HEALTH ECONOMICS
Health Econ. **17**: 1107–1128 (2008)
Published online 13 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hec.1391

THE MEASUREMENT OF EFFICIENCY AND PRODUCTIVITY OF HEALTH CARE DELIVERY

BRUCE HOLLINGSWORTH*

Centre for Health Economics, Faculty of Business and Economics, Monash University, Melbourne, Australia

SUMMARY

The measurement of efficiency and productivity of health service delivery has become a small industry. This is a review of 317 published papers on frontier efficiency measurement. The techniques used are mainly based on non-parametric data envelopment analysis, but there is increasing use of parametric techniques, such as stochastic frontier analysis. Applications to hospitals and other health care organizations and areas are reviewed and summarised, and some meta-type analysis undertaken. Cautious conclusions are that public provision may be potentially more efficient than private, in certain settings. The paper also considers conceptualizations of efficiency, and points to dangers and opportunities in generating such information. Finally, some criteria for assessing the use and usefulness of efficiency studies are established, with a view to helping both researchers and those assessing whether or not to act upon published results.

Received 19 April 2007; Revised 18 April 2008; Accepted 3 June 2008

KEY WORDS: efficiency measurement; productivity; health care; parametric; non-parametric; stochastic frontiers; data envelopment analysis

1. INTRODUCTION

As stated in a recent editorial in *Health Economics* 'Applied academic research into efficiency has grown into a thriving industry' (Hollingsworth and Street, 2006). The editorial concludes that the supply side of the market for efficiency analyses in health care is impressive, while the demand side is weak. In this paper I ask several questions including: What do we (as health economists) really mean by efficiency? What do we mean by outcome measurement, in terms of whether hospitals and doctors improve the mental and physical functioning of patients? Is it possible to move towards a 'gold standard' of practice for carrying out efficiency studies, or is this impossible at present? What are the problems we still have to overcome?

Looking at the supply side, the paper reviews published applications of efficiency measurement in health care. This encompassed a systematic search of all available and relevant databases. As in previous reviews in this area (Hollingsworth *et al*., 1999; Hollingsworth, 2003) published papers are reviewed with a view to determining methods and data used, models specified, sensitivity analysis employed, results and policy implications. In addition, results are summarised in a basic form of meta-analysis in order to synthesise results and cautiously draw out potential implications.

Previously I have hypothesised that much work undertaken and published in this area is of the 'have software—will analyse' nature (Hollingsworth, 2003). The danger of this is that over-interpreted data on 'efficiency' may lead to policy decisions based on potentially unreliable information, with potentially

*Correspondence to: Centre for Health Economics, Faculty of Business and Economics, Monash University, Building 75, Clayton Campus, Melbourne, Victoria 3800, Australia. E-mail: bruce.hollingsworth@buseco.monash.edu.au

disastrous consequences. Additionally, given the increase in data available, there may now be pressure which means 'have data—must analyse', with little thought to model specification, a critical area that is difficult to test statistically in frontier efficiency measurement. There is a positive way forward and health economists can help policy makers actually improve efficiency, provided this is an objective of the health system. As such, I suggest some non-exclusive *Use and Usefulness* criteria, to help researchers decide how to undertake useful efficiency measurement research, and to help policy makers decide whether they should actually make use of the results they are presented with.

There is, up to mid-2006, a published literature consisting of 317 journal papers and book chapters. This helps to set in place, robust foundations and guidelines for an agenda for future research in this area. As an appendix to this paper, study findings are summarised, which will hopefully be a useful resource in itself.

2. BACKGROUND

What do we mean by efficiency? Based on the seminal definitions of efficiency by Farrell (1957), technical efficiency is producing the maximum amount of output from a given amount of input or alternatively, producing a given output with minimum input quantities, such that when a firm is technically efficient, it operates on its production frontier. Allocative efficiency occurs when the input mix is that which minimises cost, given input prices or alternatively, when the output mix is that which maximises revenue, given output prices. Technical and allocative efficiency comprise 'overall efficiency'. When a firm is efficient overall, it operates on its cost or revenue frontier.

We can illustrate these efficiency concepts by considering the simple case of a single output (y) being produced from two inputs, X_1 and X_2 , see Figure 1 (from Hollingsworth *et al.*, 1999). The production function (or frontier) represents the maximum output produced from all input combinations and, in general, is: $y = f(X_1,X_2)$. Initially assume the production function is linearly homogeneous —Farrell assumed constant returns to scale but differing returns are possible. The efficient unit isoquant, $y = 1$ in Figure 1, shows the technically efficient input combinations used to produce a unit of output. Suppose that the actual observed input–output combination is at P, with input mix (X_1^0, X_2^0) and unit output $y = 1$. Production at P is technically inefficient since the firm could produce output $y = 1$ employing the same input mix but using the input quantities at point R on the isoquant. Therefore, technical efficiency, TE, at P is: $TE = OR/OP$ $(0 < TE \leq 1.)$

If $TE = 1$, the firm is on the efficient isoquant and is technically efficient; and when $TE < 1$, the firm is technically inefficient and the more inefficient the unit, the smaller the TE. A firm may also be cost-minimising. With given relative factor prices, shown in Figure 1 by the isocost line ab, the optimal (cost-minimising) input mix to produce $y = 1$ is at Q. If the unit at P is technically efficient that is, operating at R, its cost is represented by the isocost line cd, which is above minimum cost ab. As such, at its observed input mix, unit P needs to use input quantities that correspond to point S to deliver a unit of output at minimum cost. Therefore, allocative (or price) efficiency, AE, is: $AE = OS/OR$ $(0 < AE \leq 1)$. The overall cost of producing at Q relative to P is the measure of overall (economic or productive) efficiency, OE, which is the product of technical and allocative efficiency, that is: $OE = OS/OP = OR/OP \text{x} OS/OR$ $(0 < OE \leq 1)$.

This initial Farrell analysis is static but efficiency can be measured over time, i.e. the frontier may shift due to technological advances. Productivity is defined as the ratio of an index of output to an index of input usage. Change over time of this measure is productivity change. Initially economists attributed productivity changes to technological changes, i.e. shifts of the production or cost frontier. However, following Nishimuzu and Page (1982) it became increasingly accepted that productivity change can also be caused by efficiency change, that is, by shifts over time of firms relative to their frontier, and recently productivity measurement has incorporated efficiency measurement (see Grosskopf 1993, and

DOI: 10.1002/hec

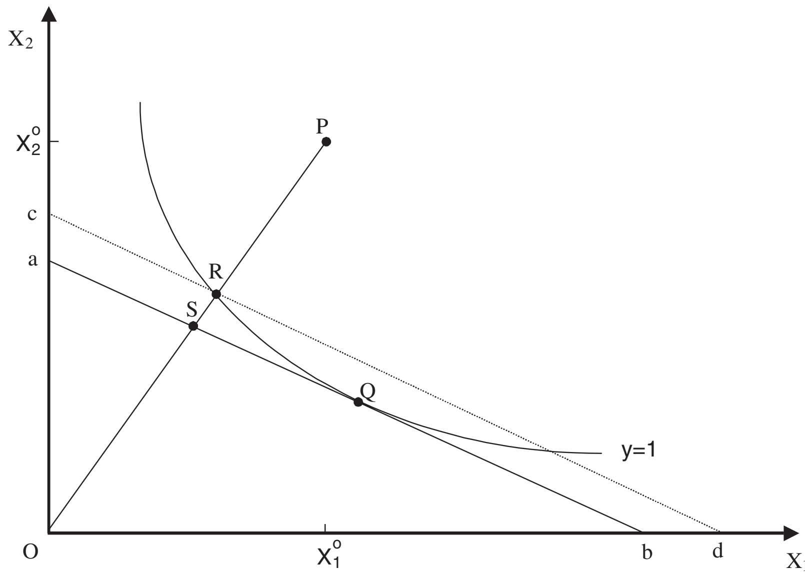

Figure 1. Radial efficiency measurement

Hollingsworth *et al*., 1999 for more detail; and Färe *et al*., 1997 for an explanation of the difference between catching up with a frontier, and how a frontier moves.).[1]

Ideally when measuring efficiency in health, we would measure health gains of individual patients, i.e. the *final* output, and we may be getting closer to being able to actually do this, given datasets linking individuals to their health care usage, now being collected in several countries. Most research published so far has used some variant of *intermediate* outputs, in terms of numbers of patients treated. Even adjusted for case mix this is not ideal, as it does not tell us whether the patient's health has improved. This critique was summarised in Newhouse (1994) which in turn is fully discussed in Jacobs *et al*. (2006). Jacobs *et al*. conclude that a cautious approach is desirable, but just because something is difficult to measure, we should not give up. They conclude by suggesting one way forward may be the use of multivariate models, which treat different objectives as part of a system of equations, but allow for correlations across equations (Hauck and Street, 2006), and the logical extension of multivariate, multi-level models, which allow for the fact that health care improvements may only be one of several objectives of health service actors. These objectives may include quality. Smith (2002) argues that in the UK, the pressure to be efficient has led to a low cost, low quality system, but this may be a result of equity considerations as well.[2]

Given the nature of what health economists mean by efficiency, it is important to establish how efficiency has been measured in health care, and in the following section published applications are reviewed with a view to summarising the methods and data used, and models specified, partly in order

[1]Considering the tradeable and non-tradeable nature of health care goods is also relevant. How can the value (or price) of an untradeable good be measured? This may help determine which method is used, for example, non-parametric measurement may be used if there is a problem of valuing (or weighting) non-tradeable inputs and outputs.

