ELSEVIER

Health Policy 56 (2001) 21–32

HEALTH policy

www.elsevier.com/locate/healthpol

Assessing efficiency in the UK breast screening programme: does size of screening unit make a difference?

Katharine Johnston [a,*], Karen Gerard [b]

[a] *Health Economics Research Centre, University of Oxford, Institute of Health Sciences, Headington, Oxford OX3 7LF, UK*

[b] *Health Care Research Unit, Community Clinical Sciences Research Division, Level B (805) South Academic Block, Southampton General Hospital, Tremona Road, Southampton SO16 6YD, UK*

Received 30 May 2000; accepted 25 October 2000

Abstract

The UK breast screening programme (UK BSP) is organised into a large number of individual screening units. Decision makers need to ensure these units are producing efficiently, particularly as the programme is anticipated to expand. Data envelopment analysis (DEA) was applied to investigate: the relative efficiency of screening units; the impact of screening unit size on efficiency; and how individual units could improve. Sixty-four screening units were categorised into 33 large and 31 small. Data were collected using a national survey and routinely collected data. The overall median efficiency score was 91%, 39 units were inefficient. Variation in efficiency scores was wide. Large units had a median efficiency score of 100% and 12 units were inefficient. Smaller units had a median efficiency score of 95% and 19 were inefficient. This difference was not statistically significant (Mann–Whitney, $P = 0.076$). Forty-two percent of large units and 21% of small units were operating at constant returns to scale (mean difference 0.20, 95% CI: 0.15–0.43). Although there is no systematic difference in efficiency by size of screening unit there are inefficiencies in both large and small units and there is scope for many individual units to improve their use of current resources. It will be necessary for decision-makers to examine the practices of individual screening units before considering options for how best to improve their resource

* Corresponding author. Tel +44-1865-226692; Fax +44-1865-226842.

E-mail address: kathy.johnston@ihs.ox.ac.uk (K. Johnston).

PII: S0168-8510(00)00137-8

use. DEA can help to identify feasible options.

Keywords: breast cancer screening; efficiency; Data Envelopment Analysis; production function

1. Introduction

The UK governments invest heavily in breast screening, spending around 63.6 million Euro's each year on the UK breast screening programme (UK BSP) to purchase a sizeable and specialist set of health care inputs. The programme is organised into a large number of individual screening units, each having autonomy over how it produces the outputs of breast screening (invitations, screens and cancers detected).

Now in its 11th year of operation, the UK BSP is currently facing a number of challenges. There is a shortage of professional staff, particularly radiologists. Anticipated changes in screening policies, such as the extension of the upper age range for screening, expansions to women at familial risk of breast cancer, as well as increasing demographic pressures, have created additional demands. These challenges mean that decision makers in the UK BSP need to have information on the performance of individual screening units and, in particular, whether resources are currently being used efficiently.

To address the relative efficiency of individual screening units, a suitable performance indicator is required. To date, the performance indicators in the UK BSP have not focused on efficiency *per se* but have measured uptake, recall and cancer detection rates. These routinely collected performance indicators have suggested that the programme continues to achieve high levels of acceptance (in terms of uptake rates) and to offer a quality service (in terms of recall and cancer detection rates), although there is variation across the UK [1–3]. Such performance indicators, however, have limitations.

First, they do not measure all outputs simultaneously and instead measure uptake, recall and cancer detection rates independently. Second, such performance indicators ignore the levels and types of inputs and, in turn, the relationship between inputs and outputs. There may be considerable differences in the way in which different units combine their inputs to produce their outputs. Furthermore, such differences may vary according to the scale of production.

