientation a decision making unit is Pareto-efficient if it is not possible to lower any one of its input levels without increasing at least another one of its input levels and/or without lowering at least one of its output levels (Thanassoulis, 2001).
Jacobs et al., 2006 refer to Coelli et al., 1998, stating that in order to obtain an indication whether a decision making unit is operating in the area of increasing, or the area of decreasing, returns to scale, a non-increasing returns to scale (NIRS) constraint can be added by altering the convexity constraint in the BCC model to:

$$\sum_{j=1}^{n} \lambda_j \leq 1$$

Scale inefficiencies can then be determined, as to whether there is increasing or decreasing returns to scale, by comparing the decision making units technical efficiency score under the BCC model to their technical efficiency score under the NIRS constraint. If they are not equal, increasing returns to scale exist; if they are equal then decreasing returns to scale apply.

The average scale efficiency score of all of the hospitals in the model was 97.6%. The results showed that 10 (35.7%) out of 28 hospitals were operating at optimal plant size, though many others were operating very close to optimal size. The pattern of scale inefficiency indicated that 9 (32.1%) hospitals were operating on increasing returns to scale and were therefore too small. A further 9 (32.1%) hospitals were operating on decreasing returns to scale and were therefore too big. It is important to note that given a relatively small sample of 28 hospitals with 6 variables and given that the models were run under variable returns to scale that scale size comes in as a restriction resulting in a loss of some degrees of freedom and thus a loss of some discrimination on efficiency. The larger the number of input and output variables used in relation to the number of hospitals in the model, the more hospitals will be assigned as fully efficient and hence the less discriminating the DEA model will be. Given the closeness of some of the scale efficiency scores caution should be used when interpreting the results.

Table 5.4

Model 1 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.9315	0.969	0.9613	DRS	Too big
U2	0.994	1	0.994	IRS	Too small
U3	1	1	1	CRS	
U4	0.9035	0.9451	0.955983	IRS	Too small
U5	1	1	1	CRS	
U6	0.936	0.9737	0.961282	DRS	Too big
U7	1	1	1	CRS	
U8	1	1	1	CRS	
U9	0.9905	1	0.9905	IRS	Too small
U10	0.9112	0.986	0.924138	DRS	Too big
U11	0.9325	0.9665	0.964822	DRS	Too big
U12	1	1	1	CRS	
U13	0.9528	0.9748	0.977431	IRS	Too small
U14	0.932	0.9743	0.956584	DRS	Too big
U15	0.8698	0.9085	0.957402	IRS	Too small
U16	1	1	1	CRS	
U17	0.9431	0.9432	0.999894	IRS	Too small
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	1	1	1	CRS	
U21	0.8388	0.8967	0.93543	IRS	Too small
U22	0.9906	1	0.9906	DRS	Too big
U23	0.9284	0.9299	0.998387	DRS	Too big
U24	1	1	1	CRS	
U25	0.8506	0.9605	0.88558	IRS	Too small
U26	0.9542	1	0.9542	DRS	Too big
U27	0.9911	0.9971	0.993983	DRS	Too big
U28	0.9286	1	0.9286	IRS	Too small
Average	**0.9564**	**0.979**	**0.976**		
%	**95.64%**	**97.9%**	**97.6%**		

5.4 Model 2

In order to test the robustness of the model to changes in the measurement of inputs model 1 was run using salary expenditure instead of full-time equivalent staff numbers.

Table 5.5 sets out the inputs and outputs in model 2 and the descriptive statistics of each of the variables.

Table 5.5

Model 2 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Medical staff salaries	24807.75	15876.04
Nursing staff salaries	35812.86	23055.7
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
Day patient attendances as adjusted for casemix	20239.54	15275.98
New outpatient attendances	21301.04	15762.35

Again as with model 1 the high standard deviations reflect the wide variation in hospital sizes included in the sample.

The results for model 2 are presented in Tables 5.6, 5.7, and 5.8.

Table 5.6

Model 2 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	1	1,11,7,12	0
U2	1	2,3,16	3
U3	1	3,16	9
U4	0.9454	2,19,28	0
U5	1	19,20,9	0
U6	0.9618	3,16,19	0
U7	1	7	6
U8	1	8,24	1
U9	1	9,19,24,	1
U10	1	3,7,12	0
U11	1	3,11,12,19,	3
U12	1	12,24	6
U13	0.9692	3,8,16,19,28	0
U14	1	7,11,12,19,24	0
U15	0.9342	3,18,19	0
U16	1	16	6
U17	0.9497	12,16,18,24,28	0
U18	1	18	5
U19	1	19,24	12
U20	1	20,24	2
U21	0.9069	2,19,28	0
U22	1	19,22,24	0
U23	0.9802	3,7,12,16,18	0
U24	1	24	10
U25	0.9274	3,18,19,28	0
U26	1	3,7,11,19	0
U27	1	20,24,27	0
U28	1	28	5
Average	**0.985**		
%	**98.5%**		

This model was input orientated with variable returns-to-scale and focused totally on quantitative data. Table 5.6 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each hospital. It also gives the peer count for each hospital which indicates the number of times that each hospital appears in the peer group of other hospitals. It is important to note that the optimal solution can include what are termed slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. U1 which is deemed to be on the efficient frontier has slacks of €31,642 in medical pay. U1 has the potential to save this money while at the same time maintaining the level of its outputs. Similarly, the other hospitals U5, U9, U10, U11, U14, U20, U26 and U27 that were deemed to be on the efficient frontier have slacks and are thus technically inefficient.

The results showed that twenty of the twenty eight hospitals (71.4%) were efficient and were thus on the best practice frontier. The average relative efficiency of all of the hospitals was 98.5%. This would appear to indicate a high level of technical efficiency overall. It also indicated that collectively all of the hospitals in the sample could produce their current output levels with a 1.5% reduction in the inputs included in the model. Whilst this is a relatively low percentage it would equate to a not inconsiderable saving in medical, nursing and bed resources. Table 5.7 sets out the potential savings to be made in each of the inefficient hospitals if they were to operate on the efficient production frontier. The total savings that could be achieved would be €6,893,000 in medical pay and €10,464,000 in nursing pay, as well as a reduction in bed numbers of 103.

The technical efficiency scores varied between the most inefficient hospital, U21, with a score of 90.69% and the efficient hospitals at 100%. This indicated that hospital U21 could produce its current level of output with a 9.31 % reduction in its inputs. Financially this would result in a saving of €4.498 million per annum and a reduction in capital costs of 23 beds.

Hospitals with emergency departments will have more difficulties in achieving efficiency. This however does not mean that they can never be efficient. Largely due to the prospect of any type of emergency arriving at the hospital they must be prepared to accept all types of patients. This may involve having excess capacity available at times but when emergency department attendances are analysed peaks and troughs in activity levels can be projected with a reasonable degree of certainty. Proper planning can ensure a high level of efficiency. Only hospitals with emergency departments have been included in this research in order to ensure that only those hospitals which produce a similar mix of outputs given input levels are compared. The results from this model show that large hospitals with large emergency departments such as U26 with 842 beds and U20 with 741 beds are deemed to be on the efficient frontier along with small hospitals such as U28 and U16 with 118 beds and 136 beds respectively.

Table 5.7

Model 2 – potential savings

Hospital	Medical Savings € '000s	Nursing Savings € '000s	Beds Savings
U1	0	0	0
U2	0	0	0
U3	0	0	0
U4	693	1128	11
U5	0	0	0
U6	588	1028	12
U7	0	0	0
U8	0	0	0
U9	0	0	0
U10	0	0	0
U11	0	0	0
U12	0	0	0
U13	350	534	7
U14	0	0	0
U15	1731	2407	22
U16	0	0	0
U17	782	1000	12
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	1582	2916	23
U22	0	0	0
U23	302	345	4
U24	0	0	0
U25	867	1106	12
U26	0	0	0
U27	0	0	0
U28	0	0	0
Total	**6893**	**10464**	**103**

Table 5.6 also indicates the number of times that each efficient hospital has been used as a reference hospital, i.e. a peer, for itself as well as for others with a similar input-output mix. This facilitates comparisons to be made between those hospitals in relation to their characteristics, operating procedures and other attributes.
For example hospital U23 would have as peers, hospitals U3, U7, U12, U16 andU18. Each of these hospitals contributes to the construction of the virtual hospital for U23. Lambda is a vector describing the percentages other producers used to construct the virtual producer. In the case of U23 Lambda for the percentages of U3, U7, U12, U16 and U18 used are respectively 35.1%, 2.49%, 7.76%, 2.75% and 51.9%.
Overall twenty hospitals achieved an efficiency score of 100%. However, many of these hospitals did not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U3, U19 and U24 are truly efficient because they are peers for nine or more other hospitals in the sample. Hospitals U2, U7, U8, U9, U11, U12, U16, U18, U20 and U28 each are peers for between one and six other hospitals and while they each achieved an efficiency score of 100% there may be scope for them to improve their efficiency.

Scale efficiency measures the impact of scale size on the productivity of a decision making unit. Scale efficiency tests indicate whether a hospital is operating at activity levels that are either contributing to higher than minimum average costs or at its most productive scale size.
Table 5.8 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. In order to calculate the technical efficiency of the hospitals in the sample model 2 was run using constant returns to scale and with an input orientation. It is clear from this table that some hospitals appeared to be operating at too large a scale to maximise the productivity of their inputs, i.e. with decreasing returns to scale, and that others appeared to be too small and were exhibiting higher than average costs. The average scale efficiency score of all of the hospitals was 97.9%. The results showed that 13 (46.4%) out of 28 hospitals were operating at optimal plant size, though many others were operating very close to optimal size. The pattern of scale inefficiency indicated that 11 (39.3%) hospitals were operating on increasing returns to scale and were therefore too small. A further 4 (14.3%) hospitals were operating on decreasing returns to scale and were therefore too big.

Table 5.8

Model 2 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	1	1	1	CRS	
U2	0.9842	1	0.9842	IRS	Too small
U3	1	1	1	CRS	
U4	0.8804	0.9454	0.931246	IRS	Too small
U5	1	1	1	CRS	
U6	0.9415	0.9618	0.978894	DRS	Too big
U7	1	1	1	CRS	
U8	1	1	1	CRS	
U9	0.9905	1	0.9905	IRS	Too small
U10	0.997	1	0.997	DRS	Too big
U11	1	1	1	CRS	
U12	1	1	1	CRS	
U13	0.9117	0.9692	0.940673	IRS	Too small
U14	0.9693	1	0.9693	DRS	Too big
U15	0.9247	0.9342	0.989831	IRS	Too small
U16	1	1	1	CRS	
U17	0.9134	0.9497	0.961777	IRS	Too small
U18	0.9689	1	0.9689	IRS	Too small
U19	1	1	1	CRS	
U20	1	1	1	CRS	
U21	0.8347	0.9069	0.920388	IRS	Too small
U22	1	1	1	CRS	
U23	0.961	0.9802	0.980412	IRS	Too small
U24	1	1	1	CRS	
U25	0.8627	0.9274	0.930235	IRS	Too small
U26	0.9656	1	0.9656	DRS	Too big
U27	1	1	1	CRS	
U28	0.8897	1	0.8897	IRS	Too small
Average	**0.964118**	**0.985**	**0.979**		
%	**96.41%**	**98.5%**	**97.9%**		

Changing the way that labour was measured only had a minor impact on the overall model results. However, using salaries as an input instead of full-time equivalents had a significant impact on some hospitals. All of the hospitals in model 1 that were deemed to be on the efficient frontier were also deemed to be on the efficient frontier in model 2. However whilst model 1 had fifteen efficient hospitals model 2 had twenty. The additional hospitals were U1, U10, U11, U14 and U27. Therefore using salaries instead of full-time equivalents as an input resulted in more hospitals appearing to be relatively efficient. What does this tell us? One explanation could be that using full-time equivalents as an input measure was hiding other payroll costs that did not impact on full-time equivalent numbers such as overtime or premium payments and that some hospitals maintained lower full-time equivalent numbers by incurring higher costs in these areas. All of these costs would have been picked up when using salaries as an input measure, thus ensuring a more accurate measurement of labour input in the model. A second explanation could be the higher use by some hospitals of agency staff that would not have been reflected in the full-time equivalent numbers. This again could have distorted the labour input. A third explanation could be that the age profile of staff disproportionately impacted on the salary levels in some hospitals. This could occur in a hospital with an older age profile and with more experienced staff on higher salaries. In this situation using full-time equivalents instead of salaries as an input would be a more accurate measure of labour efficiency in the model.

Of the additional five hospitals deemed to be relatively efficient in model 2 only hospital U27 was close to relative efficiency at 99.71% in model 1. The other hospitals: U1 at 96.9%, U10 at 98.6%, U11 at 96.65% and U14 at 97.43% were not as close to relative efficiency and thus the reasons for their improvement would need further investigation.

The average efficiency scores of both models were similar at 97.9% and 98.5% respectively and the average scale efficiency scores were also extremely close at 97.6% and 97.9% respectively. These results would appear to indicate that overall model 1 was robust in its labour specification.

5.5 Model 3

Table 5.9 sets out the inputs and outputs in model 3 and the descriptive statistics of each of the variables. In this model a qualitative measure was included for the first time. Of the qualitative measures that were deemed to be relevant only the mortality figures were available for use in the model.

Table 5.9

Model 3 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Medical staff full-time-equivalents	178.31	121.64
Nursing staff full-time equivalents	606.71	380.36
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
New outpatient attendances	21301.04	15762.35
Patient survival rate	97.61	1.06

The results for model 3 are set out in tables 5.10, 5.11 and 5.12.

Table 5.10

Model 3 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.995	18,19	0
U2	1	2,3,18	1
U3	1	3,16	10
U4	0.9954	16,18,19,28	0
U5	1	5,19	2
U6	0.9876	2,3,16,19	0
U7	0.9874	18,19	0
U8	1	8,16	6
U9	0.9979	3,8,16,19	0
U10	1	3,16,19	0
U11	1	3,19	0
U12	1	3,19	0
U13	1	8,16,19,24,28	0
U14	0.992	3,19	0
U15	1	5,18,19,28	0
U16	1	16	11
U17	0.9989	8,16,19	0
U18	1	16,18	10
U19	1	19	20
U20	1	18,19	0
U21	0.9757	5,18,19,28	0
U22	1	8,19	0
U23	1	3,18,19	0
U24	1	8,16,24	1
U25	1	16,18,19,28	0
U26	1	3,19	0
U27	0.9735	3,18,19	0
U28	1	28	6
Average	**0.99655**		
%	**99.66%**		

The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. Table 5.10 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each hospital. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. Table 5.10 also gives the peer count for each hospital which indicates the number of times that each hospital appears in the peer group of other hospitals. The results showed that nineteen of the twenty eight hospitals (68%) were efficient and were thus on the best practice frontier. The average relative efficiency of all of the hospitals was 99.66%. This would appear to indicate a very high level of technical efficiency overall. It also indicated that collectively all of the hospitals in the sample could increase their current output levels by 0.34% without increasing their inputs. Overall nineteen hospitals achieved an efficiency score of 100%. However, many of these hospitals did not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospital U19 is truly efficient because it was a peer for twenty other hospitals in the sample. Hospitals U3, U16 and U18 each were respectively peers for eleven, ten and ten other hospitals and while they each achieved an efficiency score of 100% there may be scope for them to improve their efficiency. Likewise hospitals U2, U8, U5, U24 and U28, whilst each achieving an efficiency score of 100% were respectively peers for one, two, six, one and six other hospitals in the sample and thus may also have scope to improve their efficiency.

Table 5.11 details the potential for each hospital to increase their outputs without increasing their inputs. As we can see 1615 additional inpatients and 1637 additional new outpatients could be treated without increasing any of the inputs. Similarly there would appear to be the potential to increase survival rates in a number of hospitals by from 0.11% in U17 to 2.54% U27.

Table 5.11

Model 3 – potential to increase outputs

Hospital	Additional Patients	Additional New Outpatients	Increased Survival Rate
U1	106	119	0.0049
U2	0	0	0
U3	0	0	0
U4	42	40	0.0042
U5	0	0	0
U6	152	175	0.0123
U7	152	84	0.0122
U8	0	0	0
U9	31	51	0.0030
U10	0	0	0
U11	0	0	0
U12	0	0	0
U13	0	0	0
U14	160	213	0.0078
U15	0	0	0
U16	0	0	0
U17	10	16	0.0011
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	272	308	0.0234
U22	0	0	0
U23	0	0	0
U24	0	0	0
U25	0	0	0
U26	0	0	0
U27	690	631	0.0254
U28	0	0	0
Total	**1615**	**1637**	

Table 5.12 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. In order to calculate the technical efficiency of the hospitals in the sample model 3 was also run using constant returns to scale and with an output orientation. It is clear from this table that some hospitals appeared to be operating at too large a scale to maximise the productivity of their inputs, i.e. with decreasing returns to scale, resulting in higher than average costs. The average scale efficiency score of all of the hospitals was 94.25%. The results showed that 9 (32%) out of 28 hospitals were operating at optimal plant size, though many others were operating very close to optimal size. The pattern of scale inefficiency indicated that 19 (68%) hospitals were operating on decreasing returns to scale and were therefore too big. There were no hospitals operating on increasing returns to scale and that were thus too small.

Table 5.12

Model 3 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns To scale	
U1	0.8654	0.995	0.869749	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9464	0.9954	0.950774	DRS	Too big
U5	1	1	1	CRS	
U6	0.8727	0.9876	0.883657	DRS	Too big
U7	0.8945	0.9874	0.905915	DRS	Too big
U8	1	1	1	CRS	
U9	0.8005	0.9979	0.802185	DRS	Too big
U10	0.8619	1	0.8619	DRS	Too big
U11	0.9031	1	0.9031	DRS	Too big
U12	0.9988	1	0.9988	DRS	Too big
U13	0.9776	1	0.9776	DRS	Too big
U14	0.8339	0.992	0.840625	DRS	Too big
U15	0.9115	1	0.9115	DRS	Too big
U16	1	1	1	CRS	
U17	0.9275	0.9989	0.928521	DRS	Too big
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	0.9687	1	0.9687	DRS	Too big
U21	0.8971	0.9757	0.919442	DRS	Too big
U22	0.9283	1	0.9283	DRS	Too big
U23	0.9284	1	0.9284	DRS	Too big
U24	1	1	1	CRS	
U25	0.9672	1	0.9672	DRS	Too big
U26	0.9455	1	0.9455	DRS	Too big
U27	0.8734	0.9735	0.897175	DRS	Too big
U28	1	1	1	CRS	
Average	**0.9394**	**0.9966**	**0.9425**		
%	**93.94%**	**99.66%**	**94.25%**		

5.6 Model 4

In order to test the robustness of the model to changes in the measurement of inputs model 3 was run again as model 4 with salary expenditure instead of full-time equivalent staff. The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input.