[2]It is outside the scope of this paper to present a theory of the relationships, incentives and potential trade offs between quantity, quality and equity. That would take us back to the argument of what is *best* - low cost, low quality, highly equitable, or high cost, high quality, but inequitable (see Rice, 2003; Culyer, 2005, 2006, for a further discussion of these issues – both efficiency and equity considerations are important for finding optimal solutions). In terms of ethics, as in Culyer (1992), I generally assume 'use of the term 'efficiency' can be as…ethical as the objective which is sought efficiently to be attained.' If access to health care is a right of citizenship, and maximising health given resources is the objective of health services, then it can be ethical to be efficient.

DOI: 10.1002/hec

to establish some initial criteria for assessing why research in this area has had limited use in policy terms.

3. APPLICATIONS

Since the early 1980s, efficiency analysis has been used to measure and analyse the productive performance of health care services. The number of studies has increased dramatically over the past few years (Hollingsworth and Street, 2006; Hollingsworth, 2003). The principal techniques used, based on data envelopment analysis (DEA) and stochastic frontier analysis (SFA) have been extensively summarised elsewhere in general terms (Cooper *et al*., 2004; Coelli, 2005) and in health care in particular (Hollingsworth *et al*., 1999; Jacobs *et al*., 2006). This paper will not go over the methods again, but points the interested reader to these references for more information.

One aim of this paper is to update previous reviews of methods and results of studies which measure efficiency and productivity of health services using efficiency measurement, a literature, up to and including papers published (and available) by mid-2006, of 317. Just under 70 per cent of these studies publish quantifiable scores that can be analysed.[3] The rapid increase in published studies over recent years is evident from Figure 2; 55 per cent have been published since 2000.

DEA alone is used in 48 per cent of studies, see Figure 3, a further 19 per cent of studies use DEA in some form of secondary regression, even though there are doubts about the validity of this (Simar and Wilson, 2007).[4] Non-parametric analysis is used in over 80 per cent of frontier efficiency analyses, and most studies are still cross sectional in nature. Malmquist studies are used in 8 per cent of studies, and SFA and other parametric frontier techniques are used in 18 per cent of studies, showing an increase over the last 5 years.

Fifty-two per cent of applications are in hospitals, see Figure 4. Analyses have been undertaken using data from over 30 countries (excluding cross country analyses, which make up 4 per cent of applications). This may perhaps reflect a lowering of barriers in terms of data and software availability (Hollingsworth and Street, 2006).

Most studies use *output* (or throughput) measures of physical performance, such as inpatient days or discharges. There is some use, in 9 per cent of studies, of *outcome* measures examining changes in health status, mortality or quality of care for individuals treated. In *output* terms, there is a tendency towards using inpatient days rather than inpatient cases, even though in general use of some form of case mix adjusted outcome may give a more accurate measure of intermediate output. Input variables are mainly measures of staff and capital employed, and most analysis is of technical efficiency. Only a small number of studies test methods such as weight restricted models and analysis of returns to scale or use statistical or sensitivity analysis of results.

3.1. DEA applications in health care

Initially overall measures of efficiency in hospitals are concentrated on, before going on to examine the non-hospital literature.

3.2. The hospital literature

Details of hospital efficiency studies can be found in the appendix and these show the type of hospital, country, number of hospitals in the sample, author(s) and efficiency scores. Most studies report results

[3] A list of those references not specifically mentioned here can be obtained by contacting the author.
[4] Inference may be invalid due to complicated and unknown serial correlation among the estimated efficiencies.

DOI: 10.1002/hec

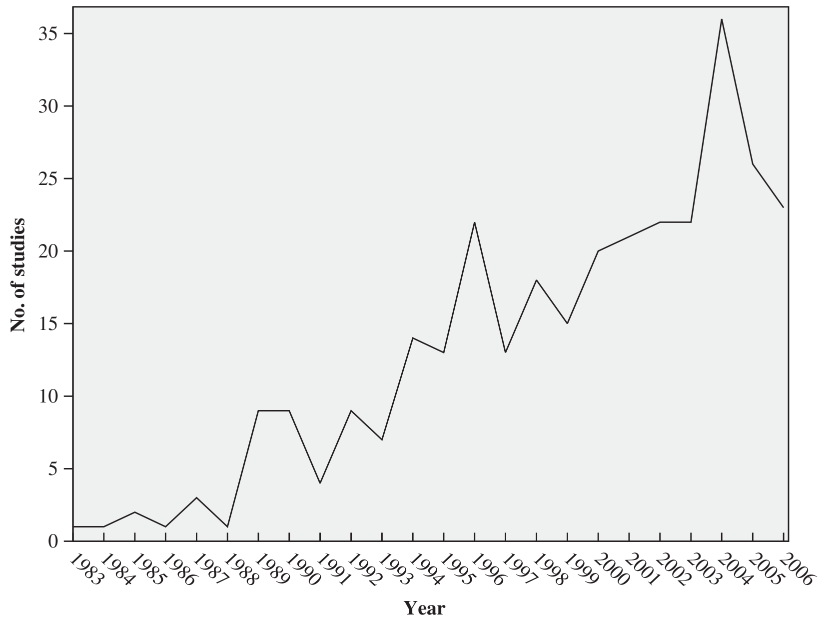

Figure 2. Number of efficiency studies 1983–2006

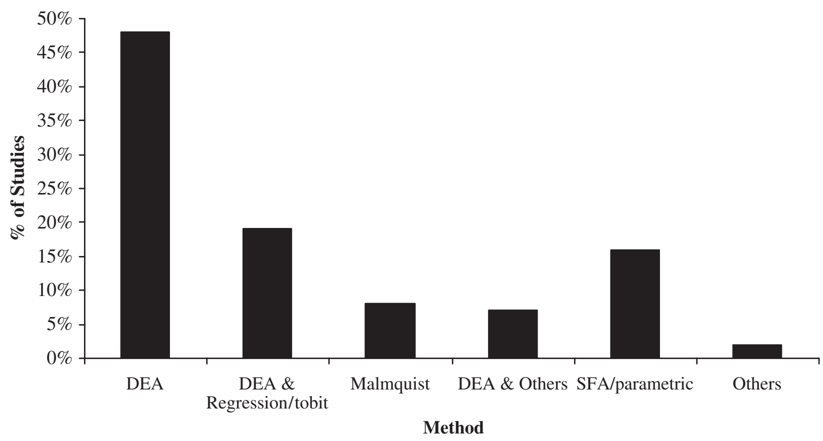

Figure 3. Methods used in reported studies

from the USA and can be categorised into different types of hospitals. The main division is made between public and private provision.[5] The public providers include Department of Veterans Affairs (VA), Department of Defense (DOD) and other Federal units.

[5]Ownership definitions used here are: public - state owned/run firms; for profit - privately run; not-for profit - in some cases are voluntary/charity run firms which serve the poor. However, health care not-for-profit firms obtain 90 per cent of revenue from sales and receipts, are privately run, are entitled to many tax exemptions and advantages, make a residual surplus and compete with for-profit hospital firms.

DOI: 10.1002/hec

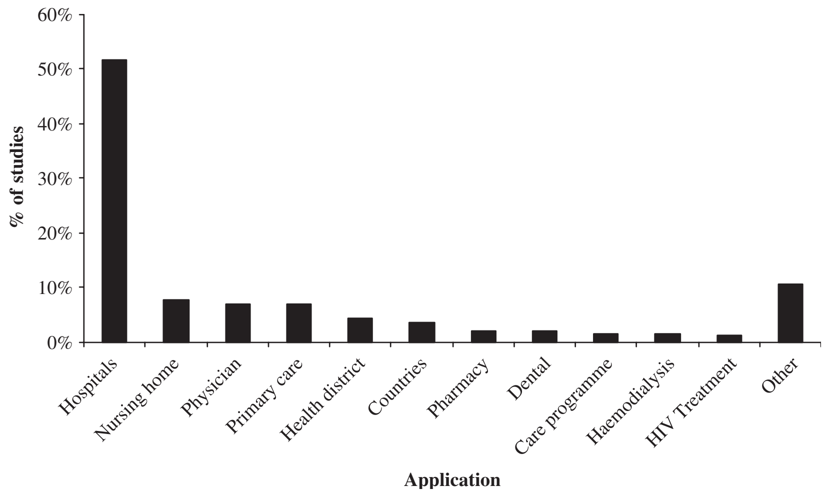

Figure 4. Areas of application

Table I. Summary statistics for hospital efficiency scores

	No.	Mean	Median	St. Dev.	Minimum
For Profit	4	0.801	0.855	0.130	0.61
Not-for-profit	12	0.825	0.86	0.109	0.60
Public	17	0.881	0.906	0.094	0.689
Defense/VA	6	0.885	0.895	0.056	0.82
Non-teaching	2	0.742	0.742	0.046	0.71
Teaching	3	0.673	0.65	0.087	0.60
Acute/general	27	0.843	0.865	0.086	0.65
Non specified	26	0.839	0.81	0.104	0.61
Psychiatric	2	0.600	0.60	0.183	0.47
All hospitals	99	0.835	0.85	0.107	0.47
USA Hospitals	59	0.826	0.85	0.115	0.47
Euro. Hospitals	28	0.860	0.872	0.084	0.72
Non USA/EU	12	0.839	0.861	0.121	0.61

The summary statistics are shown in Table I, and a box-plot of the efficiency scores by hospital category is shown in Figure 5. The mean efficiency across the whole sample is 0.835 (excluding the within-hospital studies) and the median is 0.85.