The ideal performance measure for breast screening units would be one that addressed the relationship between multiple inputs and outcomes. In addition, it is most appropriate for breast screening units to have their efficiency judged relative to those with which they are broadly similar in terms of input/output orientation. If this is not the case then *inter alia* managers of one unit may argue that it is 'unfair' to be compared with another that which they consider to be different in some important dimension. One particular difficulty with measuring the performance of breast screening units is the potential for the interpretation of cancer data to be confounded. This arises because cancer detection rates may vary geographi-

cally. They ought as a result to be adjusted for background incidence of breast cancer in the geographical area that the unit serves. (A correction factor has recently been produced by Blanks and Moss [4] and can be used to overcome this issue.) Cancer detection rates may also be confounded by the problem of a small numbers. If a unit screens a small number of women and picks up, for example, 10 cancers, the cancer detection rate would be very high. Similarly, if a unit screens a small number of women but does not detect any cancers this would lead to a zero cancer detection rate. Consequently, the cancer detection rate of very small units can be subject to marked spurious variations making it necessary to adjust for this in subsequent analysis.

2. Size of screening unit

The size of a screening unit can be measured by its eligible population, which is the number of women of screening age (aged 50–64) who reside in the catchment area. Since the interval of screening is 3-yearly, the number of women invited for screening annually is approximately one-third of the eligible population. Although a screening unit may operate both static and mobile screening, the size of screening unit is defined by the total eligible population. When the UK BSP was set up, the Forrest Committee [5], whose role it was to make initial recommendations for screening policies and practices, provided some guidance on the size of the eligible population each screening unit should serve. The size of a screening unit was estimated on the basis of its' serving a population of 471 000, an eligible screening population of 41 150 women over a 3-year cycle and implying a total of approximately 12 000 screens per year [5]. The Committee also recognised, however, that the exact specification of a unit's requirements would depend on the uptake, recall and cancer detection rates. A steady state pattern of screening has now been established throughout the UK [6]. Fig. 1 shows there is a wide range in the size of screening units with the smallest serving an eligible screening population of less than 20 000 women and the largest over 140 000.

Although the Forrest Committee suggested an eligible screening population of 41 150, it is not known what the efficient size is. Efficiency may be affected by size because larger units are more able to operate under conditions of economies of scale, that is, spread a larger output over the fixed inputs. This type of issue has been explored in terms, for example, of the size of acute hospitals and whether to concentrate acute services into fewer but larger hospitals [7,8]. Current knowledge, however, suggests that the relationship between hospital volume and health outcomes cannot be generalised and that there is only limited evidence to support the claim that larger hospitals benefit from economies of scale [9]. In fact, some economies can occur in small hospitals (less than 200 beds) [9]. Thus, it is not necessarily the case that larger screening units will be more efficient as a result of economies of scale. There may be other reasons why large and small screening units differ in terms of efficiency. For example, larger units