Table 5.13 sets out the inputs and outputs in model 4 and the descriptive statistics of each of the variables.

Table 5.13

Model 4 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Medical staff salaries	24807.75	15876.04
Nursing staff salaries	35812.86	23055.7
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
New outpatient attendances	21301.04	15762.35
Patient survival rate	97.61	1.06

The results for Model 4 are set out in tables 5.14, 5.15 and 5.16.

Table 5.14

Model 4 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.9966	3,19	0
U2	1	2,3,28	3
U3	1	3,16	17
U4	0.9957	2,19,28	0
U5	1	5,19	0
U6	0.9874	2,3,16,19	0
U7	0.9875	3	0
U8	1	8,16	6
U9	0.997	3,12,19,28	0
U10	1	3,8	0
U11	1	3,12,19	0
U12	1	3,8,12,19	4
U13	1	3,8,16,19,28	0
U14	0.9926	3,12,19	0
U15	1	3,18,19	0
U16	1	16	5
U17	0.997	3,8,12,28	0
U18	1	18	4
U19	1	19,28	16
U20	0.9864	3,19	0
U21	0.977	2,19,28	0
U22	1	8,19,22	1
U23	1	3,16,18,28	0
U24	1	8,22,28	0
U25	1	3,18,19,28	0
U26	1	3,19	0
U27	0.9742	3,18,19	0
U28	1	28	10
Average	**0.996121**		
%	**99.61%**		

Table 5.14 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each hospital. It also gives the peer count for each hospital which indicates the number of times that each hospital appears in the peer group of other hospitals. The results showed that eighteen of the twenty eight hospitals (64.3%) were relatively efficient and were thus on the best practice frontier. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency of all of the hospitals was 99.61%. This would appear to indicate a very high level of technical efficiency overall. It also indicated that collectively all of the hospitals in the sample could increase their current output levels by 0.39% without increasing their inputs. Overall eighteen hospitals achieved an efficiency score of 100%. However, many of these hospitals did not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U3, U19 and U28 were truly efficient because they were respectively peers for seventeen, sixteen and ten other hospitals in the sample. Hospitals U2, U8, U12, U16, U18 and U22 each were respectively peers for three, six, four, five, four and one other hospitals and while they each achieved an efficiency score of 100% there may be scope for them to improve their efficiency.

Table 5.15 details the potential for each hospital to increase their outputs without increasing their inputs. As we can see those hospitals that were not on the efficient frontier could potentially increase their annual throughput by 2,048 inpatients and 1,957 new outpatients. Similarly they could also potentially increase the survival rate for their patients without having to increase their inputs. This potential increase would range between 0.3% in U1 and 2.47% in U27.

Table 5.15

Model 4 – potential to increase outputs

Hospital	Additional Inpatients	Additional New Outpatients	Increased Survival Rate
U1	72	81	0.0033
U2	0	0	0
U3	0	0	0
U4	39	37	0.0042
U5	0	0	0
U6	154	177	0.0123
U7	151	84	0.0122
U8	0	0	0
U9	44	72	0.0030
U10	0	0	0
U11	0	0	0
U12	0	0	0
U13	0	0	0
U14	148	197	0.0072
U15	0	0	0
U16	0	0	0
U17	28	43	0.0029
U18	0	0	0
U19	0	0	0
U20	483	361	0.0131
U21	258	291	0.0222
U22	0	0	0
U23	0	0	0
U24	0	0	0
U25	0	0	0
U26	0	0	0
U27	671	614	0.0247
U28	0	0	0
Total	**2048**	**1957**	

Table 5.16 presents the technical efficiency for each of the hospitals in the sample under constant and variable returns to scale and calculates the scale efficiency for each hospital. In order to calculate the scale efficiency of the hospitals in the sample model 4 was run using constant returns to scale and with an output orientation. It is clear from this table that some hospitals appeared to be operating at too large a scale to maximise the productivity of their inputs, i.e. with decreasing returns to scale, and one hospital operating at too small a scale, i.e. with increasing returns to scale. The average scale efficiency score of all of the hospitals was 96.0%. The results showed that 10 (35.7%) out of 28 hospitals were operating at optimal plant size, though many others were operating at close to optimal size. The pattern of scale inefficiency indicated that 18 (64.3%) hospitals were operating on decreasing returns to scale and were therefore too big whilst there were no hospitals operating on increasing returns to scale and were therefore too small.

Table 5.16

Model 4 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.9182	0.9966	0.921333	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9478	0.9957	0.951893	DRS	Too big
U5	1	1	1	CRS	
U6	0.872	0.9874	0.883127	DRS	Too big
U7	0.9253	0.9875	0.937013	DRS	Too big
U8	1	1	1	CRS	
U9	0.909	0.997	0.911735	DRS	Too big
U10	0.9403	1	0.9403	DRS	Too big
U11	0.9771	1	0.9771	DRS	Too big
U12	1	1	1	CRS	
U13	0.9719	1	0.9719	DRS	Too big
U14	0.8858	0.9926	0.892404	DRS	Too big
U15	0.9349	1	0.9349	DRS	Too big
U16	1	1	1	CRS	
U17	0.9339	0.997	0.93671	DRS	Too big
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	0.9439	0.9864	0.956914	DRS	Too big
U21	0.9075	0.977	0.928864	DRS	Too big
U22	1	1	1	CRS	
U23	0.9673	1	0.9673	DRS	Too big
U24	0.9961	1	0.9961	DRS	Too big
U25	0.9311	1	0.9311	DRS	Too big
U26	0.9583	1	0.9583	DRS	Too big
U27	0.8599	0.9742	0.882673	DRS	Too big
U28	1	1	1	CRS	
Average	**0.956439**	**0.996121429**	**0.959988**		
%	**95.64%**	**99.61%**	**96.00%**		

Interestingly all of the hospitals deemed to be on the efficient frontier in this model were also deemed to be efficient in model 3. The only difference between the results from both models was that hospital U20 was also deemed to be on the efficient frontier in model 3 while it received an efficiency score of 98.64% in this model. Therefore using full-time equivalents instead of salaries as an input resulted in one more hospital appearing to be relatively efficient. This is the opposite of what happened in model 2 where using salaries as an input instead of full-time equivalents resulted in more hospitals appearing to be relatively efficient. There could be a number of explanations for hospital U20 appearing to be relatively efficient on this occasion. One explanation could be that using full-time equivalents as an input measure was hiding other payroll costs that did not impact on full-time equivalent numbers such as overtime or premium payments and that U20 maintained lower full-time equivalent numbers by incurring higher costs in these areas. These costs may not have been picked up when using full-time equivalents as an input measure, thus ensuring a lower cost of labour input in the model. A second explanation could be the higher use of agency staff that would not have been reflected in the full-time equivalent numbers. This again could have distorted the labour input. A third explanation could be that the age profile of staff disproportionately impacted on the salary levels in U20. This could have occurred if U20 had an older age staff profile and with more experienced staff on higher salaries. In this situation using full-time equivalents instead of salaries as an input would be a more accurate measure of labour efficiency in the model.

Another difference between models 3 and 4 was that the potential for increasing both inpatient numbers and new outpatients was greater in model 4. Given that model 4 had one less hospital on the efficient frontier and thus one more hospital with the potential to increase efficiency, this is not surprising. The hospitals with the greatest potential to increase patient output were U21 and U27 in model 3 with increases of 272 and 690 respectively, and U20, U21 and U27 in model 4 with increases respectively of 483, 258 and 671. Both models agreed that U27 had the greatest potential to increase patient output. Similarly the hospitals with the greatest potential to increase the number of new outpatients in model 3 were again U21 and U27 and in model 4 were U20, U21 and U27, with U27 showing the greatest potential. The same hospitals that showed the greatest potential to increase patient numbers and new outpatients also showed the greatest potential to increase survival rates in both models.

The average efficiency scores for both models were largely the same at 99.66% and 99.61% respectively. This would appear to indicate that model 3 was robust in its labour specifications. However, the average scale efficiency scores were not as close at 94.25% and 96.0% respectively. Model 4 had ten hospitals that were deemed to be operating at optimal size while Model 3 had nine. Both models indicated that hospitals U2, U3, U5, U8, U16, U18, U19 and U28 were of optimal size but where they differed was that model 3 also indicated that U24 was of optimal size while model 4 indicated that U12 and U22 were of optimal size.

5.7 Model 5

In Model 5, pay costs and non pay costs for each of the hospitals were included as inputs. However to ensure comparability between hospitals both pay and non-pay costs were adjusted in a number of areas. Pay costs were reduced in the Voluntary Hospitals in respect of superannuation expenditure incurred by them to ensure comparability with the statutory hospitals, as these costs are incurred centrally in respect of the statutory hospitals and are not reflected in their published expenditure. Similarly as much of the administration support for the statutory hospitals is provided centrally administration pay costs were omitted from all hospitals in the sample. Insurance costs incurred by the statutory hospitals are also paid centrally and as such are not reflected in their expenditure. The non-pay expenditure of the Voluntary hospitals was therefore reduced in respect of any insurance costs incurred.
The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 5.17 sets out the inputs and outputs in Model 5 and the descriptive statistics of each of the variables.

Table 5.17

Model 5 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Pay Costs as adjusted for superannuation and admin.	85025.46	55018.95
Non-Pay Costs as adjusted for insurance costs	42724.46	36964.79
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
New outpatient attendances	21301.04	15762.35
Patient survival rate	97.61	1.06

The results for Model 5 are set out in tables 5.18, 5.19 and 5.20.

Table 5.18

Model 5 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.9948	3,19	0
U2	1	2,3,28	7
U3	1	3	14
U4	0.9964	2,3,25	0
U5	1	5,19	2
U6	0.9865	3,18,19	0
U7	0.9871	3,18,19	0
U8	1	8	5
U9	0.996	2,3,13,19	0
U10	0.9995	3,8,19	0
U11	1	3,19	0
U12	1	3,8.19	0
U13	1	3,8,13,28	2
U14	0.9911	3,19	0
U15	1	2,5,19	0
U16	1	16	0
U17	0.9966	2,3,13,19,28	0
U18	1	18	4
U19	1	19,28	18
U20	1	2,3,19	0
U21	0.9757	2,5,19,28	0
U22	1	8,19,22	0
U23	1	2,18,19,28	0
U24	0.9952	8,19,28	0
U25	1	25	1
U26	1	3,19	0
U27	0.972	18,19,28	0
U28	1	28	8
Average	**0.996104**		
%	**99.61%**		

Table 5.18 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each hospital. It also gives the peer count for each hospital which indicates the number of times that each appears in the peer group of other hospitals. The results showed that seventeen of the twenty eight hospitals (60.7%) were efficient and thus on the best practice frontier. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency of all of the hospitals was 99.61%. This would appear to indicate a very high level of technical efficiency overall. It also indicates that collectively all of the hospitals in the sample could increase their current output by 0.39% without increasing their inputs. Overall seventeen hospitals achieved an efficiency score of 100%. However many of these hospitals do not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U19 and U3 are truly efficient because they are respectively peers for eighteen and fourteen other hospitals in the sample. Hospitals U2 and U28 are respectively peers for seven and eight other hospitals and while they both achieved an efficiency score of 100% there may be scope for them to improve their efficiency. Likewise hospitals U5, U8, U13, U18 and U25 are peers for two, five, two, four and one other hospitals and while they each achieved an efficiency score of 100% there may be scope for them to improve their efficiency.

Table 5.19 details the potential for each hospital to increase their outputs without increasing their inputs. As we can see those hospitals that are not on the efficient frontier could potentially increase their annual throughput by 1,788 inpatients and 1,890 new outpatients. Similarly they could also potentially increase the survival rate for their patients without having to increase their inputs by between 0.05% in U10 to 2.68% in U27. The potential to increase outputs is particularly large in hospital U14. This hospital has the potential to annually increase inpatients treated by 1,780 and new outpatients treated by 237 as well as increasing the survival rate for its patients by 0.87%. Hospitals U21 and U27 have the potential to increase respectively inpatients treated by 272 and 729, new outpatients by 308 and 666 and the survival rate by 2.34% and 2.68%.

Table 5.19

Model 5 – potential to increase outputs

Hospital	Additional Inpatients	Additional New Outpatients	Increased Survival Rate
U1	110	123	0.0051
U2	0	0	0
U3	0	0	0
U4	33	31	0.0035
U5	0	0	0
U6	165	190	0.0131
U7	156	86	0.0126
U8	0	0	0
U9	59	96	0.0039
U10	8	10	0.0005
U11	0	0	0
U12	0	0	0
U13	0	0	0
U14	178	237	0.0087
U15	0	0	0
U16	0	0	0
U17	32	48	0.0033
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	272	308	0.0234
U22	0	0	0
U23	0	0	0
U24	46	95	0.0047
U25	0	0	0
U26	0	0	0
U27	729	666	0.0268
U28	0	0	0
Total	**1788**	**1890**	

Table 5.20 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. It is clear from this table that some hospitals appear to be operating at too large a scale to maximise the productivity of their inputs, i.e. with decreasing returns to scale, resulting in higher than average costs. The average scale efficiency score of all of the hospitals was 94.78%. The results showed that eleven (39.3%) out of twenty eight hospitals were operating at optimal plant size, though many others were operating very close to optimal size. The pattern of scale inefficiency indicated that seventeen (60.7%) were operating on decreasing returns to scale and were therefore too big. There were no hospitals operating on increasing returns to scale and that were thus too small.

Table 5.20

Model 5 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.874	0.9948	0.878569	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9924	0.9964	0.995986	DRS	Too big
U5	1	1	1	CRS	
U6	0.8647	0.9865	0.876533	DRS	Too big
U7	0.8951	0.9871	0.906798	DRS	Too big
U8	1	1	1	CRS	
U9	0.9276	0.996	0.931325	DRS	Too big
U10	0.8256	0.9995	0.826013	DRS	Too big
U11	0.9244	1	0.9244	DRS	Too big
U12	0.9985	1	0.9985	DRS	Too big
U13	1	1	1	CRS	
U14	0.8409	0.9911	0.848451	DRS	Too big
U15	0.9639	1	0.9639	DRS	Too big
U16	1	1	1	CRS	
U17	0.9169	0.9966	0.920028	DRS	Too big
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	0.9004	1	0.9004	DRS	Too big
U21	0.8841	0.9757	0.906119	DRS	Too big
U22	1	1	1	CRS	
U23	0.9426	1	0.9426	DRS	Too big
U24	0.9415	0.9952	0.946041	DRS	Too big
U25	1	1	1	CRS	
U26	0.8912	1	0.8912	DRS	Too big
U27	0.8561	0.972	0.880761	DRS	Too big
U28	1	1	1	CRS	
Average	**0.944282**	**0.996104**	0.947772		
%	**94.42%**	**99.61%**	94.78%		

In order to compare the potential cost savings of this model with the potential cost savings as identified by staff category in Model 2, Model 5 was run again but on this occasion with an input orientation and with variable returns to scale. The results of this are set out in Tables 5.21 and 5.22.

Table 5.21

Model 5 with an input orientation

Hospital	Technical Efficiency
U1	0.9437
U2	1
U3	1
U4	1
U5	1
U6	0.880
U7	0.974
U8	1
U9	0.9118
U10	0.9883
U11	1
U12	1
U13	1
U14	0.9419
U15	1
U16	1
U17	0.9911
U18	1
U19	1
U20	1
U21	1
U22	1
U23	1
U24	0.9763
U25	1
U26	1
U27	0.9129
U28	1
Average	**0.982857**
%	**98.29**

Table 5.21 highlights the technical efficiency of each of the hospitals in the sample using an input orientation and variable returns to scale in the model. The results show that nineteen hospitals (67.86%) were on the best practice frontier and thus deemed to be relatively efficient. The average technical efficiency score for all of the hospitals was 98.29%. This would indicate that collectively all of the hospitals in the sample could continue to produce their current output levels with a 1.71% reduction in their inputs. The potential savings to be made if all of the inefficient hospitals were on the best practice frontier are set out in table 5.22. The total savings that could be achieved would be € 43,095,279 in pay and € 21,431,882 in non-pay, as well as a reduction in bed numbers of 164.

The hospital with the largest potential for making savings was U27 where pay savings of € 12,644,427 and non-pay savings of € 6,939,909 could be made. Bed numbers could also potentially be reduced in U27 by 41. The potential for pay savings varied between € 12,644,427 in U27 and € 437,257 in U17 whilst the potential for non-pay savings varied between € 6,939,909 in U27 and € 177,697 in U17. The potential to reduce bed numbers varied between 41 in U27 and 2 in U17.

Table 5.22

Model 5 – potential to make savings

Hospital	Pay Savings €	Non-pay Savings €	Bed Savings
U1	5,702,627	3,108,042	24
U2	0	0	0
U3	0	0	0
U4	0	0	0
U5	0	0	0
U6	6,918,240	2,621,880	32
U7	1,413,646	699,062	6
U8	0	0	0
U9	7,674,106	3,230,413	28
U10	919,913	363,215	4
U11	0	0	0
U12	0	0	0
U13	0	0	0
U14	6,014,919	3,603,653	22
U15	0	0	0
U16	0	0	0
U17	437,257	177,697	2
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	0	0	0
U22	0	0	0
U23	0	0	0
U24	1,370,144	688,011	5
U25	0	0	0
U26	0	0	0
U27	12,644,427	6,938,909	41
U28	0	0	0
Total	**43,095,279**	**21,431,882**	**164**

Interestingly when one compares the potential savings to be made in this model with the potential savings in model 2, as set out in table 5.7, which uses medical pay and nursing pay as well as bed numbers as inputs, it is clear that there would appear to be a greater potential to make savings in this model. Model 2 indicated that savings of €10,491,000 in medical pay and €15,064,000 in nursing pay as well as a reduction in bed numbers of 147 could be made whereas this model indicates that total savings of €64,257,000 as well as a reduction in bed numbers of 164 could be made. It is therefore clear that in order to maximise savings that the pay of all categories of staff as well as non-pay costs need to be included in the model if the savings potential of all of the hospitals is to be maximised. It is only in hospitals U6, U9 and U17 that there appears to be potential to make savings in both models. Hospitals U4, U13, U15, U21, U23 and U25 have potential to make savings in pay in model 2 while hospitals U1, U7, U10, U14, U24 and U27 have potential to make savings in both pay and non-pay in model 5. In order to make the greatest savings each hospital should focus on those areas where they have the greatest potential to make savings.