Figure 5 summarises the results for each hospital type. The box-plot shows the median, quartiles and extreme values for each hospital group. This allows us to see at a glance which hospital groups are more efficient and the range of scores. Comparing efficiency across the sector, defense/VA hospitals (which are public in nature) have the highest mean efficiency (0.885) with a median score of 0.895, compared with not-for-profit (generally private) hospitals which have a lower mean efficiency (0.825) and a lower median score (0.86). Public hospitals also have a high mean score (0.881) and a high median score (0.906), compared with not-for-profit hospitals. Not-for-profit firms treat most hospital patients in the USA (Folland *et al*., 2001) and these results reiterate earlier comparisons (Hollingsworth, 2003). Examination of the standard deviations and minimums demonstrate the room for efficiency gain. For not-for-profit hospitals the standard deviation is 0.109 and the minimum 0.60, demonstrating considerable deviation from the mean of 0.825 and so room for improvement. Potential efficiency gains are similar for public hospitals (standard deviation 0.094, minimum score 0.689 and mean of 0.881), and

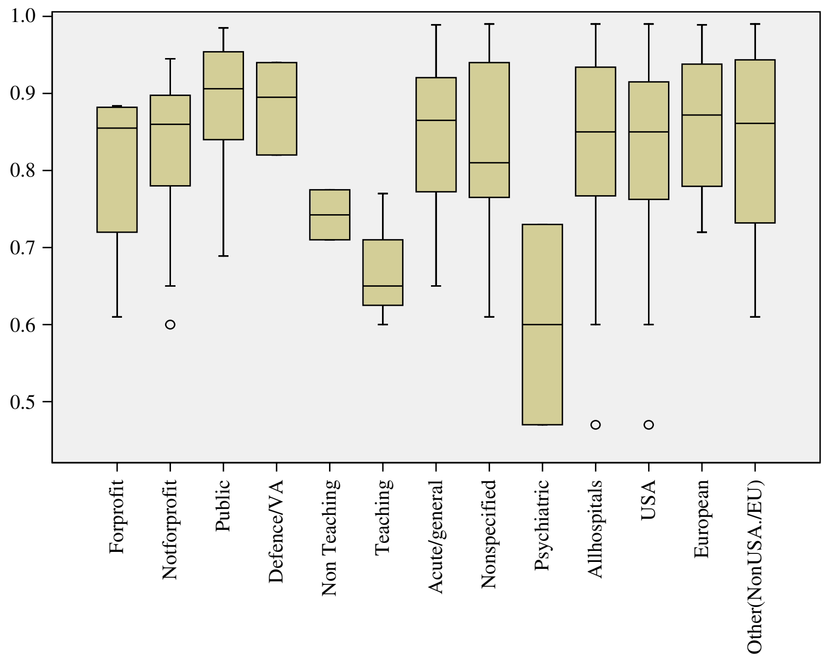

Figure 5. Box-plot of distribution of efficiency scores by category of hospital

less for defense/VA hospitals (standard deviation 0.056, minimum 0.82 and mean of 0.885). There is also some potential for gain for Acute/general hospitals (standard deviation 0.086 and a minimum of 0.65), deviating from the mean of 0.84. The small sample of studies specifically examining teaching in hospitals shows non-teaching units to be more efficient (0.742 compared with 0.673).

A further way to compare efficiency is to compare the efficiency of hospitals across countries. This also gives some indication as to the efficiency of different means of health care delivery. Most studies are from the USA where the average efficiency is 0.826, with a median of 0.85 and a minimum of 0.47. Here, the system is predominantly one of privately provided health care insurance, with a safety net of Medicaid and Medicare to cover the poor and elderly, respectively. This can be compared with Europe where health care is characterised by public provision or social insurance. In the European sample the average efficiency is 0.86, with a median of 0.872 and a minimum of 0.72. These results are slightly higher than those for the sample of USA hospitals, where there is some potential for efficiency gain, with a standard deviation of 0.115 and a minimum of 0.47, compared with 0.084 and 0.72 for the European sample. There are an increasing number of studies outside the EU and USA, especially in developing countries. However, the merits of comparing efficiency in developing countries has been debated recently in terms of the usefulness of frontier measures in this context (Asian Development Bank, 2006).

The results, both that public provision seems in general more efficient, and that European hospitals have higher average efficiency, may reflect many confounding factors, including methodological differences between studies, differences in models (which are rarely justified in terms of economic theory) or sample sizes (which are often small) impacting upon robustness and validity. Results are conditional upon basic differences in study design and samples, rather than any real variation in efficiency meaning, it is difficult to compare results beyond looking at generalities. It may also be the case that public and private enterprises have different objectives (see the work of, for example, Pestieau and Tulkens, 1993; Marchand *et al*., 1984), i.e. they operate with respect to different technological

DOI: 10.1002/hec

frontiers and simply cannot be compared, for example the objective may actually be to provide a different level of service, for example a public cheap, but basic, dental service, relative to a private more costly, but non-universal service.

3.3. The general health literature

There are several other health care areas in which DEA has been applied. Details of the general health studies can be found in the appendix and show the type of organisation, country, sample size, author(s) and efficiency scores. The summary statistics are shown in Table II and a box-plot of the distribution of efficiency scores in Figure 6.

Examination of the statistics in Table II and the box-plot of the distribution of scores in Figure 6 demonstrate that there is potential for efficiency gain. For Health Districts there is room for improvement, both in Europe and the USA (means of 0.839 and 0.742 and minimums of 0.80 and 0.5, respectively). There is also scope for efficiency gain in primary care where in Europe the mean is 0.821 compared with the USA mean of 0.712 where there is also greater potential for improvement (standard deviation 0.229 and minimum of 0.390). However, this may reflect the differences in service delivery in the USA and Europe. A more valid comparison is of nursing homes which in the USA seem less efficient, compared with those in Europe (means 0.765 and 0.821, medians[6] 0.81 and 0.83), whereas both demonstrate potential for improvement, with the minimum scores of 0.38 and 0.70 and standard deviations of 0.158 and 0.114. Previously, it has been reported that for-profit-homes appear more efficient than not-for–profits (Hollingsworth, 2003; Anderson *et al.*, 2003).

One interesting observation in the non-hospital literature is: what are the objective functions of the units of observation (for example health care area organisations), and who has managerial responsibility for efficiency? What type of inefficiency is being measured in a 'non-firm' setting? The nature and validity of cross country efficiency measurement has also been questioned (Spinks and Hollingsworth, 2005), with several theoretical and modelling concerns.

3.4. Malmquist productivity applications

A summary of Malmquist based productivity studies is provided in the appendix. As with the SFA section which follows, summaries of results are provided from a selection of more recent studies,[7] as direct meta-type analysis is not appropriate due to small numbers, and methodological incompatibility. Half of these more recent studies are conducted on samples of hospitals in the EU. Some highlights include Sola and Prior (2001) who look at quality, finding change driven by technology. Ferrari (2006) finds a small productivity improvement in UK hospitals of 2 per cent (with technological improvements, but inefficiency increases). Lyroudi *et al.* (2006), look at efficiency changes in clinics in 10 Greek public hospitals, finding efficiency improvement. Maniadakis and Thanassoulis (2004) also look at a sample of Greek hospitals, interestingly including a decomposition into cost inefficiency. Chen (2006) who looks at Taiwanese hospitals, decomposes quality (in terms of doctors and nurses per department), finding a productivity decline (2.7 per cent), with a change in the overall sample outweighing individual efficiency improvements.

The studies do seem to reflect the cutting edge of research in the area of efficiency analyses, perhaps because software packages are not as available 'off the shelf' for this type of analyses, keeping numbers of studies low, but perhaps more innovative in nature.

[6]Analysis in this way is one way to demonstrate that DEA can be very sensitive to outliers.
[7]Highlights of earlier Malmquist and SFA studies are in Hollingsworth (2003).

DOI: 10.1002/hec

Table II. Summary statistics for general health efficiency scores

	No.	Mean	Median	St. Dev.	Minimum
Care programme	2	0.623	0.623	0.032	0.60
Health Districts Euro.	4	0.839	0.838	0.04	0.80
Health Districts USA	9	0.742	0.80	0.114	0.50
Nursing Homes Euro.	6	0.821	0.83	0.114	0.70
Nursing Homes USA	19	0.765	0.81	0.158	0.38
Primary Care Euro.	6	0.821	0.815	0.104	0.675
Primary Care USA	9	0.712	0.827	0.229	0.390
Primary Care (non USA/EU)	2	0.79	0.79	0.014	0.780

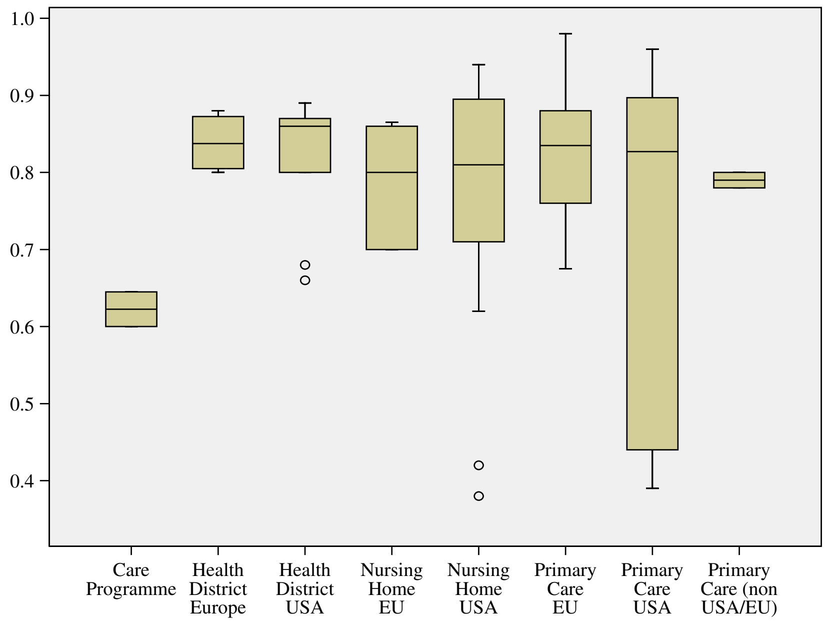

Figure 6. Box-plot of distribution of efficiency scores by general health

3.5. (SFA) and other parametric applications

A summary of studies using SFA and other parametric techniques is provided in the appendix. Of the more recent studies, most are conducted on samples of USA hospitals, a smaller number on EU hospitals and EU nursing homes.