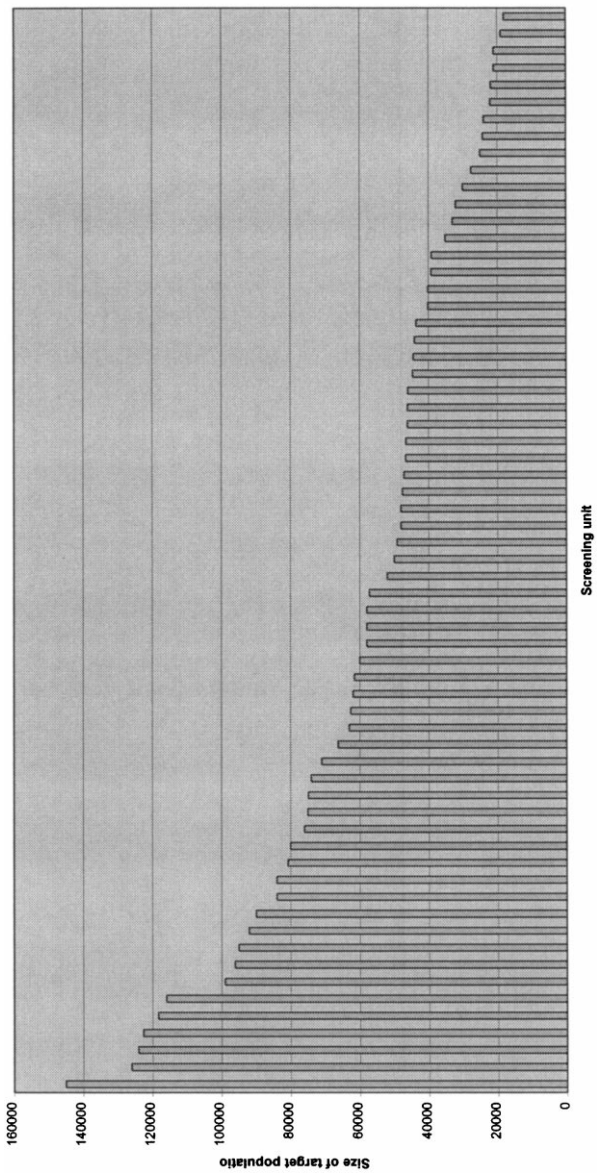

Fig. 1. Distribution of the size of UK BSP Units*, 1996. *$N = 64$ screening units meeting inclusion criteria for the study.

may be less efficient than smaller units because of a higher level of administration staff, but they may be more efficient than smaller units if they can deploy radiologists more efficiently. Smaller units may be less efficient than larger units because they have lower levels of specialist staff, such as an assessment team. Another issue with smaller units is that they may be preferred by women if they are more accessible in terms of distance travelled. All these arguments suggest that determining whether efficiency differs by size of breast screening unit is an important policy issue. There is currently no evidence of efficiency by size of screening unit in the UK or other countries with national screening programmes.

3. Assessing efficiency using data envelopment analysis

One approach to measuring efficiency is data envelopment analysis (DEA). This can be used to assess relative efficiency of productive units. Efficiency is concerned with the relationship between the inputs and outputs of a productive unit and occurs where a productive unit is organised to minimise its use of inputs from a set of outputs. In this case efficiency would be concerned with how breast screening units are organised, say, to minimise their use of inputs (deployment of administration, medical and radiographic staff with mammography machines) to produce invitations, screens and cancers detected.

DEA is a non-parametric, mathematical programming technique, using data from a sample of units to generate a single measure of efficiency for each unit and an efficiency score across units (ranging from 0%, completely inefficient, to 100% completely efficient). The efficiency score of a productive unit is defined by its position relative to the frontier of best performance, which is established mathematically by the ratio of the weighted sum of outputs to the weighted sum of inputs [10]. The efficiency score of a productive unit that is not on the frontier is estimated by comparing it with those efficient productive units on the frontier with the most similar input/output orientation (peers). Efficiency scores can be used to identify and describe efficient and inefficient units, units of best practice (peers) for a particular inefficient unit, and scale inefficiencies. The analysis may also be used to quantify potential targets for improving inefficient units [11]. DEA has been portrayed as a useful economic tool to measure the relative efficiency of health service productive units performing similar tasks [12]. Its usefulness is attributed to its appropriateness in the multiple input/output environment of much of health care provision and the simplicity of the assumptions that underlie the method.

Using data envelopment analysis, the study sought answers to three related questions:

- how well does the UK BSP overall, and individual units within it, perform?;
- does efficiency differ by size of screening unit?; and
- can the performance of individual units be improved upon within existing resource constraints?

4. Methods

4.1. Setting

In 1996 the UK BSP was delivered through 97 local breast screening units. Although mandatory activity-based information is collected from all units and routinely published (e.g. [3]), this data base lacks certain key information about the inputs of screening units to enable investigation of local production processes. A range of potentially relevant input data was obtained from a national survey of local screening policies and practices [6]. The survey produced a good response rate (from 82 (87%) of the UK screening units) and was consequently used to identify the set of screening units to be included in this study. As noted in the introduction, the potential for either very high or very low cancer detection rates across screening units causes problems in interpreting the cancer detection rates. This meant that very small units, i.e. those screening fewer than 1000 women annually, have to be omitted from the analysis. Fourteen were excluded on this basis. A further four had missing data and were excluded. The number of screening units included was thereby reduced to 64.

As the Forrest Committee had suggested a screening size of 41 150 eligible population, this was used as a guide to stratify screening units into small and large. A breast screening unit was classed as a small if it served less than 50 000 eligible women and large if over this level. This led to 33 small units and 31 large.