5.8 Model 6

In Model 6, pay and non pay costs in each hospital were included as inputs. However, on this occasion only quantitative measures were used. To ensure comparability between hospitals both pay and non-pay costs were adjusted as in Model 5. The inputs were therefore pay costs (as adjusted for administration and superannuation costs), non-pay costs (as adjusted for insurance costs) and patient bed numbers. The outputs were inpatient discharges (as adjusted for casemix), new outpatient attendances and day patient attendances (as adjusted for casemix) in each hospital. The model was input orientated with variable returns to scale. It was run with the assumption that the objective was to minimise inputs for a given level of output. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 5.23 sets out the inputs and outputs in Model 6 and the descriptive statistics of each of the variables.

Table 5.23

Model 6 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Pay Costs as adjusted for superannuation and admin.	85025.46	55018.95
Non-Pay Costs as adjusted for insurance costs	42724.46	36964.79
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
New outpatient attendances	21301.04	15762.35
Day patient attendances as adjusted for casemix	20239.54	15275.98

The results for Model 5 are set out in tables 5.24, 5.25 and 5.26.

Table 5.24

Model 6 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	1	7,12,19,24	0
U2	1	2	10
U3	1	2,3	4
U4	1	2,3,4,16	0
U5	1	5,19	4
U6	0.9117	2,3,7,12,19	0
U7	1	7	6
U8	1	8,12	0
U9	1	9,19,24,12	1
U10	1	2,7,24	0
U11	0.9931	3,7,19	0
U12	1	12	5
U13	1	2,12,13,16	1
U14	0.9713	7,19,20,24	0
U15	0.9683	2,5,19	0
U16	1	16	4
U17	0.9628	2,13,16,24,25	0
U18	1	18	1
U19	1	19,24	10
U20	1	20,24	3
U21	0.8829	2,5,19,28	0
U22	1	12,19,22,24	0
U23	0.9442	2,5,9,18,28	0
U24	1	24	9
U25	1	16,25	1
U26	1	3,19	0
U27	0.9971	5,20,24	0
U28	1	28	2
Average	**0.987**		
%	**98.7%**		

Table 5.24 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each one. The results showed that twenty of the twenty eight hospitals (71.4%) were on the best practice frontier and thus deemed to be efficient. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency of all of the hospitals was 98.7%. This would appear to indicate a high level of technical efficiency overall. It also indicates that collectively all of the hospitals in the sample could produce their current output levels with a 1.3% reduction in the inputs included in the model. Whilst this is a relatively low percentage it would equate to a not inconsiderable saving in pay costs, in non-pay costs and in bed resources. Table 5.25 sets out the potential savings to be made in each of the inefficient hospitals if they were to operate on the efficient production frontier. The total savings that could be achieved would be € 27,485,947 in pay, € 12,491,466 in non-pay and a reduction of 124 in bed numbers.
Overall nineteen hospitals achieved an efficiency score of 100%. However, many of these hospitals do not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U19, U2 and U24 are truly efficient as they are respectively peers for ten, ten and twelve other hospitals respectively in the sample. HospitalsU3, U5, U7, U12 and U16 are peers for between four and six other hospitals and while they achieved an efficiency score of 100%, there may be scope for them to improve their efficiency. Hospitals U13, U18, U20, U25 and U28 are respectively peers for between one and three other hospitals in the sample and while they each achieved an efficiency score of 100%, there may also be scope for them to improve their efficiency.

Table 5.25 sets out the potential savings to be made in each of the inefficient hospitals if they were to operate on the efficient production frontier.

Table 5.25

Model 6 – potential savings

Hospital	Pay Savings €	Non Pay Savings €	Bed Savings
U1	0	0	0
U2	0	0	0
U3	0	0	0
U4	0	0	0
U5	0	0	0
U6	5,090,672	1,929,267	27
U7	0	0	0
U8	0	0	0
U9	0	0	0
U10	0	0	0
U11	1,037,505	687,958	5
U12	0	0	0
U13	0	0	0
U14	2,971,225	1,780,118	14
U15	2,629,103	942,156	11
U16	0	0	0
U17	1,827,636	742,735	9
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	8,029,196	3,840,763	29
U22	0	0	0
U23	2,485,947	872,398	12
U24	0	0	0
U25	0	0	0
U26	0	0	0
U27	421,663	231,031	2
U28	0	0	0
Total	**24,492,947**	**11,026,426**	**109**

As we can see the potential to make savings varied in pay from € 8,029,196 in U21 to € 421,663 in U27 and in non-pay from € 3,840,763 in U21 to € 231,031 in U27. The potential to reduce bed numbers ranged from 29 in U21 to 2 in U27.

Table 5.26 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. It is clear from this table that some hospitals appear to be operating at too large a scale to maximise the productivity of their inputs while others appear to be too small and are exhibiting higher than average costs. The average scale efficiency score of all of the hospitals was 96.8%. The results showed that eleven (39.3%) out of twenty eight hospitals were operating at optimal size. The pattern of scale inefficiency indicates that nine (32.16%) hospitals were operating on increasing returns to scale and were therefore too small. A further eight (28.6%) hospitals were operating on decreasing returns to scale and were therefore too big.

Table 5.26

Model 6 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.9525	1	0.9525	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9473	1	0.9473	IRS	Too small
U5	1	1	1	CRS	
U6	0.9111	0.9117	0.999342	DRS	Too big
U7	1	1	1	CRS	
U8	1	1	1	CRS	
U9	1	1	1	CRS	
U10	0.9902	1	0.9902	DRS	Too big
U11	0.954	0.9931	0.960628	DRS	Too big
U12	1	1	1	CRS	
U13	0.9565	1	0.9565	IRS	Too small
U14	0.9388	0.9713	0.96654	DRS	Too big
U15	0.9639	0.9683	0.995456	DRS	Too big
U16	0.8831	1	0.8831	IRS	Too small
U17	0.933	0.9628	0.969049	IRS	Too small
U18	0.9415	1	0.9415	IRS	Too small
U19	1	1	1	CRS	
U20	1	1	1	CRS	
U21	0.8348	0.8829	0.94552	IRS	Too small
U22	1	1	1	CRS	
U23	0.908	0.9442	0.96166	IRS	Too small
U24	1	1	1	CRS	
U25	0.8894	1	0.8894	IRS	Too small
U26	0.8912	1	0.8912	DRS	Too big
U27	0.9883	0.9971	0.991174	DRS	Too big
U28	0.8626	1	0.8626	IRS	Too small
Average	**0.955221**	**0.987**	**0.968**		
%	**95.52%**	**98.7%**	**96.8%**		

When comparing the potential savings to be made in this model with model 5 it is clear that there is a greater potential to make savings in model 5. This is not surprising as one would expect there to be a greater potential to make savings in model 5 given that fewer hospitals were deemed to be on the efficient frontier in that model. The potential to make additional savings is significant. In model 5 there is potential to make additional savings of €18,602,332 in pay and €10,405,456 in non-pay in excess of that achievable in model 6. The hospital with the greatest potential to make savings in model 5 was U27 whilst in model 6 it was U21. The only difference between the variables used in each model was that model 5 used the survival rate as an output whereas model 6 used day patient attendances instead. These results are a good example of the sensitivity of the DEA model to changes in variables. Both models had fifteen hospitals in common that were deemed to be on the efficient frontier. However, while hospitals U11, U15, U21 and U23 were on the efficient frontier in model 5 they were deemed to be inefficient in model 6. Similarly, hospitals U1, U7, U10 and U24 were deemed to be on the efficient frontier in model 6 but were deemed to be inefficient in model 5. It is therefore of critical importance when choosing variables for the DEA model to be clear on what one is measuring. It may seem obvious but the choice of variables that need to be included in the model is dependent on the aspect of performance being assessed. In model 5, volume and the clinical quality of care relative to operating and capital expenditure was being assessed whereas in model 6 it was only the volume delivered relative to operating and capital expenditure that was assessed. The inclusion in the model of a measure of the clinical quality of care in model 5 instead of a purely volume based measure as in model 6 resulted in significantly different results for some hospitals.

5.9 Model 7

Model 7 looked at overall staff numbers. To ensure comparability between hospitals administrative staff were not included in the numbers. A similar adjustment was made in models 5 and 6 in relation to administration pay costs. As much of the administrative support for the statutory hospitals is provided centrally and not reflected in the hospitals' numbers it would be inaccurate to include administrative staff as an input in the model. The inputs were therefore medical staff full-time equivalent numbers, other full-time equivalent staff numbers (less administration staff) and patient bed numbers. The outputs were inpatient discharges (as adjusted for casemix), new outpatient attendances and survival rates in each hospital. The model was output orientated with variable returns to scale. It was run with the assumption that the objective was to maximise outputs for a given level of input. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 5.27 sets out the inputs and outputs in Model 7 and the descriptive statistics of each of the variables.

Table 5.27

Model 7 – inputs and outputs

	Arithmetic Mean	**Standard Deviation**
Inputs		
Medical staff full-time-equivalents	178.31	121.64
Other staff full-time equivalents (less admin staff)	1081.04	681.63
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
New outpatient attendances	21301.04	15762.35
Patient survival rate	97.61	1.06

The results for Model 7 are set out in tables 5.28, 5.29 and 5.30.

Table 5.28

Model 7 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.9936	3,18,19	0
U2	1	2,3,18	2
U3	1	3,16	9
U4	0.9967	16,18,19,28	0
U5	1	5,19	2
U6	0.9876	2,3,16,19	0
U7	0.9869	18,19	0
U8	1	8,16	5
U9	0.9979	3,8,16,19	0
U10	1	3,8,19	0
U11	1	3,18,19	0
U12	1	12,13	2
U13	1	8,12,13,16	2
U14	0.992	3,19	0
U15	1	18,19,28	0
U16	1	16	8
U17	0.9996	12,13,16,18,28	0
U18	1	18	11
U19	1	19,28	17
U20	1	18,19	0
U21	0.9757	5,18,19,28	0
U22	1	8,19	0
U23	1	5,18,19,28	0
U24	1	8,24,28	0
U25	1	16,18,19,28	0
U26	1	3,19	0
U27	0.972	2,3,19	0
U28	1	28	8
Average	**0.9965**		
%	**99.65%**		

Table 5.28 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each. The results showed that nineteen (67.9%) hospitals were efficient and thus on the best practice frontier. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency of all of the hospitals was 99.65%. This would appear to indicate a very high level of efficiency overall. It also indicates that collectively all of the hospitals in the sample could increase their current output levels by 0.35% without increasing their inputs.
Overall nineteen hospitals achieved an efficiency score of 100%. However many of these hospitals do not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U19 and U18 are truly efficient as they are respectively peers for seventeen and eleven other hospitals in the sample. Hospitals U3 and U28 are respectively peers for eight other hospitals and U16 is a peer for nine hospitals, and while they each achieved an efficiency score of 100% there may be scope for them to improve their efficiency. Likewise hospitals U2, U5, U12 and U13 are each peers for two other hospitals and U8 is a peer for five hospitals, and while they are each deemed to be efficient there may be scope for each of them to improve their efficiency.

Table 5.29 sets out the potential to increase outputs in each of the inefficient hospitals if they were to operate on the efficient production frontier. These hospitals could potentially increase their outputs, without increasing their inputs, by 1,671 additional inpatients and 1,758 additional new outpatients each year as well as improving the survival rates by between 0.04% as in U17 and 2.68% as in U27. Hospital U27, in particular, has potential to significantly increase its output, by 729 inpatients and 666 new outpatients. Likewise hospitals U1, U6, U7, U14 and U21 have potential to increase their inpatient numbers by 135,152, 158, 160 and 272 respectively and their new outpatients by 152, 175, 158, 213, and 308. Hospitals U4, U9 and U17 also have potential to increase their inpatient and new outpatient numbers, but to a lesser degree.
The potential to increase outputs in this model is greater than in model 3 where medical fulltime equivalents, nursing fulltime equivalents and patient bed were included as inputs whilst using the same outputs. The differences were 56 additional inpatients, 121 additional new outpatients and a 0.14% increase in survival rate.

However the potential to increase outputs in this model is lower than in models 4 and 5 where again the same outputs were used. The inputs used in models 4 and 5 were respectively medical salaries, nursing salaries and patient bed numbers in Model 4 and pay costs, non-pay costs and patient beds in Model 5. The differences were 377 additional inpatients and 199 new outpatients in Model 4 and 1,719 additional inpatients and 132 additional new outpatients in Model 5. The potential to increase the survival rate however is greater in Model 7 than in Model 4 by 0.21% while it is the same as in Model 5.

Table 5.29

Model 7 – potential to increase outputs

Hospital	Additional Inpatients	Additional New Outpatients	Increased Survival Rate
U1	135	152	0.0063
U2	0	0	0
U3	0	0	0
U4	30	29	0.0032
U5	0	0	0
U6	152	175	0.0121
U7	158	158	0.0127
U8	0	0	0
U9	31	51	0.0021
U10	0	0	0
U11	0	0	0
U12	0	0	0
U13	0	0	0
U14	160	213	0.0078
U15	0	0	0
U16	0	0	0
U17	4	6	0.0004
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	272	308	0.0234
U22	0	0	0
U23	0	0	0
U24	0	0	0
U25	0	0	0
U26	0	0	0
U27	729	666	0.0268
U28	0	0	0
Totals	**1671**	**1758**	

Table 5.30 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. This table indicates that seventeen (60.7%) out of the twenty eight hospitals are operating at too large a scale to maximise the productivity of their inputs while eleven (39.3%) hospitals are operating at their optimal size. The average scale efficiency for all of the hospitals was 94.87%.

Table 5.30

Model 7 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.8863	0.9936	0.892009	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9464	0.9967	0.949533	DRS	Too big
U5	1	1	1	CRS	
U6	0.8727	0.9876	0.883657	DRS	Too big
U7	0.894	0.9869	0.905867	DRS	Too big
U8	1	1	1	CRS	
U9	0.9199	0.9979	0.921836	DRS	Too big
U10	0.8164	1	0.8164	DRS	Too big
U11	0.9529	1	0.9529	DRS	Too big
U12	1	1	1	CRS	
U13	1	1	1	CRS	
U14	0.8339	0.992	0.840625	DRS	Too big
U15	0.9257	1	0.9257	DRS	Too big
U16	1	1	1	CRS	
U17	0.9667	0.9996	0.967087	DRS	Too big
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	0.942	1	0.942	DRS	Too big
U21	0.8971	0.9757	0.919442	DRS	Too big
U22	0.9282	1	0.9282	DRS	Too big
U23	0.9128	1	0.9128	DRS	Too big
U24	1	1	1	CRS	
U25	0.9672	1	0.9672	DRS	Too big
U26	0.9455	1	0.9455	DRS	Too big
U27	0.8678	0.972	0.892798	DRS	Too big
U28	1	1	1	CRS	
Average	**0.945554**	**0.9965**	**0.948698**		
%	**94.56%**	**99.65%**	**94.87%**		

Comparing this model with model 5, which had the same output variables, indicates a greater potential in model 5 to increase outputs whilst keeping inputs constant. The results from model 5 indicated the potential to increase inpatients by 1,788, new outpatients by 1,890 and an increased survival rate of between 0.05% and 2.68% among the hospitals deemed to be inefficient. The results from model 6 indicated the potential to increase inpatients by 1,671, new outpatients by 1,758 and an increased survival rate of between 0.04% and 2.68%. Model 5 deemed seventeen of the hospitals in the sample to be on the efficient frontier whilst model 7 deemed nineteen hospitals to be so. Given that model 5 had fewer hospitals that were deemed to be efficient it is not surprising that it has a greater potential to increase outputs among its inefficient hospitals.

Including pay costs and non-pay costs as inputs in model 5 instead of medical full-time equivalents and other staff full-time equivalents in model 7 resulted in fewer hospitals being deemed to be efficient. One explanation could be that using full-time equivalents as an input measure was hiding other payroll costs that did not impact on full-time equivalent numbers such as overtime or premium payments and that some hospitals maintained lower full-time equivalent numbers by incurring higher costs in these areas. All of these costs would have been picked up when using pay cost as an input measure, thus ensuring a more accurate measurement of labour input in the model. A second explanation could be the higher use by some hospitals of agency staff who would not have been reflected in the full-time equivalent numbers. These costs would have been picked up as part of either pay or non-pay costs. A third explanation could be that the age profile of staff disproportionately impacted on the salary levels in some hospitals. This could occur in a hospital with an older age profile and with more experienced staff on higher salaries. In this situation using full-time equivalents instead of salaries as an input would be a more accurate measure of labour efficiency in the model.

5.10 Model 8

Model 8 looked at the non-clinical quality of care being offered to the patients attending the hospitals. The inputs used were the total number of inpatients treated and subsequently discharged and the total number of outpatient attendances. The outputs used were the infection free rate and the hygiene rating for each of the hospitals. The infection free rate was derived from the HSE healthcare associated infection report on staphylococcus aureus bloodstream isolates in Ireland by acute public hospital and the hygiene rating for each hospital was available from the HSE national hygiene audit report. I would have also wished to use the level of complaints made against each hospital as a measure but this information was not available. The model was output orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 5.31 sets out the inputs and outputs in Model 8 and the descriptive statistics of each of the variables.

Table 5.31

Model 8 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Inpatient discharges	15993.07	6912.49
Outpatient attendances	81509.68	53030.58
Outputs		
Infection free rate	86.39	6.88
Hygiene rating	53.54	13.79

The results for Model 8 are set out in tables 5.32, 5.33 and 5.34.