Some highlights include Deily and McKay (2006) who include quality, in terms of in hospital mortality, finding a mean efficiency, using a hybrid function, in US urban acute hospitals of 87 per cent. Yaisawarng and Burgess (2006) account for access and quality in VA hospitals, finding efficiency of around 94 per cent. Rosko (2004) also accounts for quality in USA hospitals, but finds it has a minimal impact. McKay (2002/2003) estimates efficiency on a very large sample of 4 075 US hospitals, finding not for profits most efficient (86 per cent), followed by government (85 per cent) and for-profits (84 per cent). Linna *et al*. (2006) account for quality in Finnish hospitals, in mortality terms, finding efficiency of around 82 per cent. Gannon (2005) tries different functional forms (including Cobb Douglas and Translog) on a sample of Irish hospitals, finding large degrees of inefficiency. Street (2003) compares SFA and corrected ordinary least squares (COLS) in UK hospitals, and concludes choice of technique impacts upon efficiency rankings.

DOI: 10.1002/hec

Beyond hospital applications, Kathuria Sankar (2005) look at public health in Indian States, looking at the infant mortality rate, and the impact upon this of various factors including literacy, income, water availability and health care, modelling fixed and random effects. Several studies look at nursing homes in the EU, including Crivelli *et al*. (2002) who use a translog function for Swiss nursing homes, finding ownership not to be significant. Swiss nursing homes are also looked at by Farsi and Filippini (2004), who find private homes to be more efficient than public. Laine *et al*. (2005b) look at institutional elderly care in Finland, including quality indicators, finding a mild association between some aspects of quality and technical efficiency.

As with the Malmquist studies, the relatively low number of applications may reflect the complexity of analyses involved with SFA. However, increasing numbers of applications may be a consequence of inclusion of SFA in the better known econometric software packages and the increased potential for use with panels of data.

4. USE AND USEFULNESS CRITERIA FOR SUPPLIERS AND DEMANDERS

4.1.Suppliers

Is there a way of undertaking research in this area which would make the supply of studies more effective? In other words, are there specific criteria or guidelines, which would make efficiency measurement more user friendly? Here I establish some initial non-exhaustive criteria as a starting point, in both macro and micro terms. By *macro* I mean the overall process of undertaking the study in terms of set up and management, in a way to help ensure that the information provided will be of use in policy terms. By *micro* I mean the actual production of the efficiency scores.

Macro issues include:

1. Applied research needs to be placed in a policy context. One important element of any efficiency analyses is to get potential end users involved early on. This helps ‘ownership’ of the research from the users perspective and keeps the researcher on track. This may initially involve finding the right person or group of people (having a number of people involved reduces risks, e.g. staff moving positions). Meetings to feedback results at various stages and to different levels of users, e.g. hospital managers, health department staff, will help make sure information is provided to those who want to use it. An advisory group to initially help set up model specification may be useful.
2. Hospital managers may have concerns about health authorities using efficiency measures as ‘big sticks’ and are generally interested in more detailed information on their specific unit. Health authority staff tend to be more interested in the overall picture and comparisons between hospitals. The researcher has to balance these views, and providing all the information to everyone may help. Also ask what information it would be useful to provide that the data/modelling is not providing right now, try and accommodate this, or suggest means (e.g. extra data) which could help. Visit your model specification frequently, and provide sensitivity analyses. What value are you adding to the way efficiency is measured already?
3. Have you given your end users the information you set out to? Surveying them, perhaps including a short report, may help refine your measures. Disseminate your results as widely as possible. Make sure users know the limitations of efficiency measures; they are *a* useful policy tool, not *the* useful policy tool. Results can be manipulated so full provision of information to all may be helpful.

Micro issues include:

1. Are you asking the right questions?

DOI: 10.1002/hec

2. What is your underlying economic theory of production (or cost, does duality theory and the requirement for cost minimisation as an objective really apply).
3. Is your model specified correctly? Have you undertaken extensive sensitivity analysis? Ask your advisory group if there are any obvious omitted variables.
4. Are your data really good enough to answer the questions, particularly your output data?
5. Have you any data on quality? What will results using just quantity (throughput) data really show? Will any inefficiency be just made up of omitted quality data?
6. If you do have quality data, how will you weight it relative to quantity data, to avoid it being 'swamped' by relatively large numbers of throughput information? Unless carefully weighted, potentially vital information on quality may have little impact on results.
7. Is your sample inclusive enough, are you comparing like with like? Exploratory analyses are useful, just because all hospitals in your sample have the sample categorisation, there may be a rogue specialist unit or teaching hospital in there which will confound your results, as frontier techniques are very susceptible to outliers. Sample size is also an issue.
8. If you are happy with your data and models, what techniques will you use, parametric, non-parametric or both? If you have multiple inputs/outputs non-parametric techniques have an advantage (when comparing DEA and SFA) in terms of dis-aggregation.[8] They allow you to feedback more detailed information on areas of inefficiency. Panel data techniques will also allow you to feedback more information, not only on what happens between units, but what happens over time. Looking at trends over time is more useful than a snap shot.
9. Are you undertaking two stage analyses? If so how are you accounting for any statistical problems (see Simar and Wilson, 2007)?
10. Do you need to generate confidence intervals? Unless you are certain your sample is all inclusive, then you might wish to account for sampling variation.[9]

4.2. Demanders

In Table III I have suggested a check-list for assessing if an efficiency analysis should be made use of. This (again) is a starting point, based on the Drummond *et al.* (2005) list for assessing economic evaluations. Suppliers of efficiency studies may also wish to take note of these points.[10] The two assessment questions asked by Drummond *et al.* (see their chapter 3) 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? As Drummond *et al.* acknowledge, it is unlikely that every study can fulfil every criterion, but criteria are useful as screening devices to identify strengths and weaknesses of studies, and of course to identify the value added by comprehensive extra analysis of this nature.

5. SUMMARY AND CONCLUSIONS

The number of studies which seek to measure health service efficiency and productivity continues to increase quite dramatically. Research in this area should be reviewed carefully and the results of studies interpreted and used cautiously, as it is still an area which is under development. The outlook is

[8] A single output stochastic production frontier can be adapted to the multiple output case, making use of distance functions, however endogeneity of regressors may be a problem. There is a growing technical literature in the area of multiple output distance functions, see for example, Kumbhakar and Lovell (2000); Coelli *et al*, (2005).

[9] See Coelli *et al*. (2005: pp 202–203) for a discussion about concerns with sampling distributions, i.e. you are *measuring* the frontier when you have all the hospitals in a country, but *estimating* the frontier if not.

[10] I am talking about applied efficiency measurement here. I am not suggesting technical advances are second best forms of research, and health data are sometimes a great way of demonstrating a technical advance. I also acknowledge the pressure to publish, see Hollingsworth and Street (2006) for a discussion of this.

DOI: 10.1002/hec

Table III. A check-list for assessing efficiency measurement studies[11]

1. Is the question well defined, and answerable?
- Are the inputs and outputs clear?
- Is there a particular viewpoint stated (whose objectives are accounted for–managers, Government policy makers, patients?), is any decision making context established?

2. Is a comprehensive description of the sample given?
- Can you tell if any relevant comparator units are excluded?
- Are the samples strictly comparable, are there potential outliers?

3. Are the quality and quantity output data clear and comprehensive?
- Where do the data come from, who collected them, and why?
- Are quantity data case mix adjusted?
- Are quality data useful, e.g. can individual patients be followed through the system?

4. Are all the relevant inputs and outputs included?
- Is the range wide enough to answer the research question?
- Do they cover all relevant viewpoints (e.g. hospital mortality may be of interest to patients, scale of operation to policy makers, and range of services to managers).
- Are there measures of physical quantities of inputs as well as costs (although in a number of contexts costs alone may be appropriate)?

5. Are inputs and outputs measured accurately in appropriate units?
- Are all resources used relevant to the analysis accounted for?
- Are any data omitted? If so what is the justification?
- Are there any special circumstances, which make measurement difficult, e.g. joint use of staff? Were these circumstances handled appropriately?

6. Were inputs and outputs (or objectives) valued (or weighted) correctly?
- Were the sources of all values clearly identified? E.g. market prices for inputs, case mix weights?
- Was the value of outputs appropriate? Were the right weights placed upon the relationship between quantities (and qualities) of outputs?

7. Were analyses over time undertaken?
- Were values (and outputs) adjusted to present value?
- How are the specific techniques justified?, E.g. are random or fixed effects models used, how is scale accounted for, how is efficiency decomposed?