4.2. Specification of DEA model

To specify a suitable DEA model the decisions relating to the breast screening production process had to be taken. These relate to: (i) the selection of the most appropriate input and output variables; (ii) whether to apply weight restrictions on certain variables; and (iii) the choice of optimisation method. By using different parameters to identify a series of models under (i)–(iii) specification tests were conducted to identify the best model. These tests are reported in detail elsewhere [13]. The DEA software package, Frontier Analyst [14], was used to generate efficiency scores, score rankings and to identify peers and targets. The statistical package SPSS v10.0 allowed descriptive statistics and statistical tests to be presented. Parametric tests were used to compare descriptive variables by size of screening unit and non-parametric tests to compare skewed efficiency scores.

4.3. Selection of outputs and inputs

Efficiency scores are highly sensitive to the inputs and outputs included. Therefore careful selection of the most appropriate inputs and outputs to characterise the breast screening production process is an important step.

The ideal output would be the sum of the number of true negatives and true positives. This would require data on the sensitivity and specificity of screening for each radiologist within each screening centre. This information is not recorded in

the KC-62 returns and thus alternative output specifications had to be considered. There were four potential outputs of breast screening: number of invitations, screens, assessments and cancers detected. Screening units should aim to maximise the number of invitations, screens and the number of cancers detected. They should not, however, attempt to maximise the number of assessments since a high number of assessments may reflect poor film reading or inability to detect cancers. Assessments can therefore be viewed as inputs to the production of cancers detected rather than outputs reducing the outputs included to three: invitations, screens and cancers detected. Cancers detected were based on cancer detection rates for women aged 50–54 (the prevalent round) and were corrected for background incidence using the Blanks and Moss correction factor [4]. The corrected cancer detection rates were then applied to all women screened and multiplied by 1000 to give the expected number of cancers detected.

The survey identified a total of 22 potential inputs. A degree of freedom problem can arise if the sum of the number of inputs or outputs is too large, relative to the number of units in the sample [15]. As a result, the model may be unable to discriminate in terms of the efficiency of units. Consequently, a means of reducing the number to reflect core inputs was identified based on: (i) excluding one input, if high correlations existed between particular inputs [16]; and (ii) combining inputs if inputs act as complements (for example, if different grades of staff perform the same task) [17].

High correlations were found between dedicated mammography machines and the number of mobile vans. Consequently only dedicated mammography machines were included. The number of whole time equivalent (WTE) medical and nursing staff engaged in assessment and the number of assessments were not highly correlated and therefore both were included. Complementary staff were combined over staff grades to reduce the number of staff inputs: WTE radiologists, radiographers, administration staff, and medical and nursing staff engaged in assessment (assessment team). The total number of inputs included was therefore reduced to six: number of WTE radiologists, radiographers, administration staff, medical and nursing staff engaged in assessment work, number of dedicated mammography machines and assessments performed.

4.4. Weight restrictions

Weight restrictions in DEA are designed to eliminate situations where exceptional efficiency scores can be obtained from performance on a single input or output. Such restrictions are added by the analyst to ensure that certain inputs or outputs are taken into consideration at least to some extent [18–20]. Minimum weights of 20% were placed on two key variables: number of cancers detected and number of WTE radiologists. The former was chosen since cancer detection is the major activity of a screening unit, so for a unit to be judged without taking this into account would be misleading. The latter was chosen as, given the current shortage of radiologists, this was one input which units should be judged on with respect to efficiency.

4.5. Optimisation method

As the interest was in how inputs could be used better, an input minimisation model was selected. This allows the analyst to determine the extent to which a unit can reduce inputs while maintaining the current level of outputs. The constant return to scale (CRS) model assumes that one unit of input results in one unit of output. The variable returns to scale (VRS) model assumes that one unit of input can result in less than one unit of output (diseconomies of scale), or more than one unit of output (economies of scale). There is no theoretically correct optimisation method [21] but the advantage of the VRS model is that scale efficiencies can be identified. It may however have less discriminatory power than the CRS model as it identifies more efficient firms than CRS. A VRS specification was used as the main model but a CRS model was adopted to identify the number of units operating at CRS within the VRS model [22].