Table 5.32

Model 8 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.858	16,28	0
U2	1	16	0
U3	1	16	0
U4	1	16	0
U5	0.858	16,28	0
U6	0.852	16,28	0
U7	0.8492	16,28	0
U8	0.961	16,28	0
U9	0.8837	16,28	0
U10	0.9004	16,28	0
U11	0.8664	16,28	0
U12	0.8192	16	0
U13	0.9505	16	0
U14	0.9081	16,28	0
U15	0.9604	16	0
U16	1	16	20
U17	0.8932	16,28	0
U18	0.8822	16	0
U19	0.8259	28	0
U20	1	28	0
U21	0.8776	16,28	0
U22	1	28	0
U23	0.9342	16	0
U24	0.8525	28	0
U25	0.9543	16,28	0
U26	1	28	0
U27	1	28	0
U28	1	28	18
Average	**0.924529**		
%	**92.45%**		

Table 5.32 highlights the efficiency scores for each hospital in the sample as well as showing the peer groups for each hospital. It also gives the peer count for each hospital, which indicates the number of times that each hospital appears in the peer group of other hospitals. The results showed that nine (32.1%) hospitals were efficient and thus on the efficient production frontier. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency was 92.45%. This would appear to indicate that collectively all of the hospitals in the sample could increase their output by 7.55% without increasing their inputs. In this model this result could be interpreted as meaning that each of the hospitals could improve both their infection free rate and their hygiene rating by 7.55% based on their current inputs. Overall nine hospitals achieved an efficiency score of 100%. However many of these hospitals do not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U16 and U28 are truly efficient as they are respectively peers for twenty and eighteen other hospitals in the sample.

Table 5.33 sets out the potential increase in outputs in each of the inefficient hospitals if they were to operate on the efficient production frontier. The potential to increase the infection free rate varied from 3.63% in U8 to 13.92% in U12. Similarly the potential to increase the hygiene rating varied from 1.11% in U15 to 9.57% in U12.

Table 5.33

Model 8 – potential to increase outputs

Hospital	Increased Infection Free Rate	Increased Hygiene Rating
U1	0.1164	0.0753
U2	0	0
U3	0	0
U4	0	0
U5	0.1164	0.0753
U6	0.1199	0.0784
U7	0.1176	0.0799
U8	0.0363	0.0207
U9	0.0989	0.0616
U10	0.0867	0.0528
U11	0.1109	0.0708
U12	0.1392	0.0957
U13	0.0465	0.0139
U14	0.0809	0.0487
U15	0.038	0.0111
U16	0	0
U17	0.0918	0.0566
U18	0.1004	0.0327
U19	0.1358	0.0923
U20	0	0
U21	0.1016	0.0649
U22	0	0
U23	0.0599	0.0349
U24	0.1165	0.0782
U25	0.0425	0.0242
U26	0	0
U27	0	0
U28	0	0

Table 5.34 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. This table indicates that twenty six (92.9%) out of the twenty eight hospitals were operating at too large a scale to maximise the productivity of their inputs while only two (7.1%) hospitals were operating at their optimal size. The average scale efficiency for all of the hospitals was 45.43%.

Table 5.34

Model 8 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.2067	0.858	0.240909	DRS	Too big
U2	0.5921	1	0.5921	DRS	Too big
U3	0.338	1	0.338	DRS	Too big
U4	0.4709	1	0.4709	DRS	Too big
U5	0.1832	0.858	0.21352	DRS	Too big
U6	0.3571	0.852	0.419131	DRS	Too big
U7	0.5304	0.8492	0.624588	DRS	Too big
U8	0.6482	0.961	0.674506	DRS	Too big
U9	0.284	0.8837	0.321376	DRS	Too big
U10	0.2511	0.9004	0.278876	DRS	Too big
U11	0.1566	0.8664	0.180748	DRS	Too big
U12	0.3225	0.8192	0.393677	DRS	Too big
U13	0.4602	0.9505	0.484166	DRS	Too big
U14	0.2275	0.9081	0.250523	DRS	Too big
U15	0.2232	0.9604	0.232403	DRS	Too big
U16	1	1	1	CRS	
U17	0.389	0.8932	0.435513	DRS	Too big
U18	0.692	0.8822	0.784403	DRS	Too big
U19	0.2946	0.8259	0.356702	DRS	Too big
U20	0.3277	1	0.3277	DRS	Too big
U21	0.5095	0.8776	0.580561	DRS	Too big
U22	0.2886	1	0.2886	DRS	Too big
U23	0.3876	0.9342	0.4149	DRS	Too big
U24	0.495	0.8525	0.580645	DRS	Too big
U25	0.4824	0.9543	0.505501	DRS	Too big
U26	0.2916	1	0.2916	DRS	Too big
U27	0.4397	1	0.4397	DRS	Too big
U28	1	1	1	CRS	
Average	**0.423193**	**0.924528571**	**0.45433**		
%	**42.32%**	**92.45%**	**45.43%**		

The results from this model, which assessed the non clinical quality of care received by patients relative to volume, when compared to the other models in this research showed the fewest number of hospitals on the efficient frontier. Only nine hospitals were deemed to be on the efficient frontier. This appears to indicate that much work needs to be done by Irish hospitals if they are to improve efficiency in this area.

5.11 Model 9

Model 9 brought together pay costs and non-pay costs for each hospital as inputs along with quantitative, clinical and non-clinical qualitative outcomes as outputs. The inputs were therefore pay costs (as adjusted for administration and superannuation costs), non-pay costs (as adjusted for insurance costs) and patient bed numbers. The outputs were inpatient discharges (as adjusted for casemix), survival rates and hygiene ratings in each hospital. The model was output orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 5.35 sets out the inputs and outputs in Model 9 and the descriptive statistics of each of the variables.

Table 5.35

Model 9 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Pay Costs as adjusted for superannuation and admin.	85025.46	55018.95
Non-Pay Costs as adjusted for insurance costs	42724.46	36964.79
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
Patient survival rate	97.61	1.06
Hygiene rating	53.54	13.79

The results for Model 9 are set out in tables 5.36, 5.37 and 5.38.

Table 5.36

Model 9 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.9947	2,3,19	0
U2	1	2,16	19
U3	1	3	12
U4	0.9964	2,3,25	0
U5	1	5,28	3
U6	0.9865	2,3,19	0
U7	0.9871	2,3,19	0
U8	0.9991	2,3,16	0
U9	0.9969	2,18,19,29	0
U10	1	2,3,19	0
U11	1	3,19	0
U12	0.9952	2,3,19	0
U13	0.9972	3,16,18	0
U14	0.993	2,3,19	0
U15	1	2,5	0
U16	1	16	4
U17	0.9949	2,18,19,28	0
U18	1	18	5
U19	1	19,28	16
U20	1	2,3,19	0
U21	0.9771	2,5,19,28	0
U22	1	2,5,19,28	0
U23	1	2,18,19,28	0
U24	0.9837	2,18,19,28	0
U25	1	2,16,25	1
U26	1	2,3,19	0
U27	1	2,19,28	0
U28	1	28	9
Average	**0.996493**		
%	**99.65%**		

Table 5.36 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each. The results showed that sixteen (57.1%) hospitals were efficient and thus on the best practice frontier. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency of all of the hospitals was 99.65%. This would appear to indicate a very high level of efficiency overall. It also indicated that collectively all of the hospitals in the sample could increase their current output levels by 0.35% without increasing their inputs.
Overall sixteen hospitals achieved an efficiency score of 100%. However many of these hospitals do not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U19, U2 and U3 are truly efficient as they are respectively peers for sixteen, nineteen and twelve other hospitals in the sample. Hospital U28 is a peer for nine other hospitals and while it achieved an efficiency score of 100% there may be scope for it to improve its efficiency. Likewise hospitals U5, U16, U18 and U25 are respectively peers for three, four, five and one other hospital and while they also achieved an efficiency score of 100%, there may be scope for them to improve their efficiency.
Table 5.37 sets out the potential to increase outputs in each of the inefficient hospitals if they were to operate on the efficient production frontier. These hospitals could potentially increase their outputs, without increasing their inputs, by 1,209 additional inpatients as well as improving the survival rates by between 0.27% as in U13 and 2.21% as in U21 and increasing the hygiene ratings by between 0.05% in U8 and 1.21% in U21. The hospitals with the greatest potential to increase their inpatient numbers are U21, U6, U24, U7, U14 and U1, which could respectively increase their inpatient numbers by 256, 165, 157, 156, 140 and 112.

Table 5.37

Model 9 – potential to increase outputs

Hospital	Additional Inpatients	Increased Survival Rate	Increased Hygiene Rating
U1	112	0.0052	0.0028
U2	0	0	0
U3	0	0	0
U4	33	0.0035	0.0019
U5	0	0	0
U6	165	0.0131	0.0072
U7	156	0.0126	0.0068
U8	8	0.0009	0.0005
U9	46	0.0031	0.0016
U10	0	0	0
U11	0	0	0
U12	65	0.0047	0.0025
U13	23	0.0027	0.0008
U14	140	0.0069	0.0037
U15	0	0	0
U16	0	0	0
U17	48	0.0050	0.0027
U18	0	0	0
U19	0	0	0
U20	0	0	0
U21	256	0.0221	0.0121
U22	0	0	0
U23	0	0	0
U24	157	0.0158	0.0086
U25	0	0	0
U26	0	0	0
U27	0	0	0
U28	0	0	0
Totals	**1209**		

Table 5.38 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. This table indicates that twenty (71.4%) out of the twenty eight hospitals are operating at too large a scale to maximise the productivity of their inputs while eight (28.6%) hospitals are operating at their optimal size. The average scale efficiency for all of the hospitals was 92.39%.

Table 5.38

Model 9 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns To scale	
U1	0.8762	0.9947	0.880869	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9953	0.9964	0.998896	DRS	Too big
U5	1	1	1	CRS	
U6	0.8693	0.9865	0.881196	DRS	Too big
U7	0.9008	0.9871	0.912572	DRS	Too big
U8	0.934	0.9991	0.934841	DRS	Too big
U9	0.9227	0.9969	0.925569	DRS	Too big
U10	0.826	1	0.826	DRS	Too big
U11	0.9244	1	0.9244	DRS	Too big
U12	0.9209	0.9952	0.925342	DRS	Too big
U13	0.873	0.9972	0.875451	DRS	Too big
U14	0.8444	0.993	0.850352	DRS	Too big
U15	0.9639	1	0.9639	DRS	Too big
U16	1	1	1	CRS	
U17	0.8451	0.9949	0.849432	DRS	Too big
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	0.9007	1	0.9007	DRS	Too big
U21	0.8841	0.9771	0.90482	DRS	Too big
U22	0.7929	1	0.7929	DRS	Too big
U23	0.9426	1	0.9426	DRS	Too big
U24	0.7971	0.9837	0.810308	DRS	Too big
U25	1	1	1	CRS	
U26	0.8922	1	0.8922	DRS	Too big
U27	0.8771	1	0.8771	DRS	Too big
U28	1	1	1	CRS	
Average	**0.920811**	**0.996492857**	**0.923909**		
%	**92.08%**	**99.65%**	**92.39%**		

Comparing the results from this model with those from model 5, which had the same input variables, indicates that model 5 had a greater potential to increase inpatients amongst those hospitals deemed to be inefficient. Model 5 indicated a potential to increase inpatient numbers by 1,788 whereas this model indicated a potential to increase inpatient numbers by 1,209. The potential to increase the survival rate in this model ranged from 0.09% in U8 to 2.21% in U21, whilst in model 5 this ranged from 0.05% in U10 to 2.68% in U27.

5.12 Model 10

Model 10 looked at staff numbers as inputs along with quantitative, clinical and non-clinical qualitative outcomes as outputs. To ensure comparability between hospitals administrative staff were not included in the numbers. A similar adjustment was made in model 7. As much of the administrative support for the statutory hospitals is provided centrally and not reflected in the hospitals' numbers it would be inaccurate to include administrative staff as an input in the model. The inputs were therefore medical staff full-time equivalent numbers, other full-time equivalent staff numbers (less administration staff) and patient bed numbers. The outputs were inpatient discharges (as adjusted for casemix), survival rates and hygiene ratings in each hospital. The model was output orientated with variable returns to scale. The model was also run under constant returns to scale in order to measure scale efficiencies.

Table 5.39 sets out the inputs and outputs in Model 10 and the descriptive statistics of each of the variables.

Table 5.39

Model 10 – inputs and outputs

	Arithmetic Mean	Standard Deviation
Inputs		
Medical staff full-time-equivalents	178.31	121.64
Other staff full-time equivalents (less admin staff)	1081.04	681.63
Number of patient beds	369.54	196.67
Outputs		
Inpatient discharges as adjusted for casemix	16552.5	10167.89
Patient survival rate	97.61	1.06
Hygiene rating	53.54	13.79

The results for Model 10 are set out in tables 5.40, 5.41 and 5.42.

Table 5.40

Model 10 – results

Hospital	Technical Efficiency	Peer Group	Peer Count
U1	0.9945	18,19	0
U2	1	2,16,28	14
U3	1	3	5
U4	0.9968	2,18,19,28	0
U5	1	5,28	1
U6	0.9876	2,3,19	0
U7	0.9872	2,18,19	0
U8	0.9957	16,18	0
U9	0.9969	2,3,18,19	0
U10	1	2,3,19	0
U11	1	18,19	0
U12	1	2,18,19	0
U13	0.998	18	0
U14	0.993	2,3,19	0
U15	1	18,19,28	0
U16	1	16	5
U17	0.9975	2,16,18	0
U18	1	16,18	16
U19	1	19,28	17
U20	1	18,19	0
U21	0.9771	2,18,19,28	0
U22	1	2,18,19	0
U23	1	2,18,19,28	0
U24	0.9837	2,18,19,28	0
U25	1	2,16,18,28	0
U26	1	3,19	0
U27	1	2,5,19,28	0
U28	1	28	10
Average	**0.996714**		
%	**99.67%**		

Table 5.40 highlights the technical efficiency scores for each hospital in the sample as well as showing the peer groups for each. The results showed that seventeen (60.7%) hospitals were efficient and thus on the best practice frontier. It is important to note that the optimal solution can include slacks. These are the extra amounts by which an input (output) can be reduced (increased) to attain technical efficiency after all inputs (outputs) have been reduced (increased) in equal proportions to reach the production frontier. Hospitals that are deemed to be technically efficient in the model are inefficient if they contain slacks. The average relative efficiency of all of the hospitals was 99.67%. This would appear to indicate a very high level of efficiency overall. It also indicated that collectively all of the hospitals in the sample could increase their current output levels by 0.33% without increasing their inputs.
Overall seventeen hospitals achieved an efficiency score of 100%. However many of these hospitals do not appear in peer groups for other hospitals and thus would not appear to be efficient at all. It is far more likely that hospitals U18, U19, U28 and U2 are truly efficient as they are respectively peers for sixteen, seventeen, ten and fourteen other hospitals in the sample. Hospitals U5, U3 and U16 are respectively peers for one, five and five other hospitals and while they each achieved an efficiency score of 100% there may be scope for them to improve their efficiency.

Table 5.41 sets out the potential to increase outputs in each of the inefficient hospitals if they were to operate on the efficient production frontier. These hospitals could potentially increase their outputs, without increasing their inputs, by 1,127 additional inpatients as well as improving the survival rates by between 0.20% as in U13 and 2.21% as in U21 and increasing the hygiene ratings by between 0.13% in U17 and 1.21% in U21. The hospitals with the greatest potential to increase their inpatient numbers are U21, U24, U7, U6, U14 and U1, which could respectively increase their inpatient numbers by 256, 157, 155, 152, 140 and 116.

Table 5.41

Model 10 – potential to increase outputs

Hospital	Additional Inpatients	Increased Survival Rate	Increased Hygiene Rating
U1	116	0.0054	0.0029
U2	0	0.00	0.00
U3	0	0.00	0.00
U4	29	0.0031	0.0017
U5	0	0.00	0.00
U6	152	0.0121	0.0066
U7	155	0.0125	0.0068
U8	37	0.0042	0.0023
U9	46	0.0031	0.0016
U10	0	0.00	0.00
U11	0	0.00	0.00
U12	0	0.00	0.00
U13	16	0.0020	0.06
U14	140	0.0069	0.0037
U15	0	0.00	0.00
U16	0	0.00	0.00
U17	23	0.0025	0.0013
U18	0	0.00	0.00
U19	0	0.00	0.00
U20	0	0.00	0.00
U21	256	0.0221	0.0121
U22	0	0.00	0.00
U23	0	0.00	0.00
U24	157	0.0158	0.0086
U25	0	0	0.00
U26	0	0	0.00
U27	0	0	0.00
U28	0	0	0.00
Totals	**1127**		

Table 5.42 presents the technical efficiency for each of the hospitals in the sample under both constant and variable returns to scale and calculates the scale efficiency for each hospital. This table indicates that twenty one (75%) out of the twenty eight hospitals were operating at too large a scale to maximise the productivity of their inputs while seven (25%) hospitals were operating at their optimal size. The average scale efficiency for all of the hospitals was 92.67%.

Table 5.42

Model 10 – scale efficiencies

Hospital	Technical Efficiency CRTS	Technical Efficiency VRTS	Scale Efficiency	Returns to scale	
U1	0.8853	0.9945	0.890196	DRS	Too big
U2	1	1	1	CRS	
U3	1	1	1	CRS	
U4	0.9512	0.9968	0.954254	DRS	Too big
U5	1	1	1	CRS	
U6	0.8746	0.9876	0.885581	DRS	Too big
U7	0.909	0.9872	0.920786	DRS	Too big
U8	0.8871	0.9957	0.890931	DRS	Too big
U9	0.9218	0.9969	0.924666	DRS	Too big
U10	0.8179	1	0.8179	DRS	Too big
U11	0.9516	1	0.9516	DRS	Too big
U12	0.9402	1	0.9402	DRS	Too big
U13	0.8747	0.998	0.876453	DRS	Too big
U14	0.8355	0.993	0.84139	DRS	Too big
U15	0.9257	1	0.9257	DRS	Too big
U16	1	1	1	CRS	
U17	0.9165	0.9975	0.918797	DRS	Too big
U18	1	1	1	CRS	
U19	1	1	1	CRS	
U20	0.942	1	0.942	DRS	Too big
U21	0.9001	0.9771	0.921195	DRS	Too big
U22	0.7981	1	0.7981	DRS	Too big
U23	0.9364	1	0.9364	DRS	Too big
U24	0.8127	0.9837	0.826167	DRS	Too big
U25	0.9601	1	0.9601	DRS	Too big
U26	0.9455	1	0.9455	DRS	Too big
U27	0.8805	1	0.8805	DRS	Too big
U28	1	1	1	CRS	
Average	**0.923804**	**0.996714286**	**0.926729**		
%	**92.38%**	**99.67%**	**92.67%**		

Changing the way that labour was measured only had a minor impact on the model results. All of the hospitals in model 9 deemed to be on the efficient frontier were also deemed to be on the efficient frontier in model 10. However whilst model 9 had sixteen efficient hospitals model 10 had seventeen efficient hospitals. The additional hospital deemed to be efficient in model 10 was U12, which was deemed to be 99.52% efficient in model 9. The average efficiency scores were largely the same at 99.65% and 99.67% respectively and the average scale efficiency scores were also similar at 92.39% and 92.67% respectively.