8. Do techniques add incremental value?
- For example is data envelopment analysis used? Or stochastic frontier analysis? Which cross sectional or panel data (over time) techniques are used?
- Are the techniques used justified clearly, for example what incremental value do they add beyond how efficiency is currently measured?

9. Was allowance made for uncertainty?
- Were appropriate statistical analyses undertaken?
- Were sensitivity analyses performed, which dimensions are tested?
- Were the results sensitive to the statistical/sensitivity analysis?

10. Did the presentation and discussion of study results include all issues of concern to users?
- Were the conclusions based on an overall measure, or individual comparisons of efficiency?
- Were the results compared with others who have investigated the same question?
- Did the study discuss the generalisability of the results to other settings?
- Did the study allude to other important factors in the decision or choice under consideration, e.g. ethical issues, or access issues, or equity?
- Did the study discuss issues of implementation, such as the feasibility of adopting efficiency changes, given existing operational constraints, and whether freed resources could be redeployed to other more efficient programmes?

[11]This Checklist relies heavily on Box 3.1 in Drummond *et al.* (2005) for more than just inspiration.

improving, as sample sizes increase, and statistical techniques evolve; perhaps a little more faith can be placed in well-conducted studies in this area. However, the estimated results are still sensitive to changes in the basic assumptions and specifications of the models used and the characteristics of the environment in which the units operate. Thus, as concluded previously, the results may only be valid for the units under investigation raising generalisability issues.

The review of results here should be treated with a certain amount of caution, as there are no accepted criteria for comparing efficiency studies, as there are for assessing the criteria for including clinical trials in meta-analyses. However, results here may be useful in identifying certain trends. There is still a lot of room for improvement, but as the use of Malmquist and SFA panel data techniques is showing, there is cause for optimism in terms of validity and robustness of results as longitudinal data allows more

DOI: 10.1002/hec

information to be generated than a simple cross sectional snap shot. Accuracy of the estimated performance measures is multifactorial in nature, i.e. whose objectives are we accounting for when specifying our functional forms? Model specification is still an important methodological issue, and there is still room for increased use of more advanced methods, such as panel data analysis. If we look at the general trends in health care results, public rather than private provision of health care seems more efficient for hospitals. However, these results may be confounded by different objectives of providers.

I started off by asking if there was a 'gold standard' in this area, and there may well be. A number of criteria are suggested for judging whether research published in this area is potentially useful in a policy context. It should be noted that, as with the original economic evaluation criteria on which they are modelled, these criteria should be used as a means to interpret results, not a check list for dismissing the usefulness of individual studies on a generic basis; what is of no use to one user may be very useful to another, working from a different viewpoint in a different health system. As such I have not reviewed every study in this context, I simply use the experience of reading them to suggest the means by which studies may be made more user friendly, in terms of identification of their strengths and weaknesses.

In terms of 'best practice' for undertaking efficiency studies, it may be that the use of multiple techniques might help indicate trends in inefficiency. If the multiple techniques (parametric and non-parametric, including techniques which can account for multiple objectives) point to the same inefficient organisations, and the organisations cannot sensibly explain them away (i.e. omitted variables, policy shocks), then perhaps we are picking up some form of inefficiency. Of course, it may be that in certain circumstances one method is obviously more useful, for example, when there are multiple outputs, SFA may not be appropriate due to problems with having to aggregate variables. Justification of the method used is sometimes difficult at present as there are few criteria for which are 'best', although, in practice different measurement methods often show similar results. Another danger at present is relying on exact numbers; small differences in inefficiency may not truly reflect inefficiency, and should be viewed with caution. Trends over time may be more reliable.

There are still data problems, not only on the output side, where use of case mix adjusted outputs is increasing. We should not forget input data requirements, in terms of human capital for example.[12] There is still a large disparity between the number of studies published in this area and the actual use of efficiency measures in this area, but anecdotally there does appear to be increased interest in use of efficiency and productivity measures internationally by policy makers, and this, if handled properly may be useful, although issues concerning appropriateness of use in developing countries are worth bearing in mind, given a potentially totally different set of system objectives.

As economists we need to bear in mind the basics of what we mean by efficiency, but of course, we do not own the term, and many other professions such as management scientists publish useful work in this area. However, not only must we decide how we measure economic efficiency and productivity, but why, and how important it is relative to other societal objectives in terms of the delivery of health care. Weightings are important, both within efficiency measures and between objectives within a social welfare function. How do we decide upon these weights? Choice of method impacts upon this for the measurement of efficiency, non-parametric measures use organisational specific weights, parametric measures common weights (Smith and Street, 2006). For the placing of efficiency among other objectives in a health care system, various methods have been suggested, but whose preferences are we to use, the public, those elected to represent them, health care professionals? The WHO attempted a study of this nature, with the World Health Report (2000), leading to much debate.

There are many challenges not discussed here, such as the impact on measuring performance overall, of the principal-agent relationship and the impact this might have on the relative weighting of objectives (see Smith and Street, 2006, for a discussion of this); efficiency measurement as discussed here is really

[12]A point not discussed in this paper is the value of information, and the use of Bayesian methods. Perhaps this could help with certain data problems.

DOI: 10.1002/hec

only a partial measure of overall performance. Also, at what level are performance measures most useful (Hollingsworth and Parkin, 2003; Propper and Wilson, 2006) – for example, it may be difficult to convince organizations that measures will not be used as a stick rather than a carrot. Is productivity measurement (by which I mean performance over time) a link between static efficiency measurement and actual improvement in performance?

Is there a way forward? As well as refining techniques so that we measure efficiency accurately in its own right, and recognising and measuring the relative importance of efficiency as an objective, we need to translate our work carefully to those who may actually be interested in it. As noted by Burgess (2006) and Hollingsworth and Street (2006) the actual demand for efficiency measurement is well out of line with supply. If we can surmount some of these problems, convincing ourselves and others, and bearing in mind other perhaps equally important objectives, it may be that, in the end, *patients* may actually benefit from increased efficiency in our health services.

APPENDIX

Studies not detailed in Hollingsworth (1999; 2003) are listed here

Hospital type	Country	Number	Author	Efficiency Scores
Federal/Defense/ Veterans' Administration				
	USA	131 hospitals in 1998, 121 in 2001	Harrison and Ogniewski, 2005	1998 0.86 (sd 0.11) 2001 0.86 (sd 0.11)
Not- for-profit				
	USA	480 hospitals in 1998, 471 in 2001	Harrison and Sexton, 2006	1998 0.72 (sd 0.15) 2001 0.74 (sd 0.16)
Acute/General				
	USA	89 hospitals	Chen *et al.*, 2005	VRS TE 0.81–-0.85 CRS TE 0.75–0.80 SE 0.93–0.94
	USA	170 hospitals	Ferrier *et al.*, 2006	TE 0.95 SE 0.98
	USA	348 hospitals	Mobley and Magnussen, 2002	TE 0.908
Public				
	Taiwan	1996–483 hospitals 1997–473 hospitals	Chang *et al.*, 2004	Efficiency range: 0.58–0.93. Private hospitals more efficient, may be down to case mix or quality
	Ireland	33 hospitals	Gannon, 2005	Efficiency range: 0.94–0.97
	USA	1998–280 Federal hospitals, 2001–245 Federal hospitals	Harrison *et al.*, 2004	Means: 1998 0.68 2001 0.79
	Kenya	54 hospitals	Kirigia *et al.*, 2002	TE: 0.956 SE: 0.968
	Finland	114 Public health centre hospitals	Laine *et al.*, 2005a	Mean: 0.72
	Norway and Finland	47 in Finland, 51 in Norway	Linna *et al.*, 2006	CRS: Norway 0.83 Finland 0.86 VRS: Norway 0.92 Finland 0.92
	China	6 hospitals	Liu and Mills, 2005	1978: 0.97 1997: 0.73
	Norway	51 hospitals	Martinussen and Midttun, 2004	1999: 0.827 2000: 0.835 2001: 0.841

DOI: 10.1002/hec

Hospital type	Country	Number	Author	Efficiency Scores
	Thailand	68 hospitals	Valdmanis *et al.*, 2004	Mean: 0.954 (sd 0.69)
	Namibia	26 district hospitals	Zere *et al.*, 2006	TE: range 0.627–-0.743 SE: range 0.732–-0.837
	South Africa	56 hospitals	Kirigia *et al.*, 2000	TE: 0.906
Teaching				
	USA	254 hospitals	Grosskopf *et al.*, 2004	TE: CRS 0.6, VRS 0.71 SE: 0.85
Non-specific				
	USA	Hospitals in 306 metropolitan areas	Bates *et al.*, 2006	Mean Effic. 0.888
	USA	38 (then 19) merged hospitals	Ferrier and Valdmanis, 2004	1996 OE 0.79 TE 0.88 SE 0.9 1997 OE 0.8 TE 0.88 SE 0.87
	Greece	17 hospitals	Kontodimopoulos *et al.*, 2006b	Hospital services mean: 0.749 (sd 0.112)
	USA	53 non-metropolitan hospitals	O'Neill and Dexter, 2004	Median: 0.99
	Ghana	17 district hospitals	Osei *et al.*, 2005	VRS TE: 0.61 (sd 0.12) SE: 0.81 (sd 0.25)
	Spain	29 hospitals	Prior, 2006	1990, without quality: 0.929; with quality 0.95 1993, without quality: 0.907; with quality 0.92
	Oman	20 hospitals	Ramanathan, 2005	CRS mean: 0.872 VRS mean: 0.926
	Botswana	13 hospitals	Ramanathan *et al.*, 2003	Mean: 0.99
	Italy	117 hospitals	Siciliana, 2006	CRS: range 0.65-1 VRS: Range: 0.63-1 FDH range: 0.75-1
	Germany	1,700 hospitals	Staat, 2006	Mean: 0.79
	Germany and Switzerland	105 German and 251 Swiss hospitals	Steinmann *et al.*, 2004	German range 0.79-0.828, Swiss range: 0.719–0.752
	Germany	2 ,020–2, 145 hospitals	Helmig and Lapsley, 2001	Range: 0.769–1
	France and USA		Dervaux, 2004	France inefficiency: 19.8% (SE 9.5%, TE 7.1%, congestion 3.2%); US inefficiency: 23.7% (SE 6.3%, TE 14%, congestion 3.4%). Direct comparison difficult.
Psychiatric	USA	506 hospitals	Ferrier and Valdmanis, 2002	Mean: 0.47 Private NFP: 0.65 Private FP: 0.57 Public: 0.23–-0.28
	Greece	90 facilities	Kontodimopoulos *et al.*, 2006a	Mean: 0.73 Public: 0.688 Private: 0.866

AE, Allocative efficiency; CE, Cost efficiency; FDH, Free disposable hull; FP, For Profit; NFP, Not-for-profit; OE, Overall efficiency; SE, Scale efficiency; TE, Technical efficiency; VA, Veterans administration.