Model 9 had a greater potential to increase outputs than model 10 amongst those hospitals deemed to be inefficient which is not surprising given that it had more hospitals that were deemed to be inefficient. Model 9 could potentially increase its output from the inefficient hospitals by 1,209 inpatients while model 10 could increase its output by 1,127 inpatients. The potential to increase survival rates varied between 0.09% and 2.21% in model 9 and between 0.20% and 2.21% in model 10. Similarly, the potential to increase the hygiene rating varied between 0.05% and 1.21% in model 9 and between 0.13% and 1.21% in model 10. The potential to increase both the survival rate and the hygiene rating was greatest in hospital U21 in both models. The results would appear to indicate that both models are robust to changes in the inputs used whether these are pay costs and non-pay costs or full-time equivalent staff numbers.

5.13 Summary of Results

Table 5.43 sets out a summary of the results from each of the models. This table shows the number of times that each hospital is deemed to be on the efficient frontier and its mean efficiency score. This should give a good indication, given the use of different inputs and outputs in each model, which hospitals are most likely to be efficient.

Table 5.43

Summary of efficiency results from all models

Hospital	Model										Mean
	1	2	3	4	5	6	7	8	9	10	
U1	0.969	1	0.995	0.9966	0.9948	1	0.9936	0.858	0.9947	0.9945	0.980
U2	1	1	1	1	1	1	1	1	1	1	1.000
U3	1	1	1	1	1	1	1	1	1	1	1.000
U4	0.9451	0.9454	0.9954	0.9957	0.9964	1	0.9967	1	0.9964	0.9968	0.987
U5	1	1	1	1	1	1	1	0.858	1	1	0.986
U6	0.9737	0.9618	0.9876	0.9874	0.9865	0.9117	0.9876	0.852	0.9865	0.9876	0.962
U7	1	1	0.9874	0.9875	0.9871	1	0.9869	0.8492	0.9871	0.9872	0.977
U8	1	1	1	1	1	1	1	0.961	0.9991	0.9957	0.996
U9	1	1	0.9979	0.997	0.996	1	0.9979	0.8837	0.9969	0.9969	0.987
U10	0.986	1	1	1	0.9995	1	1	0.9004	1	1	0.989
U11	0.9665	1	1	1	1	0.9931	1	0.8664	1	1	0.983
U12	1	1	1	1	1	1	1	0.8192	0.9952	1	0.981
U13	0.9748	0.9692	1	1	1	1	1	0.9505	0.9972	0.998	0.989
U14	0.9743	1	0.992	0.9926	0.9911	0.9713	0.992	0.9081	0.993	0.993	0.981
U15	0.9085	0.9342	1	1	1	0.9683	1	0.9604	1	1	0.977
U16	1	1	1	1	1	1	1	1	1	1	1.000
U17	0.9432	0.9497	0.9989	0.997	0.9966	0.9628	0.9996	0.8932	0.9949	0.9975	0.973
U18	1	1	1	1	1	1	1	0.8822	1	1	0.988
U19	1	1	1	1	1	1	1	0.8259	1	1	0.983
U20	1	1	1	0.9864	1	1	1	1	1	1	0.999
U21	0.8967	0.9069	0.9757	0.977	0.9757	0.8829	0.9757	0.8776	0.9771	0.9771	0.942
U22	1	1	1	1	1	1	1	1	1	1	1.000
U23	0.9299	0.9802	1	1	1	0.9442	1	0.9342	1	1	0.979
U24	1	1	1	1	0.9952	1	1	0.8525	0.9837	0.9837	0.979
U25	0.9605	0.9274	1	1	1	1	1	0.9543	1	1	0.984
U26	1	1	1	1	1	1	1	1	1	1	1.000
U27	0.9971	1	0.9735	0.9742	0.972	0.9971	0.972	1	1	1	0.989
U28	1	1	1	1	1	1	1	1	1	1	1.000
Average	**0.979**	**0.985**	**0.9966**	**0.9961**	**0.9961**	**0.987**	**0.9965**	**0.9245**	**0.9965**	**0.9967**	**0.985**
%	**97.9%**	**98.5%**	**99.66%**	**99.61%**	**99.61%**	**98.7%**	**99.65%**	**92.45%**	**99.65%**	**99.67%**	**98.5%**

Six hospitals were deemed to be on the efficient production frontier and thus 100% efficient by all ten models. These hospitals were U2, U3, U16, U22, U26 and U28. Interestingly these hospitals varied in size from 118 to 842 inpatient beds.

Four hospitals, U5, U18, U19 and U20, were deemed to be on the efficient production frontier by nine of the models. One hospital, U12 was deemed to be on the production frontier by eight of the models. Four hospitals, U8, U10, U11 and U25, were deemed to be on the efficient production frontier by seven of the models. Three hospitals, U15, U23 and U24, were deemed to be on the efficient production frontier by six of the models. One hospital, U13, was deemed to be on the efficient production frontier by five models. One hospital U27 was deemed to be on the efficient frontier by four models. Two hospitals, U7 and U9 were deemed to be on the efficient production frontier by three models. One hospital, U1 and U4 were deemed to be on the efficient production frontier by two models. One hospital, U14, was deemed to be on the efficient production frontier by one model. Finally, three hospitals, U6, U17 and U21 were not deemed to be on the efficient production frontier by any model.

While the results indicate a relatively high level of efficiency overall with an average efficiency score of 98.5% across all ten models only six hospitals were on the efficient frontier for all models. There is therefore scope for the remaining hospitals to improve their efficiency.

One point that is obvious from table 5.43 is that hospitals appear to be less efficient in model 8. One explanation for this may be that the number of inputs and outputs were reduced from six in the other models to four in this model. The larger the number of input and output variables used in relation to the number of hospitals in the model, the more hospitals will be assigned as fully efficient and hence the less discriminating the DEA model will be. That may be why hospitals appear to be less efficient in this model. This is an important point for policy makers to note.

Table 5.44 sets out the scale efficiency for each of the hospitals using the results from the different models.

Table 5.44

Summary of scale efficiencies of all models

Hospital	Model No.										Mean
	1	2	3	4	5	6	7	8	9	10	
U1	0.9613	1	0.8697	0.9213	0.8786	0.9525	0.892	0.2409	0.8809	0.8902	0.849
U2	0.994	0.9842	1	1	1	1	1	0.5921	1	1	0.957
U3	1	1	1	1	1	1	1	0.338	1	1	0.934
U4	0.956	0.9312	0.9508	0.9519	0.996	0.9473	0.9495	0.4709	0.9989	0.9543	0.911
U5	1	1	1	1	1	1	1	0.2135	1	1	0.921
U6	0.9613	0.9789	0.8837	0.8831	0.8765	0.9993	0.8837	0.4191	0.8812	0.8856	0.865
U7	1	1	0.9059	0.937	0.9068	1	0.9059	0.6246	0.9126	0.9208	0.911
U8	1	1	1	1	1	1	1	0.6745	0.9348	0.8909	0.950
U9	0.9905	0.9905	0.8022	0.9117	0.9313	1	0.9218	0.3214	0.9256	0.9247	0.872
U10	0.9241	0.997	0.8619	0.9403	0.826	0.9902	0.8164	0.2789	0.826	0.8179	0.828
U11	0.9648	1	0.9031	0.9771	0.9244	0.9606	0.9529	0.1807	0.9244	0.9516	0.874
U12	1	1	0.9988	1	0.9985	1	1	0.3937	0.9253	0.9402	0.926
U13	0.9774	0.9407	0.9776	0.9719	1	0.9565	1	0.4842	0.8755	0.8765	0.906
U14	0.9566	0.9693	0.8406	0.8924	0.8485	0.9665	0.8406	0.2505	0.8504	0.8414	0.826
U15	0.9574	0.9898	0.9115	0.9349	0.9639	0.9955	0.9257	0.2324	0.9639	0.9257	0.880
U16	1	1	1	1	1	0.8831	1	1	1	1	0.988
U17	0.9999	0.9618	0.9285	0.9367	0.92	0.969	0.9671	0.4355	0.8494	0.9188	0.889
U18	1	0.9689	1	1	1	0.9415	1	0.7844	1	1	0.969
U19	1	1	1	1	1	1	1	0.3567	1	1	0.936
U20	1	1	0.9687	0.9569	0.9004	1	0.942	0.3277	0.9007	0.942	0.894
U21	0.9354	0.9204	0.9194	0.9289	0.9061	0.9455	0.9194	0.5806	0.9048	0.9212	0.888
U22	0.9906	1	0.9283	1	1	1	0.9282	0.2886	0.7929	0.7981	0.873
U23	0.9984	0.9804	0.9284	0.9673	0.9426	0.9617	0.9128	0.4149	0.9426	0.9364	0.898
U24	1	1	1	0.9961	0.946	1	1	0.5806	0.8103	0.8262	0.916
U25	0.8856	0.9302	0.9672	0.9311	1	0.8894	0.9672	0.5055	1	0.9601	0.904
U26	0.9542	0.9656	0.9455	0.9583	0.8912	0.8912	0.9455	0.2916	0.8922	0.9455	0.868
U27	0.994	1	0.8972	0.8827	0.8808	0.9912	0.8928	0.4397	0.8771	0.8805	0.874
U28	0.9286	0.8897	1	1	1	0.8626	1	1	1	1	0.968
Average	**0.976**	**0.979**	**0.9425**	**0.96**	**0.9478**	**0.968**	**0.9487**	**0.4543**	**0.9239**	**0.9267**	**0.903**
%	**97.6%**	**97.9%**	**94.25%**	**96%**	**94.78%**	**96.8%**	**94.87%**	**45.43%**	**92.39%**	**92.67%**	**90.3%**

There were no hospitals that were deemed to be scale efficient by all ten of the models. However hospitals U3, U5, U16 and U19 were deemed to be scale efficient by nine of the models. Interestingly only U3 and U16 of these hospitals were deemed to be technically efficient by all of the models in Table 5.43. Hospitals U5 and U19 were deemed to be relatively efficient by nine of the models. Likewise hospitals U2, U22, U26 and U28 whilst they were deemed to be technically efficient they are not deemed to be scale efficient by all of the models. In fact U26 is not deemed to be scale efficient by any of the models. Overall the average scale efficiency score across all of the models is 90.30%. Whilst this is a relatively high score it is clear that all of the hospitals have potential to improve their scale efficiency levels.

Again this table highlights the lower scale efficiency of hospitals in model 8. As with the technical efficiency score, the lower number of inputs and outputs in the model possibly ensured that the model was more discriminating in measuring efficiency. The less flexible constant returns to scale model ensured an even lower measure of technical efficiency in this model.

5.14 Ranking of Hospitals

Table 5.45 sets out the rankings of the hospitals in each model based on their efficiency scores.

Table 5.45

Hospital Rankings

Hospital	Model										Mean
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
U1	21	1	23	21	23	1	23	22.5	23	23	18.15
U2	1	1	1	1	1	1	1	1	1	1	1
U3	1	1	1	1	1	1	1	1	1	1	1
U4	24	25	22	22	20	1	22	1	20	21	17.80
U5	1	1	1	1	1	1	1	22.5	1	1	3.15
U6	20	23	25	25	26	27	25	25	26	25	24.7
U7	1	1	26	24	25	1	26	26	25	26	18.10
U8	1	1	1	1	1	1	1	10	17	22	5.60
U9	1	1	21	19.5	21	1	21	18	19	20	14.25
U10	17	1	1	1	18	1	1	16	1	1	5.80
U11	22	1	1	1	1	22	1	21	1	1	7.20
U12	1	1	1	1	1	1	1	28	21	1	5.70
U13	18	22	1	1	1	1	1	13	18	18	9.40
U14	19	1	24	23	24	23	24	15	24	24	20.10
U15	27	26	1	1	1	24	1	11	1	1	9.40
U16	1	1	1	1	1	1	1	1	1	1	1
U17	25	24	20	19.5	19	25	20	17	22	19	21.05
U18	1	1	1	1	1	1	1	19	1	1	2.80
U19	1	1	1	1	1	1	1	27	1	1	3.60
U20	1	1	1	26	1	1	1	1	1	1	3.50
U21	28	28	27	27	27	28	27	20	28	28	26.80
U22	1	1	1	1	1	1	1	1	1	1	1
U23	26	21	1	1	1	26	1	14	1	1	9.30
U24	1	1	1	1	22	1	1	24	27	27	10.60
U25	23	27	1	1	1	1	1	12	1	1	6.90
U26	1	1	1	1	1	1	1	1	1	1	1
U27	16	1	28	28	28	21	28	1	1	1	15.30
U28	1	1	1	1	1	1	1	1	1	1	1

The mean rankings indicate the best performing hospitals to be U2, U3, U16, U22, U26 and U28. These hospitals were on the efficient frontier in every model.

The worst performing hospitals whose means were 15 or greater were U1, U4, U6, U7, U14, U17, U21 and U28. Hospitals U6, U14, U17 and U21 performed particularly badly with a mean ranking of greater than 20. The performance of these hospitals gives some cause for concern and may indicate serious inefficiencies in their operation.

5.15 Model Validity

Spearman rank correlations are used to validate the models in terms of internal validity. These correlations are calculated using Wessa, P (2012) free statistics software. *"Validity of findings may be divided into internal validity – do the methods alter the results? And external validity - are the results applicable more generally?* (Parkin and Hollingsworth, 1997: 1428)

As an analyst can use different configurations of data and methods within the same data set internal validity is critically important. External validity is also important as DEA is a technique that is meant to have more general applicability. A test for external validity is to show consistency over time. As this research only applies to one time period it was not possible to test for this. However, a test for internal validity was carried out by comparing the results obtained from each model using different combinations of inputs and outputs. The results of the Spearman's rank correlation coefficients are set out in table 5.46.

"A true test of validity requires the comparison of genuinely competing alternatives."

(Parkin and Hollingsworth, 1997: 1429

Table 5.46

Spearman's Rank Correlation Coefficients

Models	1	2	3	4	5	6	7	8	9
2	0.849								
3	0.523	0.497							
4	0.451	0.455	0.926						
5	0.456	0.406	0.937	0.854					
6	0.797	0.710	0.639	0.594	0.574				
7	0.523	0.497	1	0.926	0.937	0.639			
8	0.168	0.202	0.350	0.222	0.402	0.326	0.350		
9	0.375	0.465	0.739	0.648	0.781	0.465	0.739	0.579	
10	0.391	0.474	0.760	0.672	0.802	0.465	0.760	0.476	0.949

The results indicate a positive correlation between all of the models even though some models were more strongly correlated than others. The correlation between models 3 and 4; 3 and 5; 3 and 7; 4 and 7; 5 and 7, and 9 and 10, all show a high positive results greater than 0.9, suggesting internal validity. The correlation between models 1 and 2; 1 and 6; 2 and 6; 3 and 9; 3 and 10; 4 and 5; 5 and 9; 5 and 10; 7 and 9, and 7 and 10, also shows results to be significantly greater than zero and all greater than 0.7. The one model that shows a weak positive correlation with all of the other models at less than 0.5 is model 8. However, the specifications in this model were considerably different to the other models. It is possibly incorrect to include this model in a comparison with the other models given that it may not have been a genuinely competing alternative. Overall whilst there is strong correlation between some of the models this is not true for all of the models. Therefore these results, particularly as they are from a single time period, need to be treated cautiously.

5.16 Conclusion

What is very clear from this research is the sensitivity of the DEA models to changes in the variables used. While it has been shown that the models are relatively robust when it comes to either using full-time equivalent staff numbers or pay expenditure as labour inputs it is evident that changing other measures can have a major impact on the model results. This highlights how critical it is to ensure that relevant variables are chosen for each model. The aspect of performance being measured dictates the variables that need to be included in a model. If, for example, one is measuring the volume of care delivered then it is important to include variables relating to numbers of patients treated whereas if the quality of care is being assessed then variables such as survival rates or numbers of patients not having to return unexpectedly to hospital should be included. If the non clinical quality of care is being measured then variables such as the number of patients remaining infection free during treatment or the number of patients not complaining should be included.

As DEA measures relative efficiency the rankings of the hospitals is important. The ranking of the hospitals as set out in table 5.45 clearly shows those that are performing well and those that appear to be performing badly. Those that appear to be performing badly need to be further analysed and reviewed in conjunction with their peers to identify areas of inefficiency.

Spearman rank correlations are used to validate the models in terms of internal validity. Overall whilst there is strong correlation between some of the models this is not true for all of the models. Therefore these results, particularly as they are from a single time period, need to be treated cautiously.

Chapter 6

Conclusions

6.1 Summary

The increasing cost of healthcare is a major concern for most economies and Ireland is no different in this regard. In the period 1990 to 2006 we have seen a rapid growth in expenditure on health in the Republic of Ireland. The non-capital expenditure on the public general hospitals programme per head of population increased from €495 to €1,280 in this period. Likewise public health capital expenditure increased from €3,671 million in 1997 to €12,337 million in 2006. The upward trajectory in health costs is continuing despite the current economic downturn and is showing no signs of slowing down. As a result of this increasing emphasis is being placed on cost containment and on the delivery of greater efficiency and productivity in the health service. Given the large amount of funds being provided to hospitals it is no surprise that their use of resources and their relative performance is coming under intense scrutiny. In this environment performance needs to be clearly defined and how one measures this performance has to be understood by all stakeholders. In order to be certain that this is being achieved there also needs to be a reliable performance measurement process.

This study sought to develop a performance measurement model that incorporated relevant input and output variables. To achieve this focus groups, questionnaires and an academic expert group were employed to ascertain what stakeholders felt would be relevant to measuring performance in a hospital. This approach was used because it was felt that previous research using DEA did not generally have any process for deciding on relevant inputs and outputs, thus resulting in possible bias. The process used in this research, which was carried out in three steps, highlighted those areas that individuals felt should form part of any measurement model. The results from the process threw up some interesting trends. The first step involving four focus groups did not highlight any unexpected results with all expected inputs and outputs being mentioned. The results from the second step which was the issuing of a questionnaire to hospital stakeholders did however highlight the different perceptions of relevant performance measures held by each type of stakeholder. Not surprisingly in relation to inputs former patients rated first in order of preference nurses followed by doctors, modern equipment and finally the number of beds. Hospital staff rated modern equipment as being most relevant followed by the total number of staff, doctors, and the number of beds. It is interesting that while doctors were again rated in the top four, albeit in third place, nurses did not feature at all in the top four relevant inputs of this group. This may have been a reaction from other staff who may have felt that all staff would be important and not just nurses. Hospital staff did rate nurses in fifth place in order of relevance. The results from the senior health service officials were very different to those of the other groups. They rated in order of relevance pay costs, non pay costs, the number of beds and total costs. This clearly shows a narrow focus on finance by this group which clearly highlights their priorities. Maybe it also shows how far removed they are from the coal-face of providing health services. Interestingly the hospital directors group chose the same top four inputs as did the former patients except that they chose the doctors as being most relevant instead of the nurses. They rated nurses, number of beds and modern equipment as all being equally relevant. The third step was to establish an expert academic group to validate the inputs and outputs deemed relevant in the first two steps of the process. The same questionnaire was issued to this group and they were requested to state whether each of the measures were relevant, informative and/or necessary for the performance measurement model. A number of input and output measures were then recommended for consideration to be included in the model. Whilst the inputs that were deemed to be relevant were quantitative, measurable and available the same could not be said about the outputs. The main outputs highlighted by the various stakeholders were qualitative in nature and not generally available. They were either not being recorded or if they were, they were not available to the general public. Whilst it is extremely difficult to gauge performance levels when data on patient satisfaction is not being recorded on a national basis, it is even more disconcerting when data on patient outcomes, which are recorded, are not available to the public. This data is readily available in the UK through the NHS and there is no reason why such information is not available in the Republic of Ireland. I believe that it is unacceptable that, for example, post operative survival rates for individual surgeons are not available in the Republic of Ireland when such information is readily available in the UK and Northern Ireland. Lame arguments have been made by the medical profession for the non-release of such information, which are largely driven by fear but these fears are far outweighed when one looks at recent high profile cases. The issues that occurred in the Bristol Royal Infirmary, where the infant mortality rate for serious heart surgery was twice the national average, could still arise in this country and not be detected. Likewise the lack of data on patient outcomes could result in someone like Harold Shipman continuing to practice without being discovered. In this country we have seen Michael Neary continuing to practice as an Obstetrician whilst having an unusually high rate of both caesarean deliveries and symphysiotomies. Having high quality data collected prospectively would prevent the occurrence of these situations. The New York State Cardiac Surgery Reporting System's programme highlighted the fact that low volume surgeons had higher risk adjusted mortality than high volume surgeons did. Over the first four years of the programme twenty four low volume surgeons stopped doing cardiac surgery in New York. By 1992, the fourth year of the programme, New York had the lowest risk-adjusted mortality of any state in the USA and the most rapid rate of decline of any state with below average mortality.

Other outcome information such as the unplanned re-admission rate of patients to hospital within a short time period is not and should be available. This information could be used as a proxy in any performance measurement model for the quality of the medical service being provided in the hospital. There is no reason why this information is not available to everyone.

Due to this lack of qualitative information I was restricted in what I could use in the performance measurement model. With a lot of difficulty, however, I succeeded in getting mortality figures per hospital. This information is generally not available to the public. Information on infection rates and hygiene scores per hospital was publicly available. Inpatient and day-case data, as adjusted for casemix, and outpatient data, though unadjusted for casemix, was also available.

Qualitative feedback is also an important method for finding those people who find the results useful and anything else that might have been missed which is useful. NHS tools for assessing efficiency have been criticised as lacking sound conceptual bases and paying little attention to the needs of health service staff in understanding and using them. Hollingsworth and Parkin (2003) carried out research among 57 trusts and 14 health authorities in the Northern and Yorkshire Region of the NHS for 1994-1996. Routine data from these trusts and health authorities was used to create information on efficiency based on DEA. The trusts and health authorities were then surveyed to elicit their views on current measures of efficiency and on the potential use of the DEA-based information. The results showed that overall 80% of those surveyed gave high scores for the potential usefulness of DEA-based measures compared with between 9% and 45% for existing methods. The quality of presentation of the information was also consistently high. Using qualitative feedback analysis can be carried out on what people think is useful afterwards compared with what they thought was useful in the first place.

What is very clear from this research is the sensitivity of the DEA models to changes in the variables used. While it has been shown that the models are relatively robust when it comes to either using full-time equivalent staff numbers or pay expenditure as labour inputs it is evident that changing other measures can have a major impact on the model results. This highlights how critical it is to ensure that relevant variables are chosen for each model. The aspect of performance being measured dictates the variables that need to be included in a model. If, for example, one is measuring the volume of care delivered then it is important to include variables relating to numbers of patients treated whereas if the quality of care is being assessed then variables such as survival rates or numbers of patients not having to return unexpectedly to hospital should be included. If the non clinical quality of care is being measured then variables such as the number of patients remaining infection free during treatment or the number of patients not complaining should be included.
As DEA measures relative efficiency the rankings of the hospitals is important. Those that appear to be performing badly need to be further analysed and reviewed in conjunction with their peers to identify areas of inefficiency.
Spearman rank correlations are used to validate the models in terms of internal validity. Overall whilst there is strong correlation between some of the models this is not true for all of the models. Therefore these results, particularly as they are from a single time period, need to be treated cautiously.

In the research I introduced a process for determining the relevant input and output measures to be included in the DEA model. I also developed a model which included qualitative measures for the first time in the Irish context and tested this model for its sensitivity to different specifications. I tested ten different DEA models and showed that a high level of efficiency existed amongst Irish hospitals but that only six hospitals were on the efficient frontier in all models. I also measured scale efficiency under each model and showed those hospitals that were of optimal size as well as those that were either too big or too small. All of the models tested indicated that the majority of hospitals were either too big or too small. Clearly, while the models were based on one year’s data, the best measure for the external validity of the DEA models would be a longitudinal study.

> *Consistency over time is accepted as a test of the external validity of DEA, because, although some changes over time would be expected, it would be unlikely that these would be dramatic over succeeding years.*
>
> (Hollingsworth and Parkin, 2003: 234)

In this research I expanded knowledge and added to the academic literature on the measurement of hospital performance generally and specifically in relation to acute public hospitals in the Republic of Ireland.

6.2 Comparison with hospital casemix model results

The Irish hospital casemix model results, as set out in table 2.3, list the hospitals in order of their relative efficiency with the most efficient hospital receiving the largest monetary reward and the least efficient hospital receiving the highest monetary penalty. Comparing these results with the results from this research highlight a number of differences. These are set out in Table 6.1. The DEA model would indicate that only six hospitals were deemed to be on the efficient frontier in all ten models tested. These hospitals were St. Luke’s Hospital Kilkenny, Wexford General Hospital, Louth County Hospital, Adelaide & Meath National Children’s Hospital, St. James’s Hospital and St. Columcille’s Loughlinstown. While four of these hospitals, St. Luke’s, Wexford, Louth and St. James’s, also scored highly in the Casemix model coming in fourth, second, eight and fifth places respectively, the other two hospitals being the Adelaide & Meath and Loughlinstown scored badly, coming in twenty eight and twenty third places respectively. This highlights a significant difference between the results of both models. Similarly the Longford and Westmeath General hospital was deemed to be the most efficient hospital in the Casemix model but finished in twenty third place in the DEA model. There is clearly a performance measurement issue here.

Table 6.1

Comparison between DEA and Casemix Results

Hospital		Technical Efficiency Score	DEA Placing	Casemix Placing
U1:	Waterford Regional Hospital	0.980	22	13
U2:	St. Luke's Hospital, Kilkenny	1.000	1	4
U3:	Wexford General Hospital	1.000	1	2
U4:	South Tipperary General Hospital	0.987	13	15
U5:	Cork University Hospital	0.986	15	6
U6 :	Kerry General Hospital	0.962	26	7
U7 :	Mercy University Hospital	0.977	24	19
U8 :	South Infirmary-Victoria University Hospital	0.996	8	18
U9 :	Sligo General Hospital	0.987	13	22
U10 :	Letterkenny General Hospital	0.989	9	3
U11 :	Galway University Hospital	0.983	17	11
U12 :	Mayo General Hospital	0.981	20	12
U13 :	Portiuncula Hospital	0.989	9	17
U14 :	Mid Western Regional Hospital	0.981	20	24
U15 :	Our Lady of Lourdes Hospital Drogheda	0.977	24	27
U16 :	Louth County Hospital	1.000	1	8
U17 :	Cavan General Hospital	0.973	27	20
U18 :	Our Lady's Hospital, Navan	0.988	12	25
U19 :	Mater Hospital, Dublin	0.983	17	10
0U20 :	Beaumont Hospital, Dublin	0.999	7	9
U21 :	Connolly Hospital, Blanchardstown	0.942	28	14
U22 :	Adelaide & Meath National Children's Hospital	1.000	1	28
U23 :	Longford Westmeath General Hospital, Mullingar	0.979	23	1
U24 :	Tullamore General Hospital	0.982	19	26
U25 :	Portlaoise General Hospital	0.984	16	16
U26 :	St. James's Hospital Dublin	1.000	1	5
U27 :	St. Vincent's University Hospital Dublin	0.989	9	21
U28 :	St. Columcille's Loughlinstown	1.000	1	23

Other hospitals which show a significant difference in performance between both models are the South Infirmary-Victoria University Hospital, which finished in eighteenth place in the casemix model and finished in eighth place in the DEA model, and St. Vincent's University Hospital, which finished in twenty first place in the casemix model and finished in ninth place in the DEA model. At the same time it is clear that many of the hospitals received similar scores using either model.
The question has to be asked as to which is the more accurate model? I would argue that the DEA model more accurately reflects the efficiency of each hospital. It includes qualitative and capital measures unlike the Casemix model as well as a range of quantitative measures. It has also been tested using different mixes of variables. The Casemix model also differentiates between large Academic teaching hospitals and other hospitals in calculating the monetary penalties and rewards, which I believe questions the accuracy of some of the figures. The DEA model does not differentiate between hospital sizes and running the model using variable returns to scale takes into account any impact that increasing or decreasing returns to scale may have. The less restrictive variable returns to scale frontier allows the best practice level of outputs to inputs to vary with the size of the organisations in the sample.

6.3 Report for senior policy makers in health and hospital chief executives

In the current difficult economic climate the performance of hospitals is rightly coming intense scrutiny. What was acceptable practice during the boom years is now no longer so. All hospitals will be expected to deliver more services at a lower cost and to show that comparatively they are performing better than their counterparts. The question is how do we know that we are performing better than others and more importantly how do we show that we are doing so? This raises interesting questions about what we mean by measuring performance. Are we measuring efficiency, effectiveness, economy or quality? The answer to this question can be dependant on the target audience. The concept of performance can be viewed from a different perspective by different disciplines. An economist, an accountant, an industrial engineer, a psychologist or a hospital manager may all have a different perspective on how performance should be measured. Of this group the hospital manager takes the broadest perspective of performance which would include efficiency, effectiveness, economy and quality. A hospital may be very efficient at providing a service in the short term but may be highly dysfunctional in the long term if the quality of the service is allowed to fall. Similarly a hospital may be highly effective in achieving its targets but it may be using far too many resources in doing so. At the same time that a hospital may appear to be efficient and effective it may have serious quality issues. Achieving a high throughput of patients with great efficiency comes to nothing if an unacceptably low survival rate is the result.

As part of this research both quantitative and qualitative measures were used in the study. In order to reduce bias in the model a three step approach was adopted for the process of choosing relevant input and output measures. The first step was to establish four stakeholder focus groups who recommend relevant measures. The second step was to develop a questionnaire based on these recommendations which was sent to four groups of stakeholders for completion. The final step was to send the same questionnaire to an academic expert group for completion, asking them to state whether they felt that the measures proposed were relevant, informative and/or necessary. Resulting from this process a performance measurement model was developed, which was then tested for sensitivity using DEA with various combinations of input and output measures.

Having run ten different models the results indicated a high average level of technical efficiency across all of the hospitals at 98.5%. At the same time it was clear that there remained potential amongst many hospitals to improve their efficiencies and to make sizeable savings in both expenditure and staff levels and to increase output. Scale efficiency was also relatively high at 90.3% even though no hospital was deemed to be scale efficient across all models and the majority of hospitals were deemed to be either too big or too small. There was clearly definite scope for improvements.

A major limitation of the model was the lack of available quality and medical outcome measures. It is unacceptable that survival rates are not publicly available for each hospital and for each surgeon practising in every hospital. This information is available in the UK and Northern Ireland and should be available in the Republic of Ireland. Likewise patient satisfaction rates are not nationally available. The lack of such information weakens the performance measurement model as it omits many measures that have been deemed relevant by the users of the service. Such information is also necessary if we are to ensure that bad medical practices are avoided or discontinued. We need to ensure that what happened in the Bristol Royal Infirmary, where it is believed that over ninety children died unnecessarily, or the case of Harold Shipman where many women lost their lives, does not recur.

The performance measurement models currently being used in the Republic of Ireland are the Casemix model and Healthstat. Both of these measures perform an important function but they each have their drawbacks. While the Casemix model is a comprehensive indicator of relative efficiency across all public hospitals it does not take account of quality measures or capital used. Similarly, while healthstat does take quality and effectiveness into account it does not provide a comprehensive measure of overall efficiency for each hospital. Given that it only provides partial measures of overall performance healthstat also runs the risk of providing conflicting messages that make it more difficult to draw conclusions about an organisation's overall performance. We therefore need to take both measures into account when forming an opinion on overall performance in each hospital.

The DEA model used in this research provides a comprehensive measure of overall performance whilst also allowing for the inclusion of qualitative measures in the model. It therefore provides the opportunity of combining both the quantitative features of the Casemix model with the qualitative features of healthstat. The DEA model can readily incorporate multiple inputs and outputs and to calculate technical efficiency it only requires information on input and output quantities, not prices. Using this model, possible sources of inefficiency can be determined as well as efficiency levels. By identifying the peers for organisations that are not observed to be efficient, it provides a set of potential role models that an organisation can look to, in the first instance, for ways of improving its operations.

In the current difficult economic climate where hospitals are competing for scarce funding it is critically important that a reliable performance measurement model is used. Funding methods of hospitals are changing from block government grants to a fee per patient system where only the most efficient hospitals will survive. The DEA performance measurement model is a reliable method of measuring hospital efficiency that can assist hospitals in highlighting their inefficiencies, which they can then correct, and in identifying peers that they can then try to emulate.

6.4 Implications of this research for the measurement of productivity

This research has implications for our thinking about the measurement of productivity and performance in general and particularly in the health services. The research highlights the need to ensure that only those areas that are deemed to be relevant should be used in any performance measurement model. The focus groups and questionnaire outcomes showed clearly the differing views of stakeholders and academic experts on what should be measured and what determines a good performance. They also showed that different performance measures may be appropriate for different target audiences.

The implications of this research for health services in Ireland are:

(1) The results of the focus groups and questionnaires clearly showed that the outcome measures deemed most relevant by the stakeholders are not being measured in the Republic of Ireland. Health policymakers need to focus on qualitative measures that provide information on health outcomes to the public. Hospital performance measurement must focus on these measures.
(2) The results of the focus groups and questionnaires also highlighted the wide divergence in views between different groups of stakeholders as to what were the most relevant input and output measures that should be included in any performance measurement model. This is an important factor that should be borne in mind when addressing the needs of specific groups.
(3) Having health outcome information collected prospectively could ensure that situations such as those that arose at the Bristol Royal Infirmary and with Harold Shipman in the UK and with Michael Neary in Ireland would not occur. The success of the New York Cardiac Reporting System's programme is proof of this.
(4) The research highlights the need to develop a comprehensive performance measurement model for hospitals in Ireland. The DEA model used in this research is one such model. The current Casemix model being used only focuses on quantitative measures and does not take account of qualitative measures or capital used. The comparison of the results in table 5.1 highlights the differences found between both models. The Healthstat model that is also used provides some qualitative partial measures but it is not a comprehensive measure.
(5) What the research also highlights is the lack of debate in Ireland regarding the measurement of performance, not just in hospitals, but in the wider health sector. There is a minimal amount of academic literature available on the subject and this knowledge deficit needs to be addressed.

6.5 Limitations of the research

There are a number of limitations that need to be borne in mind when interpreting the results of this research. They are:

(1) Outpatient activity numbers have not been adjusted for casemix. While casemix data on inpatient and day case attendances are publicly available

this is not the case for outpatient attendances. This omission may negatively affect the results for hospitals with a high proportion of complex outpatients and may positively impact on those hospitals that do not have such complex cases.

(2) The non-availability of qualitative measures seriously restricts our ability to design a comprehensive performance measurement model. It was not possible to include many output measures in the model that were deemed to be relevant measures by the various stakeholders.

(3) The use of survival rates as a proxy measure for quality of service. While this measure has been used widely as a proxy measure for quality it does have its critics. For example the survival rate in very dependant on the medical specialties in the hospitals and as has been highlighted by Jarman et al. (1999) on the number of emergency medical admissions to the hospital and the ratio of doctors to beds.

(4) The use of bed numbers as a proxy for capital employed. This measure has also been widely used as a proxy for capital employed even though it is clearly not the most accurate measure. It is however simple to measure and given the difficulties involved in measuring capital employed it is generally accepted as a reasonable proxy.

(5) The choice of input and output variables. While the process used for specifying the inputs and outputs was valid and reliable the variables that were finally used in the model were dependant on available data. Many output variables that were deemed to be relevant by the stakeholders were not available and thus could not be included in the model. There was therefore potential for specification error where important variables may have been omitted.