Organisation type	Country	Number	Author	Efficiency Scores
Cross country	OECD	24 countries	Afonso and St Aubyn, 2005	CRS: 0.815 VRS: 0.832–0.946
	OECD	24 countries	Bhat, 2005	Mean: 0.901
	OECD	27 countries	Retzlaff-Roberts *et al.*, 2004	Mean: 0.89–0.958
	Developing countries	51 countries	Alexander *et al.*, 2003	Effic: 1.033–1.036
Primary Care				
	USA	16 primary care physicians	Collier, 2006a, b	Mean: 0.96

DOI: 10.1002/hec

Organisation type	Country	Number	Author	Efficiency Scores
	USA	115 physician practices	Andes *et al*., 2002	Mean: 0.39
	USA	156 primary care practices (& 346 specialist care)	Rosenman and Friesner, 2004	TE: 0.736 AE: 0.913 CE: 0.75 SE: 0.876
	USA	27 primary care physicians	Wagner *et al*., 2003	Range: 0.65–1
	USA	21 primary care physicians	Wagner and Shimshak, 2000	Score: 0.903
	Sierra Leone	37 public health units	Renner *et al*., 2005	TE: 0.72 SE: 0.82
	Austria	591 GPs	Staat, 2003	Mean: 0.84
	Kenya	32 Public health centers	Kirigia *et al*., 2004	Mean TE: 0.8 SE: 0.9
Nursing Homes				
	Finland	10 Long term care facilities (64 units)	Björkgren *et al*., 2004	Mean 0.86–0.87
	Netherlands	71 homes	Blank and Valdmanis, 2005	TE: 1 SE: 1 AE: 0.95 CE: 0.95
	USA	487 homes	Anderson *et al*., 2003	Mean: 0.72 FP: 0.77 NFP: 0.74
Vaccination				
	Bangladesh	117 sites/clinics	Valdmanis *et al*., 2003	CRS TE: 0.33 (sd 0.26) VRS TE: 0.50 (sd 0.29) SE 0.64 (sd 0.27)
Dental services				
	USA	279, 999 patient encounters	Coppola *et al*., 2003	Mean: 0.788 (sd 0.13)
	EU	6 countries	Parkin and Devlin, 2003	Health care median: 0.848 Oral health median: 0.483
Renal care				
	UK	70 haemodialysis units	Gerard and Roderick, 2003	Mean: 0.9
	Greece	118 haemodialysis units	Kontodimopoulos and Niakas, 2005	Overall Mean: 0.704 (sd 0.139) Public: 0.65 Private: 0.82
	USA	49 renal dialysis facilities	Ozgen, 2006	Mean range: 0.876– 0.896 (sd 0.01–0.09)

AE, Allocative efficiency; CE, Cost efficiency; FDH, Free disposable hull; FP, For Profit; NFP, Not-for-profit; OE, Overall efficiency; SE, Scale efficiency; TE, Technical efficiency; VA, Veterans administration.

Organisation type	Country	No of units	Author	Results
Cross country				
	—	143 countries	Grosskopf *et al*., 2006	Inefficiency: 11-–22%, Developed countries 0-–3%, less developed 22-–44%
Hospital	Taiwan	40 hospitals	Chen, 2006	2.6% overall fall in productivity, technology falls outweigh efficiency improvements.
	UK	53 hospitals	Ferrari, 2006	2% increase in total productivity (technology increase 3%, efficiency fall 1%).
	Greece	10 public hospital clinics	Lyroudi *et al*., 2006	Efficiency is improving, range: 1-–1.56
	Greece	30 hospitals	Maniadakis and Thanassoulis, 2004	Malmquist: 0.97 Cost Malmquist: 0.96

Organisation type	Country	Number	Author	Efficiency Scores
	Spain	20 hospitals	Sola and Prior, 2001	Malmquist: 1.3, change in quality 0.98, TE change 0.98, technology change 1.36. Technology driving global fall in index.
	Spain	68 public hospitals	Ventura *et al*., 2004	TE: 0.829 Pure TE: 0.89 SE: 0.93
Pharmacy				
	Spain	80 pharmacy laboratories	Gonzalez and Gascon, 2004	CRS: 0.68 VRS: 0.84 SE: 0.81 Technology change minimal.
Dialysis	USA	140 dialysis facilities	Ozgen and Ozcan, 2004	Mean: 0.918. Technology regress, efficiency increase, may mean increase in quality.
Primary care	Austria	591 GPs	Staat, 2003	Productivity Index 1.05, individual efficiency improves (0.98), technology falls (1.08).

AE, Allocative efficiency; CE, Cost efficiency; FDH, Free disposable hull; FP, For Profit; NFP, Not-for-profit; OE, Overall efficiency; SE, Scale efficiency; TE, Technical efficiency; VA, Veterans administration.

Organisation type	Country	No of units	Author	Results
Hospitals				
	Holland	General hospitals in 27 health care regions	Blank and Eggink, 2004	TE: 0.86 AE: 0.92 There is technical regress through time
	USA	57 rural hospitals	Butler and Li, 2005	58% inefficient Returns to scale measured
	USA	140 hospitals (1999), 139 (2000), 137 (2001)	Deily and McKay, 2006	Mean efficiency 87%. Cost inefficiency has a positive and significant effect on mortality.
	Ireland	33 hospitals	Gannon, 2005	Various models, including time varying, and compares with DEA, concluding that DEA is not controlling for certain factors.
	USA	22 VA networks, 138 medical centres	Gao *et al*., 2006	93% efficiency.
	Finland	48 acute hospitals	Linna *et al*., 2006	Average efficiency: 82%. May be economies of scope in University hospitals.
	USA	4,075 hospitals	McKay *et al*., 2002/3	Efficiency: All: 0.859-–0.852 NFP: 0.865–-0.859 FP: 0.837–-0.833 Govt.: 0.853–-0.844
	USA	616 hospitals	Rosko, 2004	1990 Efficiency: 85.65% 1999 Inefficiency: 88.22% decreases associated with HMO penetration, and time; increases associated with FP status and Medicare share.
	USA	1,368 urban general hospitals	Rosko and Proenca, 2005	Mean Efficiency: 0.852. Network/system users were more efficient.

DOI: 10.1002/hec

Organisation type	Country	Number	Author	Efficiency Scores
	UK	226 acute hospitals	Street, 2003	COLS: 0.694 SF: 0.874-0.903. COLS overstates inefficiency if hospitals are prone to random events, and choice of technique affects ranking.
	USA	131 VA hospitals	Yaisawarng and Burgess, 2006	Mean effic: 0.94.
	USA	382 USA non-teaching hospitals	Koop *et al*., 1997	Mean effic: 0.85 NFP: 0.86 FP: 0.79 Govt.: 0.87 Model specification is robust (Bayesian SFA).
Cross country	WHO	141 countries (1993–1997), 50 more in 1997	Gravelle *et al*., 2003	It is premature to reach conclusions on the production of health, given available methods.
Nursing homes				
	USA	653 homes	Anderson *et al*., 1999	Bayesian SFA, Overall effic.: 0.69 Indep: 0.784 FP: 0.901 NFP: 0.725
	Switzerland	886 homes	Crivelli *et al*., 2002	Median efficiency 87%. Ownership not significantly related to inefficiency.
	Switzerland	36 homes	Farsi and Filippini, 2004	Private NFP are more efficient than public by 3%. There are potential scale economies.
	Finland	122 institutional care wards for the elderly	Laine *et al*., 2005a	Technical efficiency: 84%.
Primary care				
	Spain	180 primary care health teams	Puig-Junoy and Ortûn, 2004	Overall effic: 0.92 Public: 0.928 Contracted out: 0.837
States				
	India	16 states public health systems	Kathuria and Sanker, 2005	Inefficiency range: 0.686–1 in fixed effects, 0.725–1 in random effects, 0.719–1 in maximum likelihood.

AE, Allocative efficiency; CE, Cost efficiency; FDH, Free disposable hull; FP, For Profit; NFP, Not-for-profit; OE, Overall efficiency; SE, Scale efficiency; TE, Technical efficiency; VA, Veterans administration.