(6) DEA assumes no measurement error. There is a likelihood that some error existed.

6.6 Need for further research

I would recommend that future research should examine the following areas:

(1) A longitudinal study of Irish public hospitals examining the impact of the economic downturn on their efficiency levels.

(2) An examination of the location or environmental factors that are beyond the hospital manager's control that are impacting on the efficiency levels of Irish public hospitals.
(3) A longitudinal study of Irish public hospitals to see what factors contributed to some hospitals improving their efficiency levels over time and others not.
(4) A study comparing the efficiency levels of public and private Irish hospitals.
(5) The use of more qualitative measures in performance measurement models.
(6) A study of hospitals that are performing best with a view to developing a guide to best practice for all hospitals.
(7) When comparing the results from this research with the results from the Casemix model currently in use in the Republic of Ireland a number of significant differences were evident. The source of these differences should be further investigated.

6.7 Final Conclusions

It is clear from the results obtained from the focus groups, the questionnaires and the academic expert group that each stakeholder group has differing opinions on what are the most relevant input and output measures that need to be included in any performance measurement model. This observation, though it may appear obvious, is of critical importance in that it does highlight the difficulty in designing a measurement model that is acceptable to all parties. It is clear that any performance measurement model must be designed within parameters that are clearly understood by any intended audience.

The lack of publicly available qualitative information in Ireland is unacceptable. It would be in the interests of the general public and the health care system that information on patient medical outcomes, mortality rates, patient satisfaction rates and other qualitative measures would be publicly available in this country, as they are from the National Health Service throughout the UK. The non-availability of this data makes it extremely difficult to develop a comprehensive performance measurement model that would include all relevant output measures.

Based on available qualitative and quantitative data the results indicated a high level of technical efficiency among the public acute hospitals in Ireland. The average technical efficiency score was 98.5%. At the same time there existed scope for the hospitals to make savings and increase output. Some hospitals that were not technically efficient had a greater potential to make further efficiencies. The DEA model highlighted the peers for these inefficient hospitals. Only five hospitals were

deemed to be technically efficient in all of the models tested. These best performing hospitals ranged in size from 118 to 842 beds, thus indicating that hospital size did not unduly impact on efficiency levels. No hospital was deemed to be scale efficient in all of the models even though the average scale efficiency for all of the hospitals was relatively high at 90.3%. The results of all of the models tested indicated that the majority of the hospitals were either too big or too small.
When comparing the results from this research with the results from the Casemix model currently in use in the Republic of Ireland a number of significant differences were evident. These differences could call into question the accuracy of the Casemix model. This is an important finding given that a portion of public hospital funding in Ireland is based on these Casemix results.

This research achieved its objectives of developing a comprehensive model for measuring hospital performance and using this model to measure the performance of public acute hospitals in Ireland in 2007. The research has provided detailed information on the technical and scale efficiency of the hospitals both individually and overall and has added to current literature on performance measurement.

Arising from this research the main recommendations that I would make would be that information on medical outcomes, survival rates, patient satisfaction and other relevant qualitative information should be made publicly available in Ireland; that hospitals should focus on improving their technical and scale efficiencies, and that performance measurement models should be developed that would include more qualitative data.

References

Aaron, H.J. and Ginsburg, P.B. (2009). Is health spending excessive? If so, what can we do about it? Health Affairs, 2009, Vol. 28(5): 1260-1275.

Akazili, J., Adjuik, M., Jehu-Appiah, C. and Zere, E. (2008). Using data envelopment analysis to measure the extent of technical efficiency of public health centres in Ghans. BMC International Health and Human Rights, 8: 11.

Al-Shammari, M. (1999). A multi-criteria data envelopment analysis model for measuring the productive efficiency of hospitals. International Journal of Operations and Production Management 19: 879-890.

Amaratunga, D., Haigh, R., Sarshar, M., and Baldry, D. (2002). Application of the balanced score-card concept to develop a conceptual framework to measure facilities management performance within NHS facilities. International Journal of Healthcare Quality Assurance, Bradford: Vol. 15, iss. 4/5: pp. 141-152.

Ansoff, I.H. (1985). Strategic response in a turbulent environment. Handbook of Business Strategy, Guth, W., New York: Warren, Gorham and Lamont.

Athanassopoulos, A. and Gounaris, C. (2001). Assessing the technical and allocative efficiency of hospital operations in Greece and its resource allocation implications. European Journal of Operational Research, 133: 416-431.

Balk, W.L. (1975, March-April). Technological trends in productivity measurement. Public Personnel Management, pp. 128-133.

Banker, R.D., Charnes, A., and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment analysis. Management Science, 30(9): pp. 1078-1092.

Banker, R.D., Charnes, A., Cooper, W., Swarts, J. and Thomas, D. (1989). An introduction to data envelopment analysis with some models and their uses. Research in Governmental and Non-Profit Accounting, 5: 125-163.

Banker, R., Gadh, V. and Gorr, W. (1993). A Monte Carlo comparison of two production frontier methods: Corrected ordinary least squares and data envelopment analysis. European Journal of Operational Research, 1993: 67: 332-343.

Bannick, R.R. and Ozcan, Y.A. (1995). Efficiency analysis of federally funded hospitals: comparison of DoD and VA hospitals using data envelopment analysis. Health Services Management Research, 8: 73-85.

Belcher, J.G., Jr. (1982). The productivity management process. Oxford: Planning Executives Institute.

Blaxter, L., Hughes, C. and Tight, M. (1996). How to Research. Open University Press, Bucks.

Blendon, R.J., Minah, K. and Benson, J.M (2001). The Public Versus the World Health Organisation on Health System Performance. Health Affairs 2001; 20: 10-20.

Brennan Commission (2003). Report of the commission on financial management and control systems in the Health Service, Government Stationery Office, Dublin.

Bristol Royal Infirmary Report (2001). The Report of the Public Inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995. Bristol Inquiry Unit.

Bullock, R.J., and Batten, D.B. (1983). Organizational productivity: A measurement review. Paper presented at the meeting of the Southwest Division of the Academy of Management, Houston, TX.

Burgess, J.F. and Wilson, P.W. (1996). Hospital ownership and technical efficiency. Management Science, 42: 110-123.

Campbell, D.T. and Stanley, J.C. (1963). Experimental and quasi-experimental designs for research on teaching. In N.L. Gage, ed., Handbook of Research on Teaching. Chicago: Rand McNally.

Campbell, J.P. (1977). On the nature of organizational effectiveness. In P.S. Goodman, J.M. Pennings, & Associates (Eds.), New perspectives on organizational effectiveness. (pp. 13-55). San Francisco: Jossey-Bass.

Campbell, J.P. and Campbell, R.J. (1988c). Industrial-organisational psychology and productivity: The goodness of fit. In J.P. Campbell and R.J. Campbell (Eds.), Productivity in organisations (pp. 82-94). San Francisco: Jossey-Bass.

Carter, N. (1991). Learning to measure performance: the use of indicators in organisations. Public Administration, 69: 85-101.

Casu, B., Shaw, D. and Thanassoulis, E. (2005). Using a group support system to aid input-output identification in DEA. Journal of the Operational Research Society, 56: pp. 1363-1372.

Chang, H.H. (1998). Determinants of hospital efficiency: the case of central government-owned hospitals in Taiwan. Omega-International Journal of Management Science, 26: 307-317.

Chang, H., Cheng, M-A. and Das, S. (2004). Hospital ownership and operating efficiency: evidence from Taiwan. European Journal of Operational Research, 159: 513-527.

Charnes, A., Cooper, W.W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), pp. 429-444.

Chen, A., Hwang, Y.C. and Shao, B. (2005). Measurement and sources of overall and input inefficiencies: evidence and implications in hospital services. European Journal of Operational Research, 161: 447-468.

Chern, J.-Y. and Wan, T.T. (2000). The impact of the prospective payment system on the technical efficiency of hospitals. Journal of Medical System, 24: 159-172.

Chirikos, T.N. and Sear, A.M. (2000). Measuring hospital efficiency: a comparison of two approaches. Health Services Research, 34: 1389-1408.

Codman, E.A., A study in hospital efficiency (1917). Vol. 53. 1996, Oakbrook Terrace, Illinois: reprinted by the Joint Commission on Accreditation of Healthcare Organisations.

Coelli, T. (1996). A guide to DEAP version 2.1: Data Envelopment Analysis Composite Program, Centre for Efficiency and Productivity Analysis, CEPA Working Paper 9608, University of New England.

Coelli, T., Rao, D. and Battesse, G. (1998). An introduction to Efficiency and Productivity Analysis, Boston, Kluwer Academic Publishers.

Coelli, T., Rao, D.S.P., O'Donnell, C. J. and Battesse, G. (2005). An Introduction to Efficiency and Productivity Measurement. New York: Springer.

Commission on Health Funding (1989). Report of the Commission on Health Funding. Government Stationery Office, Dublin.

Commission for Health Improvement (2002). Patient and public perspectives on health care performance. Report prepared for the Commission for Health Improvement. Department of Public Health, University of Oxford.

Connolly, T., Conlon, E.J., and Deutsch, S.J. (1980). Organizational effectiveness: A multiple constituency approach. Academy of Management Review, 5: pp. 211-217.

Cook, T.D. and Campbell, D.T. (1979). Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally.

Cook, R.D. and Weisberg, S. (1982). Residuals and influences in regression. London: Chapman and Hall.

Coulter, P.B. (1979). Organizational effectiveness in the public sector: The example of municipal fire protection. Administrative Science Quarterly, 24(1): pp. 65-81.

Craig, C.E., and Harris, R.C. (1973). Total productivity measurement at the firm level. Sloan Management Review, 14(3): pp. 13-28.

Dalmau-Matarrodona, E. and Puig-Junoy, J. (1998). Market structure and hospital efficiency: evaluating potential effects of deregulation in a national health service. Review of Industrial Organisation, 13: 447-466.

Debreau, G. (1951). The coefficient of resource utilisation. Econometrics, 19: pp. 273-292.

Denison, E.F. (1984). Productivity analysis through growth accounting. In A.P. Brief (Ed.), Productivity research in the behavioral and social sciences (pp. 7-55). New York: Praeger.

Denzin, N.K. and Lincoln, Y.S. (1998). Strategies of Qualitative Inquiry. Sage Publications, California.

Department of Health, (2002), Delivering the NHS Plan. Crown copyright.

Department of Health and Children, (2001), Quality and Fairness, A Health System for You. Government Stationery Office, Dublin.

Department of Health and Children, (2004). The Modernisation of the National Casemix Programme in Ireland. Government Stationery Office, Dublin.

Department of Health and Children, (2007), Health in Ireland, Key Trends 2007. Government Stationery Office, Dublin.

Department of Health and Children, (2009), Annual Report. Government Stationery Office, Dublin.

Deprez, F.L. (1986). Office productivity. Information Services and Use, 6: pp. 83-102.
Dittman, D.A., Capettini, R. and Morey, R.C. (1991). Measuring efficiency in acute care hospitals: An application of data envelopment analysis, J Human Res Admin, 14: 89-108.

Drummond, M., Sculpher, M, Torrance, G., O'Brien, B and Stoddart, G. (2005). Methods for the Economic Evaluation of health care programmes. Oxford University Press: Oxford, UK.

Dyson, R.G., and Thanassoulis, E. (1988). Reducing weight flexibility in data envelopment analysis. Journal of Operational Research Society, 39(6), pp. 563-576.

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S. and Shale, E.A. (2001). Pitfalls and protocols in DEA. European Journal of Operational Research, 2001, 132(2): 245-259.

Fare, R., Grosskopf, S. and Logan, J. (1983). The relative efficiency of Illinois electric utilities. Resources and Energy 5: 349- 367.

Fare, R., Grosskpf, S., Norris, M. and Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialised countries. American Economic Review, 84: 66-83.

Farrell, M.J. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society 120: 253-281.

Ferrier, G.D. and Valdmanis, V.G. (1996). Rural hospital performance and its correlates. Journal of Productive Analysis, 7: 63-80.

Ferrier, G.D. and Valdmanis, V.G. (2002). Exploring psychiatric hospital performance using data envelopment analysis and cluster analysis. Journal of Economic Medicine, 20: 143-153.

Ferrier, G.D. and Valdmanis, V.G (2004). Do mergers improve hospital productivity? Journal of Operational Research Society, 55: 1071-1080.

Folland, S.T. and Hofler, R.A. (2001). How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production. Health Economics, 10: 683-698.

Fried, H.O., Lovell, C.A., Schmidt, S.S., Yaisawarng, S. (2002). Accounting for environmental effects and statistical noise in data envelopment analysis. Journal of Productivity Analysis, 17: 157-174.

Friesner, D., Roseman, D. and McPherson, M.Q. (2008). Are hospitals seasonally efficient? Evidence from Washington State. Applied Economics, 40: 699-723.

Gabris, G.T., Mitchell, K., and McLemore, R. (1985). Rewarding individual and team productivity: The Biloxi merit bonus plan. Public Personnel Management, 14(3): 231-244.

Gannon, B. (2005). Testing for variation in technical efficiency of hospitals in Ireland. The Economic and Social Review, Dublin, Vol. 36, 3: pp. 273-295.

Giuffrida, A. (1999). Productivity and efficiency changes in primary care: a Malmquist index approach, Health Care Management Science, 2: 11-26.

Goddard, M. and Smith, P. (2002). Some Principles of performance measurement and performance improvement, Centre for Health Economics, University of York.

Gregan, T. and Bruce, R. (1997). Technical Efficiency in the hospitals of Victoria, in 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.

Grosskopf, S. and Valdmanis, V. (1987). Measuring hospital performance – a non parametric approach. Journal of Health Economics 6: 89-107.

Grosskopf, S. and Valdmanis, V. (1993). Evaluating hospital performance with casemix adjusted outputs. Medical Care 31: 525-532.

Grosskopf, S., Margaritis, D. and Valdmanis, V. (2001). Comparing teaching and non-teaching hospitals: a frontier approach, teaching vs. non-teaching hospitals. Health Care Management Science 4: 83-90.

Gruca, T.S. and Nath, D. (2001). The technical efficiency of hospitals under a single payer system: the case of Ontario community hospitals. Health Care Management Science, 4: 91-101.

Guzzo, R.A., Jette, R.D., and Katzell, R.A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta analysis. Personal Psychology, 38: pp. 275-291.

Guzzo, R.A. (1988). Productivity research: Reviewing psychological and economic perspectives. In J.P. Campbell and R.J. Campbell (Eds.), Productivity in organisations (pp. 63-81). San Francisco: Jossey-Bass.

Hajialiafzali, H., Moss, J.R. and Mahmoud, M.A. (2007). Efficiency measurement for hospitals owned by the Iranian Social Security organisation. Journal of Medical Systems, 31: 166-172.

Hansmann, H.B. (1980). The role of Non-profit Enterprise. The Yale Law Journal, 89: 835-901.

Harris, J., Ozgen, H. and Ozcan, Y. (2000). Do mergers enhance the performance of hospital efficiency? Journal of the Operational Research Society, 51: 801-811.

Harrison, J.P., Coppola, M.N. and Wakefield, M (2004). Efficiency of Federal hospitals in the United States. Journal of Medical Systems, 28: 411-422.

Hatam, N., Moslehi, S., Askarian, M., Shokrpour, N., Keshtkaran, A. and Abbasi, M. (2010). The efficiency of general public hospitals in Fars province, Southern Iran. Iranian Red Crescent Medical Journal, 12(2): 138-144.

Hollingsworth, B. and Parkin, D. (1998). Developing efficiency measures for use in the NHS, a report to the NHS Executive Northern and Yorkshire Region R&D Directorate. Newcastle: Health Economics Group, School of Health Sciences, University of Newcastle.

Hollingsworth, B. and Parkin, D. (2001). The efficiency of the delivery of neonatal care in the UK. Journal of Public Health Medicine, 23: 47-50.

Hollingsworth, B. and Parkin, D. (2003).Efficiency and productivity change in the English National Health Service: can data envelopment analysis provide a robust and useful measure? Journal of Health Services Research and Policy, 2003 (8): 230-236.

Hollingsworth, B. and Street, A. (2006). The market for efficiency analysis of health care organisations. Health Economics, 15: 1055-1059.

Hollingsworth, B. and Peacock, S. (2008). Efficiency Measurement in Health and Health Care. New York: Routledge.

Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. Health Economics, 17(10): 1107-1128.

Horton, R. (2001). The real lessons from Harold Frederick Shipman. Lancet 357: 82-83.

Hurst, E.G. (1980). Attributes of performance measures. Public Productivity Review, 4(1), 43-50.

Ilgen, D.R., and Klein, H.J. (1988). Individual motivation and performance: Cognitive influences on effort and choice. In J.P. Campbell and R.J. Campbell (Eds.), Productivity in organisations (pp. 143-176). San Francisco: Jossey-Bass.

Jacobs, R. (2000). Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. Discussion Paper 177, Centre for Health Economics, University of York.

Jacobs, R., Smith, P. and Street, A. (2006). Measuring Efficiency in Health Care: Analytic Techniques and Health Policy. Cambridge.

Jarman, B., Gault S., Alves, B., Hider, A., Dolan, S., Cook, A., Hurwitz, B., and Iezzoni, L.I. (1999). Explaining differences in English hospital death rates using routinely collected data. BMJ 318: pp. 1515-1520.

Kanji, G.K. (1998). Measurement of business excellence. Total Quality Management, 9(7): pp. 633-643.

Kanji, G.K. and Sa, P.M. (2002). Kanji's Business Scorecard. Total Quality Management, 13(1): pp. 13-27.

Kaplan, R., and Norton, D. (1992). The balanced scorecard: measures that drive performance. Harvard Business Review, January/February: pp. 71-79.

Keeley, M. (1978). A social-justice approach to organizational evaluation. Administrative Sciences Quarterly, 23: pp. 272-292.

Kendrick, J.W. (1977). Understanding productivity. Baltimore: Johns Hopkins University Press.

Kendrick, J.W. (1984). Improving company productivity. Baltimore: Johns Hopkins University Press.

Kerr, C.A., Glass, J.C., McCallion, G.M. and McKillop, D.G. (1999). Best practice measures of resource utilisation for hospitals: a useful complement in performance assessment, Public Administration, 77: 639-650.

Kirigia, J.M., Emrouznejad, A. and Sambo, L.G. (2002). Measurement of technical efficiency of public hospitals in Kenya: using data envelopment analysis. Journal of Medical Systems, 26: 39-45.

Kirigia, J.M. et al. (2008). A performance assessment method for hospitals: the case of municipal hospitals in Angola. Journal of Medical System, 32: 509-519.

Kontodimopoulos, N. et al. (2006). Balancing efficiency of health services and equity of access in remote areas of Greece. Health Policy 76: 49-57.

Kopelman, R.E. (1986). Managing productivity in organisations: A practical, people-oriented perspective. New York: McGraw-Hill.

LeCompte, M.D. and Goetz, J.P. (1982). Problems of reliability and validity in ethnographic research. Review of Educational Research, 52; 31-36.

Lee, K.-H. et al. (2008). The association between hospital ownership and technical efficiency in a managed care environment. Journal of Medical System, 33: 307-315.

Li, T. and Rosenman, R. (2001). Cost inefficiency in Washington hospitals: a stochastic frontier approach using panel data. Health Care Management Science, 4: 73-81.

Liu, X. and Mills, A. (2005). The effect of performance related pay of hospital doctors on hospital behaviour: a case study from Shandong, China. Human Resources for Health 3:11.

Linna, M. (1998). Measuring hospital cost efficiency with panel data models. Health Economics, 7: 415-427.

Linna, M. and Hakkinen, U. (1999). Determinants of cost efficiency of Finnish hospitals: a comparison of DEA and SFA. Working paper, National Research and development centre for welfare and health.

McCallion, G.M., McKillop, D.G., Glass, J.C. and Kerr, C. (1999). Rationalising Northern Ireland hospital services towards larger providers: best- practice efficiency studies and current policy. Public Money Management 27-32.

Mahoney, T.A. (1988). Productivity defined: The relativity of efficiency, effectiveness and change. In J.P. Campbell & R.J. Campbell (Eds.), Productivity in organisations (pp. 13-38). San Francisco: Jossey-Bass.

Magnussen, J. (1996). Efficiency measurement and the operationalisation of hospital production. Health Services Research, 31: 21-37.

Mali, P. (1978). Improving total productivity. New York: Wiley.

Maniadakis, N., and Thanassoulis, E. (1998). Changes in productivity of a sample of Scottish hospitals: A cost index approach. Research Paper No. 288, Research Bureau, Business School, University of Warwick, Coventry.

Maniadakis, N., Hollingsworth, B. and Thanassoulis, E. (1999). The impact of the internal market on hospital efficiency, productivity and service quality. Health Care Management Science, 2: 75-85.

Maniadakis, N., and Thanassoulis, E. (2000). Assessing productivity changes in UK hospitals reflecting technology and input prices. Applied Economics, 32: pp. 1575-1589.

Masiye, F. (2007). Investigating health system performance: an application of data envelopment analysis to Zambian hospitals. BMC Health Services Research 7: 58.

Maxwell, J.A. (1992). Understanding and validity in qualitative research. Harvard Educational Review, 62; 279-300.

Maxwell, J.A. (1996). Qualitative Research Design: An Interactive Approach. Thousand Oaks, California: Sage.

Mobley, L.R. and Magnussen, J. (1998). An international comparison of hospital efficiency: does institutional environment matter? Applied Economics, 30: 1089-1100.

Muckler, F.A. (1982). Evaluating productivity. In M.D. Dunnette and E.A. Fleishman (Eds.), Human performance and productivity: Vol.1. Human capability assessment (pp. 13-47). Hillsdale, NJ: Erlbaum.

Murray, C.J.L., Kawabata, K. and Valentine, N (2001). Peoples Experience Versus Peoples Expectations. Health Affairs (Millwood) 2001; 20(3): 21-24.

National Institute for Health and Clinical Excellence (2008). Guide to the methods of technology appraisal. publications@nice.org.uk
National Task Force on Medical Staffing (2003). Report of the National Task Force on Medical Staffing, Government Stationery Office, Dublin.

Nightingale, F. (1863). Notes on hospitals, 3rd edition. London: Longman, Green, Longman, Roberts and Green.

Norman, R.G. and Bahiri, S. (1972). Productivity measurement and incentives. London: Butterworths.

Ozcan, Y.A. (1992). Sensitivity analysis of hospital efficiency under alternative output/input and peer groups: A review, Int J Knowledge Trans Utilis, 5: 1-29.

Ozcan, Y.A., McCue, M.J. and Okasha, A.A. (1996). Measuring the technical efficiency of psychiatric hospitals. Journal of Medical Systems, 20: 141-150).

Parker, D. and Newbrander, W. (1994). Tackling wastage and inefficiency in the health sector. World Health Forum, Vol.15: 107-113.

Parkin, D. and Hollingsworth, B. (1997). Measuring production efficiency of acute hospitals in Scotland, 1991-1994: validity issues in data envelopment analysis. Applied Economics, 29:1425-1433.

Pedraja-Chaparro, F., Salinas-Jiminez, J. and Smith, P. (1999). On the quality of the data envelopment analysis model. Journal of the Operational Research Society, 50: 636-644.

Pennings, J.M., and Goodman, P.S. (1977). Towards a workable framework. In P.S. Goodman, J.M. Pennings, & Associates (Eds.), New perspectives on organizational effectiveness (pp. 146-184). San Francisco: Jossey-bass.

Perakyla, A. (1997). Reliability and Validity in research based on Tapes and Transcripts, Ch. 13 in Silverman, D. (Ed) (1997) Qualitative Research: Theory, Method and Practice. Sage Publications, London.

Preziosi, R.C. (1985). Productivity improvement training takes preparation. Training and Development Journal, 39(2):pp. 101-102.

Prior, D. (1996). Technical efficiency and scope economies in hospitals. Applied Economics, 28: 1295-1301.

Pritchard, R.D., Jones, S.D., Roth, P.L., Stuebing, K.K., and Ekeberg, S.E. (1988). The effects of feedback, goal setting, and incentives on organizational productivity. Journal of Applied Psychology Monograph Series, 73(2): pp. 337-358.

Pritchard, R.D., Jones, S.D., Roth, P.L., Stuebing, K.K., and Ekeberg, S.E. (1989). The evaluation of an integrated approach to measuring organizational productivity. Personnel Psychology, 42(1): pp. 69-115.

Pritchard, R.D. (1990). Organizational Productivity. In M.D. Dunnette & L.M. Hough (1992) Handbook of Industrial and Organizational Psychology (2nd. edition) Volume 3: pp. 443-471.

Prospectus Consultants (2003). Audit of structures and functions in the health system, on behalf of the Department of Health and Children, Government Stationery Office, Dublin.

Puig-Junoy, J. (2000). Partitioning input cost efficiency into its allocative and technical components: an empirical DEA application to hospitals. Socio-Economic Planning Sciences, 34: 199-218.

Richardson, J., Wildman, J. and Robertson, I.K (2003). A critique of World Health Organisation evaluation of health system performance. Health Economics 2003; 12(5): 355-366.

Robson, C. (2002). Real World Research. Second Edition, Blackwell Publishing.

Rosow, J.M. (Ed.) (1981). Productivity: Prospects for growth. New York: Van Nostrand, Reinhold.

Saretsky, G. (1972). The OEO P.C. experiment and the John Henry effect. Phi Delta Kappan, 53(9): 579-581.

Schmidt, P. and Lin, T. (1984). Simple tests of alternative specifications in stochastic frontier models. Journal of Econometrics, 24: 349-361.

Schmidt, P. (1985). Frontier production functions. Econometric Reviews, 4: 289-328.

Schneider, B. (1984). Industrial and organizational psychology perspective. In A.P. Brief (Ed.), Productivity research in the behavioral and social sciences (pp. 174-206). New York: Praeger.

Shetty, Y.K., and Buehler, V.M. (1985). Productivity through people: Practices of well-managed companies. Westport, CT: Quorom Books.

Silver, M.S. (1984). Productivity Indices: Methods and applications. Aldershot, England: Gower Publishing.

Smith, P.C. (1997). Model mis-specification in data envelopment analysis. Annals of Operations Research, 73: 233-252.

Smith, P.C. (1998). Data envelopment analysis in health care: An introductory note. Centre for Health Economics, University of York.

Spottiswoode, C. (2000). Improving police performance: a new approach to measuring police efficiency. London: Public Services Productivity Panel, HM Treasury.

Thanassoulis, E., Boussofiane, A. and Dyson, R.G. (1995). Exploring output quality targets in the provision of peri-natal care in England using data envelopment analysis. European Journal of Operational Research, 80: 588-607.

Thanassoulis, E. (2001). Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software, Dordrecht: Kluwer Academic.

Tuttle, T.C. (1981). Productivity measurement methods: Classification, critique and implications for the Air Force (AFHRL-TR-81-9). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Tuttle, T.C. (1982). Measuring productivity and quality of working life. National Forum, 62(2): pp. 5-7.

Tuttle, T.C. (1983). Organizational productivity: A challenge for psychologists. American Psychologist, 38: pp. 479-486.

Valdmanis, V. (1992). Sensitivity analysis for DEA models: an empirical example using public versus private NFP hospitals. Journal of Public Economics, 48: 185-205.

Valdmanis, V., Kumanarayake, L. and Lertiendumrong, J. (2004). Capacity in Thai public hospitals and the production of care for poor and non-poor patients. Health Services Research 39: 2117-2134.

Wanless, D. and the Health Trends Review team at HM Treasury (2002). Securing our Future Health: Taking a Long-Term View. London: HM Treasury.

Wei, C.-K. (2006). Measuring efficiency and productivity change in Taiwan hospitals: a nonparametric frontier approach. Journal of American Academy of Business, Cambridge, 10(1): 317-323.

Werther, W.B., Ruch, W.A., and McClure, L. (1986). Productivity through people. New York: West.

West, M., King, E., Dawson, J and De Chermont, K. (2006). Teams at the level of the organisation: Organisational outcomes of team-based working.

Wessa, P. (2011), Free statistics software, Office for Research Development and Education.

World Health Organisation (1999). Health 21. Health for all in the 21st century. An Introduction. Copenhagen: World Health Organisation.

World Health Organisation (2000).The World Health Report 2000. Health systems: improving performance, Geneva: World Health Organisation.

Zelman, W.N., Pink, G.H., and Matthias, C.B. (2003). Use of the balanced scorecard in health care. Journal of Health Care Finance, New York: Summer, Vol. 29, Iss. 4: p. 1.

Zere, E., Mbeeli, T., Shangula, K., Mandlhate, C., Matirua, K., Tjivambi, B. and Kapenambili, W. (2006). Technical efficiency of district hospitals: evidence from Namibia using Data Envelopment Analysis. Cost effectiveness and Resource Allocation, 4: 5.

Zuckerman, S., Hadley, J. and Iezzoni, L. (1994). Measuring hospital efficiency with frontier cost functions. Journal of Health Economics, 13: 255-280.

Appendices

Appendix 1

Casemix Adjustments 2009

Source: Casemix: H.I.P.E.
Health Service Executive

Appendix 2

Patient Questionnaire

Questionnaire

Dear Sir/Madam,

As someone who was recently a patient at the South Infirmary-Victoria University Hospital I would be grateful if you would take a few moments to complete the following questionnaire and return it in the enclosed stamped addressed envelope. We are always aiming to improve our services and I would hope that the information that we will receive would help to improve our performance measurement systems within the hospital.

1. When measuring hospital performance which of the following do you think are relevant? (Tick as many items as you wish)

(a) Number of doctors ( )

(b) Number of nurses ( )

(c) Number of Radiographers ( )

(d) Number of support staff ( )

(e) Total number of staff ( )

(f) Number of beds ( )

(g) Modern equipment ( )

(h) Drug costs ( )

(i) Pay costs ( )

(j) Non-pay costs ( )

(k) Total costs ( )

2. Are there any other things that you think are important that are not included above in question 1? Please write them in below:

…………………………………………………………………………………………
……
…………………………………………………………………………………………
……
…………………………………………………………………………………………
……
…………………………………………………………………………………………
……
…………………………………………………………………………………………
……
…………………………………………………………………………………………
……
…………………………………………………………………………………………
……

3. When measuring hospital performance which of the following do you think are important? (Tick as many items as you wish)

(a)	Hygiene	( )
(b)	Approval by professional bodies of the standards of the hospital	( )
(c)	The food	( )
(d)	Inpatient waiting times	( )
(e)	Numbers of people who are treated without having to stay in hospital overnight	( )
(f)	Health & safety within the hospital	( )
(g)	Length of waiting lists	( )
(h)	Patients referred to the hospital when they do not need hospital treatment	( )
(i)	How quickly patients are treated	( )
(j)	How quickly patients are released from hospital after their treatment	( )
(k)	Patient satisfaction	( )
(l)	Outpatient waiting times	( )
(m)	Infection levels at the hospital	( )
(n)	The amount of time for which operating theatres are available for operations	( )
(o)	Total patient numbers	( )
(p)	How easy it is to get to the hospital	( )

(q)	The number of new patients attending the outpatient's department	( )
(r)	The number of patients returning for further outpatient appointments	( )
(s)	The number of patients having to return to the hospital unexpectedly	( )
(t)	The number of patients who die at the hospital following treatment	( )
(u)	Time waiting to be seen by a doctor in the accident & emergency department	( )
(v)	Time taken by hospital staff to answer phone calls	( )
(w)	The ability of the hospital to operate within its financial budget	( )
(x)	Staff courtesy	( )
(y)	Staff communications with patients and their families	( )
(z)	Car parking facilities	( )
(a1)	Staff friendliness	( )
(b1)	Nursing care	( )
(c1)	Level of complaints	( )

4. Are there any other ways of assessing our performance as a hospital that you think are not included above in question 3? Please write them in below:

……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

Appendix 3

Letter to Patients

Office of Chief Executive,
South Infirmary-Victoria University Hospital,
Old Blackrock Road,
Cork,
Ireland.

Aston University
Illustration removed for copyright restrictions

9th. September, 2008

Measuring Hospital Performance

Dear Sir/Madam,

As someone who was recently a patient at the South Infirmary-Victoria University Hospital I would be grateful if you would take a few moments to complete the following questionnaire and return it in the enclosed stamped addressed envelope. This should take less than 10 minutes. All completed and returned questionnaires received by the 30th September, 2008 will be entered in a draw for a meal for two in The Barn Restaurant, Cork.

We are always aiming to improve our services and I would hope that the information provided by this questionnaire would help to improve our performance measurement systems within the hospital. This information will be used by me as part of a research project that I am completing at Aston University, Birmingham that is looking at how hospital performance is measured and how this might be improved.

<u>Please note that participation in this study is voluntary.</u>

All data will be treated confidentially and in accordance with the Irish Data Protection Act 1988 and the UK Data Protection (Amendment) Act 2003.

I would like to thank you for your co-operation. If you have any questions please do not hesitate to contact me at the above telephone number or e-mail address.

Yours faithfully,

Gerard O'Callaghan

Appendix 4

Letter to Expert Group

Office of Chief Executive,
South Infirmary-Victoria University Hospital,
Old Blackrock Road,
Cork,
Ireland.

6th. March, 2009

Measuring Hospital Performance

Dear Sir/Madam,

I am currently undertaking the DBA programme at Aston University, Birmingham. My research is based on measuring hospital performance. The focus of my research is twofold. Firstly it aims to add to existing theory on organisation performance measurement by developing a new measurement model and secondly it aims to apply this new measurement model to the measurement of the performance in the 37 acute hospitals in Ireland.

My proposed methodology involves a number of steps. I will be using the Data Envelopment Analysis (DEA) model as a measurement method. The choice of the unit of assessment and the identification of the inputs and outputs are critical to this model. In order to ensure that the most relevant inputs and outputs are used in the model I proceeded in three steps. The first step was to establish four focus groups, the second step was to circulate questionnaires to the main stakeholders and now the third step is to write to experts, like yourself, to validate the inputs and outputs that have been deemed to be relevant. This will take only 10 minutes of your time.

Below are the inputs and outputs deemed to be relevant at the first two stages of the process and I would be grateful if you would, as an expert in this area, validate the results. I would also be interested in whether you believe any relevant input or output measures have been omitted from the model.

- Could you please indicate next to each input or output measure whether or not you think it is a good measure of hospital performance i.e., it is relevant, informative and necessary as a component of a measure of hospital performance.
- Do this by putting a tick or cross in the appropriate category opposite each input or output. (Clearly this judgement will depend on how the measure is operationalised but assume for now that the operationalisation is reasonably accurate).
- Can you also add at the end any inputs or outputs you feel should be included which are not?

Inputs

	Relevant	Informative	Necessary
Number of doctors			
Number of beds			
Number of nurses			
Modern equipment			
Number of staff			
Total costs			
Number of support staff			
Non pay costs			
Pay costs			
Number of radiographers			
Drug costs			

Outputs

	Relevant	Informative	Necessary
Hygiene			
Staff communications with patients and their families			
Infection levels at the hospital			
Health and Safety within the hospital			
Patient satisfaction			
How quickly patients are treated			
Nursing care			
Length of waiting lists			
Time taken to be seen by a doctor in the Accident and Emergency department			
Inpatient waiting times			
Approval by professional bodies of the standards of the hospital			
Staff courtesy			
The ability of the hospital to operate within its financial budget			
Outpatient waiting times			
Number of people who are treated without having to stay in the hospital overnight			
Level of complaints			
Staff friendliness			
How quickly patients are released from hospital after their treatment			
The food			
The amount of time for which operating theatres are available for operations			
The number of patients having to return to hospital unexpectedly			
Number of new patients attending the Outpatients' department			
Number of patients who die at the hospital following treatment			
Number of patients returning for further outpatients' appointments			
Total patient numbers			
Car parking facilities			

Time taken by hospital staff to answer phone calls			
Patients referred to the hospital when they do not need hospital treatment			
How easy it is to get to the hospital			

Please note that participation in this study is voluntary.

All data will be treated confidentially and in accordance with the Irish Data Protection Act 1988 and the UK Data Protection (Amendment) Act 2003.

I would like to thank you for your co-operation. If you have any questions please do not hesitate to contact me at the above telephone number or e-mail address.

Yours faithfully,

Gerard O'Callaghan

Appendix 5

Total Hospital Bed Numbers

Aston University

Illustration removed for copyright restrictions

Source: Health Service Executive