ACKNOWLEDGEMENTS

Bruce Hollingsworth is a Victorian Health Promotion Foundation Research Fellow. Research funding from the Australian Research Council is also acknowledged. Many thanks to Katharina Hauck for comments, Jim Burgess for comments as discussant when this paper was presented at the International Health Economics Association World Congress in Copenhagen, 2007, Bill Greene for an important technical point and two anonymous referees whose suggestions helped sharpen the paper a great deal.

REFERENCES

Afonso A, St Aubyn M. 2005. Non-parametric approaches to education and health efficiency in OECD countries. *The Journal of Applied Economics* **8**(2): 227–246.

Alexander CA, Busch G, Stringer K. 2003. Implementing and interpreting a data envelopment analysis model to assess the efficiency of health systems in developing countries. *IMA Journal of Management Mathematics* **14**: 49–63.
Anderson RI, Lewis D, Webb JR. 1999. The efficiency of nursing home chains and the implications of non-profit status. *Journal of Real Estate Portfolio Management* **5**(3): 235–245.
Anderson RI, Weeks HS, Hobbs BK, Webb JR. 2003. Nursing home quality, chain affiliation, profit status and performance. *Journal of Real Estate Research* **25**(1): 43–60.
Andes S, Metzger LM, Kralewski J, Gans D. 2002. Measuring efficiency of physician practices using data envelopment analysis. *Manage Care,* 48–54.
Asian Development Bank. 2006. *Measuring Policy Effectiveness in Health and Education*. ADB: Philippines.
Bates LJ, College B, Mukherjee K, Santerre RE. 2006. Market structure and technical efficiency in the hospital services industry:a DEA approach. *Medical Care Research and Review* **63**(4): 499–524.
Bhat VN. 2005. Institutional arrangements and efficiency of health care delivery systems. *The European Journal of Health Economics* **3**: 215–222.
Bjorkgren MA, Fries BE, Hakkinen U, Brommels M. 2004. Case-mix adjustment and efficiency measurement. *Scan Journal of Public Health* **32**(6): 464–471.
Blank JL, Eggink E. 2004. The decomposition of cost efficiency: an empirical application of the shadow cost function model to Dutch general hospitals. *Health Care Management Science* **7**: 79–88.
Blank JL, Valdmanis V. 2005. A modified three-stage data envelopment analysis. *The European Journal of Health Economics* **6**(1): 65–72.
Burgess JF. 2006. Productivity analysis in health care. In *The Elgar Companion to Health Economics,* Jones A (ed). Edward Elgar: Cheltenham.
Butler TW, Li L. 2005. The utility of returns to scale in DEA programming: an analysis of Michigan rural hospitals. *European Journal of Operational Research* **161**(2): 469–477.
Chang H, Cheng M-A, Das S. 2004. Hospital ownership and operating efficiency: evidence from Taiwan. *European Journal of Operational Research* **159**: 513–527.
Chen A, Hwang YC, Shao B. 2005. Measurement and sources of overall and input inefficiencies: evidences and implications in hospital services. *European Journal of Operational Research* **161**(2): 447–468.
Chen SN. 2006. Productivity changes in Taiwanese hospitals and the national health insurance. *The Service Industry Journal* **26**(4): 459–477.
Coelli T, Rao DSP, O'Donnell CJ, Battesse G. 2005. *An Introduction to Efficiency and Productivity Measurement*. Springer: New York.
Collier DA, Collier CE, Kelly TM. 2006a. Benchmarking physician performance. *The Journal of Medical Practice Management Part 1*. 185–189.
Collier DA, Collier CE, Kelly TM. 2006b. Benchmarking physician performance, part 2. *The Journal of Medical Practice Management* 273–279. [Erratum appears in *The Journal of Medical Practice Management* 2006; **21**(6): 322].
Cooper WW, Seiford LM, Zhu J. 2004. *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers: Massachusetts, USA.
Coppola MN, Ozcan YA, Bogacki R. 2003. Evaluation of performance of dental providers on posterior restorations: does experience matter? A data envelopment analysis (DEA) approach. *Journal of Medical Systems* **27**(5): 445–456.
Crivelli L, Filippini M, Lunati D. 2002. Regulation, ownership and efficiency in the Swiss nursing home industry. *International Journal of Health Care Finance and Economics* **2**(2): 79–97.
Culyer AJ. 1992. The morality of efficiency in health care-some uncomfortable implications. *Health Economics* **1**: 7–18.
Culyer AJ. 2005. *The Dictionary of Health Economics*. Edward Elgar: Cheltenham, UK.
Culyer AJ. 2006. The bogus conflict between efficiency and vertical equity. *Health Economics* **15**(11): 1155–1158.
Deily ME, McKay NL. 2006. Cost inefficiency and mortality rates in Florida hospitals. *Health Economics* **15**: 419–431.
Dervaux B. 2004. Comparing French and US hospital technologies: a directional input distance function approach. *Applied Economics* **36**: 1065–1081.
Drummond M, Sculpher M, Torrance G, O'rien B, Stoddart G. 2005. *Methods for the Economic Evaluation of health care programmes*. Oxford University Press: Oxford, UK.
Färe R, Grosskopf S, Lindgren B, Poullier J-P. 1997. Productivity growth in health care delivery. *Medical Care* **35**(4): 354–366.
Farrell MJ. 1957. The measurement of productive efficiency. *Journal of Royal Statistical Society (Series A)* **120**(3): 253–281.

DOI: 10.1002/hec

Farsi M, Filippini M. 2004. An empirical analysis of cost efficiency in non-profit and public nursing homes. *Annals of Public and Cooperative Economy* **75**(3): 339–365.
Ferrari A. 2006. Market oriented reforms of health services: a non-parametric analysis. *The Service Industry Journal* **26**(1): 1–13.
Ferrier GD, Rosko MD, Valdmanis VG. 2006. Analysis of uncompensated hospital care using a DEA model of output congestion. *Health Care Management Science* **9**: 181–188.
Ferrier GD, Valdmanis VG. 2002. Exploring psychiatric hospital performance using data envelopment analysis and cluster analysis. *Journal of Medical Economics* **20**(3/4): 143–153.
Ferrier GD, Valdmanis VG. 2004. Do mergers improve hospital productivity? *Journal of Operation Research Society* **55**(10): 1071–1080.
Folland S, Goodman AC, Stano M. *The Economics of Health and Health Care*. Prentice-Hall: New Jersey, 2001.
Gannon B. 2005. Testing for variation in technical efficiency of hospitals in Ireland. *Economics Social Reviews* **36**(3): 273–294.
Gao J, Campbell J, Lovell CAK. 2006. Equitable resource allocation and operational efficiency evaluation. *International Journal of Health Care Finance and Economics* **7**(1/2): 143–167.
Gerard K, Roderick P. 2003. Comparison of apparent efficiency of haemodialysis satellite units in England and Wales using data envelopment analysis. *International Journal of Technology Assessment in Health Care* **19**(3): 533–539.
Gonzalez E, Gascon F. 2004. Sources of productivity growth in the Spanish pharmaceutical industry (1994–2000). *Research Policy* **33**(5): 735–745.
Gravelle H, Jacobs R, Jones AM, Street A. 2003. Comparing the efficiency of national health systems: a sensitivity analysis of the WHO approach. *Applied Health Economics and Health Policy* **2**(3): 141–147.
Grosskopf S. Efficiency and Productivity. 1993. In *The Measurement of Productive Efficiency: Techniques and Applications*, Fried HO, Lovell CAK, Schmidt S (eds.) Oxford University Press: New York.
Grosskopf S, Margaritis D, Valdmanis V. 2004. Competitive effects on teaching hospitals. *European Journal of Operational Research* **154**(2): 515–525.
Grosskopf S, Self S, Zaim O. 2006. Estimating the efficiency of the system of healthcare financing in achieving better health. *Applied Economics* **38**(13): 1477–1488.
Harrison JP, Coppola MN, Wakefield M. 2004. Efficiency of federal hospitals in the United States. *Journal of Medical Systems* **28**: 411–422.
Harrison JP, Ogniewski RJ. 2005. An efficiency analysis of Veterans Health Administration hospitals. *Military Medicine* **170**(7): 607–611.
Harrison JP, Sexton C. 2006. The improving efficiency frontier of religious not-for-profit hospitals. *Hospital Topics* **84**: 2–10.
Hauck K, Street A. 2006. Performance assessment in the context of multiple objectives: a multivariate multilevel analysis. *Journal of Health Economics* **25**(6): 1029–1048.
Helmig B, Lapsley I. 2001. On the efficiency of public, welfare and private hospitals in Germany over time: a sectoral data envelopment analysis study. *Health Services Management Research* **14**: 263–274.
Hollingsworth B. 2003. Non-parametric and parametric applications measuring efficiency in health care. *Health Care Management Science* **6**(4): 203–218.
Hollingsworth B, Dawson P, Maniadakis N. 1999. Efficiency measurement of health care: a review of non-parametric methods and applications. *Health Care Management Science* **2**(3): 161–172.
Hollingsworth B, 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* **8**: 230–236.
Hollingsworth B, Street A. 2006. The market for efficiency analysis of health care organisations. *Health Economics* **15**(10): 1055–1059.
Jacobs R, Smith PC, Street A. 2006. *Measuring Efficiency in Health Care*. Cambridge University Press: Cambridge, UK.
Kathuria V, Sankar D. 2005. Inter-state disparities in health outcomes in rural India: an analysis using a stochastic production frontier approach. *Development Policy Review* **23**(2): 145–163.
Kirigia JM, Emrouznejad A, Sambo LG. 2002. Measurement of technical efficiency of public hospitals in Kenya: using Data Envelopment Analysis. *Journal of Medical Systems* **26**: 39–45.
Kirigia JM, Emrouznejad A, Sambo LG, Munguti N, Liambila W. 2004. Using data envelopment analysis to measure the technical efficiency of public health centers in Kenya. *Journal of Medical Systems* **28**(2): 155–166.
Kirigia JM, Lambo E, Sambo LG. 2000. Are public hospitals in Kwazulu-Natal province of South Africa technically efficient. *African Journal of Health Science* **7**: 25–32.

Kontodimopoulos N, Bellali T, Labiris G, Niakas D. 2006a. Investigating sources of inefficiency in residential mental health facilities. *Journal of Medical Systems* **30**(3): 169–176.
Kontodimopoulos N, Nanos P, Niakas D. 2006b. Balancing efficiency of health services and equity of access in remote areas in Greece. *Health Policy* **76**(1): 49–57.
Kontodimopoulos N, Niakas D. 2005. Efficiency measurement of hemodialysis units in Greece with data envelopment analysis. *Health Policy* **71**(2): 195–204.
Koop G, Osiewalski J, Steel MFJ. 1997. Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics* **76**: 77–105.
Kumbhakar S, Lovell CAK. 2000. *Stochastic Frontier Analysis*. Cambridge University Press: Cambridge.
Laine J, Finne-Soveri UH, Bjorkgren M, Linna M, Noro A, Hakkinen U. 2005a. The association between quality of care and technical efficiency in long-term care. *International Journal for Quality in Health Care* **17**(3): 259–267.
Laine J, Linna M, Hakkinen U, Noro A. 2005b. Measuring the productive efficiency and clinical quality of institutional long-term care for the elderly. *Health Economics* **14**(3): 245–256.
Linna M, Hakkinen U, Magnussen J. 2006. Comparing hospital cost efficiency between Norway and Finland. *Health Policy* **77**(3): 268–278.
Liu X, 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 (12 pages).
Lyroudi K, Glaveli N, Koulakiotis A, Angelidis D. 2006. The productive performance of public hospital clinics in Greece: a case study. *Health Services Management and Research* **19**: 67–72.
Maniadakis N, Thanassoulis E. 2004. A cost Malmquist productivity index. *European Journal of Operational Research* **154**: 396–409.
Marchand M, Pestieau P, Tulkens H. 1984. The performance of public enterprises: normative, positive and empirical issues. In *The Performance of Public Enterprises: Concepts and Measurement,* Marchand M, Pestieau P, Tulkens H (eds). Elsevier: The Netherlands.
Martinussen PE, Midttun L. 2004. Day surgery and hospital efficiency: empirical analysis of Norwegian hospitals, 1999–2001. *Health Policy* **68**(2): 183–196.
McKay NL, Deily ME, Dorner FH. 2003. Ownership and changes in hospital inefficiency. *Inquiry* **39**(4): 388–399.
Mobley LR, Magnussen J. 2002. The impact of managed care penetration and hospital quality on efficiency in hospital staffing. *Journal of Health Care Finance* **28**: 24–42.
Newhouse JP. 1994. Frontier estimation: how useful a tool for health economics?. *Journal of Health Economics* **13**(3): 317–322.
Nishimuzu M, Page JM. 1982. Total factor productivity growth, technological progress and technical efficiency change: dimensions of productivity change in Yugoslavia 1965–1978. *Economics Journal* **92**: 920–936.
O'Neill L, Dexter F. 2004. Market capture of inpatient perioperative services using DEA. *Health Care Management Science* **7**: 263–273.
Osei D, d'Almeida S, George MO, Kirigia JM, Mensah AO, Kainyu LH. 2005. Technical efficiency of public district hospitals and health centres in Ghana: a pilot study. *Cost Effectiveness and Resource Allocation* **3**: 9 (13 pages).
Ozgen H. 2006. Does chain affiliation make a difference in efficiency of dialysis providers in the USA. *Social Science and Medicine* **62**(9): 2112–2124.
Ozgen H, Ozcan YA. 2004. Longitudinal analysis of efficiency in multiple output dialysis markets. *Health Care Management Science* **7**: 253–261.
Parkin D, Devlin N. 2003. Measuring efficiency in dental care. In *Advance in Health Economics*, Scott A, Maynard A, Elliot R (eds). Wiley: West Sussex.
Pestieau P, Tulkens H. 1993. Assessing and explaining the performance of public enterprises: some recent evidence from the productive efficiency viewpoint. In *Public Goods, Environmental Externalities and Fiscal Competition*, Chander P, Drèze J, Lovell CK, Mintz J (eds). Springer: USA.
Prior D. 2006. Efficiency and total quality management in health care organisations: A dynamic frontier approach. *Annals of Operational Research* **145**: 281–299.
Propper C, Wilson D. 2006. The use of performance measures in health care systems. In *The Elgar Companion to Health Economics*, Jones A (ed). Edward Elgar: Cheltenham.
Puig-Junoy J, Ortun V. 2004. Cost efficiency in primary care contracting: a stochastic frontier cost function approach. *Health Economics* **13**(12): 1149–1165.
Ramanathan R. 2005. Operations assessment of hospitals in the Sultanate of Oman. *International Journal of Operations and Production Management* **25**(1): 36–54.
Ramanathan TV, Chandra KS, Thupeng WM. 2003. A comparison of the technical efficiencies of health districts and hospitals in Botswana. *Development Southern Africa* **20**(2): 307–320.

DOI: 10.1002/hec

Renner A, Kirigia JM, Zere EA, Barry SP, Kirigia DG, Kamara C *et al*. 2005. Technical efficiency of peripheral health units in Pujehun district of Sierra Leone: a DEA application. *BMC Health Services Research* **5**: 5 (11 pages).
Retzlaff-Roberts D, Chang CF, Rubin RM. 2004. Technical efficiency in the use of health care resources: a comparison of OECD countries. *Health Policy* **69**(1): 55–72.
Rice T. 2003. *The Economics of Health Reconsidered*. Health Administration Press: Chicago.
Rosenman R, Friesner D. 2004. Scope and scale inefficiencies in physician practices. *Health Economics* **13**(11): 1091–1116.
Rosko MD. 2004. Performance of US teaching hospitals: a panel analysis of cost inefficiency. *Health Care Management Review* **1**: 7–16.
Rosko MD, Proenca J. 2005. Impact of network and system use on hospital X-inefficiency. *Health Care Management Review* **30**(1): 69–79.
Siciliana L. 2006. Estimating technical efficiency in the hospital sector with panel data. A comparison of parametric and non-parametric techniques. *Applied Health Economics and Health Policy* **5**(2): 99–116.
Simar L, Wilson P. 2007. Estimation and inference in two stage, semi parametric models of production processes. *Journal of Econometrics* **136**: 31–64.
Smith PC. 2002. Measuring health system performance. *The European Journal of Health Economics* **3**: 145–148.
Smith PC, Street A. 2006. Concepts and challenges in measuring the performance of health care systems. In *The Elgar Companion to Health Economics*, Jones A (ed). Edward Elgar: Cheltenham.
Sola M, Prior D. 2001. Measuring productivity and quality changes using data envelopment analysis: an application to Catalan hospitals. *Financial Account Management* **17**(3): 219–245.
Spinks J, Hollingsworth B. 2005. *Health production and the socioeconomic determinants of health in OECD countries: the use of efficiency models*. Working Paper No. 151, Centre for Health Economics, Faculty of Business and Economics, Monash University: Australia.
Staat M. 2003. The efficiency of treatment strategies of general practitioners. *The European Journal of Health Economics* **4**: 232–238.
Staat M. 2006. Efficiency of hospitals in Germany: a DEA-bootstrap approach. *Applied Economics* **38**: 2255–2263.
Steinmann L, Dittrich G, Karmann A, Zweifel P. 2004. Measuring and comparing the (in)efficiency of German and Swiss hospitals. *The European Journal of Health Economics* **3**: 216–226.
Street A. 2003. How much confidence should we place in efficiency estimates? *Health Economics* **12**(11): 895–907.
Valdmanis V, Kumanarayake L, Lertiendumrong J. 2004. Capacity in Thai public hospitals and the production of care for poor and nonpoor patients. *Health Services Research* **39**(6): 2117–2134.
Valdmanis V, Walker D, Fox-Rushby J. 2003. Are vaccination sites in Bangladesh scale efficient? *International Journal of Technology Assessment in Health Care* **19**(4): 692–697.
Ventura J, González E, Cárcaba A. 2004. Efficiency and program-contract bargaining in Spanish public hospitals. *Annals of Public and Cooperative Economy* **75**(4): 549–573.
Wagner JM, Shimshak DG. 2000. Physician profiling using data envelopment analysis: a case study. *International Journal of Health Technology and Management* **2**(1/2/3/4): 358–374.
Wagner JM, Shimshak DG, Novak M. 2003. Advances in physician profiling: the use of DEA. *Socio-Economic Planning Sciences* **37**: 141–163.
World Health Report Health Systems. 2000. *Improving Performance World Health Organisation*. WHO: Geneva.
Yaisawarng S, Burgess JF. 2006. Performance-based budgeting in the public sector: an illustration from the VA health care system. *Health Economics* **15**(3): 295–310.
Zere E, Mbeeli T, Shangula K, Mandlhate C, Mutirua K, Tjivambi B, Kapenambili W. 2006. Technical efficiency of district hospitals: evidence from Namibia using data envelopment analysis. *Cost Effectiveness and Resource Allocation* **4**: 5 (9 pages).