5. Results

The efficiency scores obtained relate to 64 breast screening units of the UK BSP. This sample represents 66% of all units within the programme or 67% of units located in England, 71% in Scotland, all Welsh units and no Northern Ireland units. Thirty-one screening units (48%) were classified as large units, mean size of 83 301 eligible women. Thirty-three units (55%) were classified as small units, corresponding mean size of 36 618 eligible women (see Table 1). The smallest unit

Table 1
Size of target population, observed inputs and outputs by size of screening unit[a]

	Large units ($n = 31$)	Small units ($n = 33$)	t-test mean difference	95% confidence interval
Target population				
Mean	83 301	36 618		
Range	52 000–145,000	18 000–49 900		
Inputs	Mean (SD)	Mean (SD)		
WTE radiologist	1.2 (0.4)	0.7 (0.7)	0.50	0.22–0.78
WTE radiographers	7.9 (2.1)	3.4 (1.0)	4.46	3.63–5.29
WTE administration	6.5 (2.1)	3.3 (1.2)	3.33	2.35–4.09
WTE assessment team	1.8 (0.9)	1.1 (1.1)	0.62	0.11–1.13
No. assessments	1109 (578)	501 (203)	607	385–830
No. mammography machines	4.2 (1.4)	2.3 (0.8)	1.96	1.38–2.55
Outputs				
No. invitations	26 580 (8014)	11 931 (3408)	14 649	11 605–17 693
No. screens	19 813 (5855)	9572 (2695)	10 240	7916–12 565
No. cancers detected	91.4 (51)	41.8 (23)	49.5	29.3–69.9

[a] *Note*: WTE = whole time equivalent numbers.

Table 2
Efficiency scores generated under variable returns to scale model[a]

Grouping	Mean (SD)	Median (IQR)	Minima	Ranking	Units operating at CRS
All units ($n = 64$)	82.1 (20)	91.2 (67.9–100)	28.0	100 25	
				80–99 12	
				60–79 17	
				40–59 7	
				<40 3	
Large unit ($n = 31$)	92.1 (14)	100 (86.9–100)	51.0	100 19	13 (42%)
				80–99 8	
				60–79 2	
				40–59 2	
Small unit ($n = 33$)	84.5 (18)	95.6 (67.6–100)	48.0	100 15	7 (21%)
				80–99 5	
				60–79 9	
				40–59 4	

[a] Mann–Whitney U value, 389.0; P-value 0.076.

provided a service to a target population of 18 000 women, the largest unit was to 145 000.

Additional information presented in Table 1, adjusted for group size, is for mean observed values for the six inputs and the three outputs selected to describe and analyse breast screening production. Aside from a mean difference in the use of assessment teams that was found to be statistically insignificant between groups, remaining differences in mean values were statistically significant. For example, on the output side, a mean of 26 580 invitations were sent from large units. This resulted in a mean response of 19 813 screens being performed (average uptake rate 74%) and 91.4 breast cancers detected per 10 000 screens taken. Corresponding activity for smaller units was, on average, 16 903 invitations sent to eligible women, an average of 9572 women being screened per year (average uptake rate 56%) and, on average, 41.8 cancers detected per 10 000 women screened. Differences are also found in the way inputs are combined between small and large units.

Table 2 reports the relative efficiency under a VRS model of the breast screening programme overall and grouped by size of unit. These data allow analysis of performance of the average screening unit and individual units within the sample. A mean efficiency score of 82.1 (or median score of 91.2) was estimated overall. Thus, on average, if production practices in all screening units had followed the best practice of the 25 most efficient units identified, then mean outputs would have been produced using only 82.1% of mean inputs. Instead, the data identify 39 units under-performing to some extent when compared to peers. Twelve units were credited with an efficiency score ranging between 80 and 99; 17 units scored between 60 and 79; seven units between 40 and 59 and three less than 40% the efficiency of best practice counterparts. The least efficient unit had a score of 28.0.

The efficiency scores by group size altered the measures of central tendency and spread but these differences were of borderline statistical significance. The mean efficiency score for large units was 92.1% implying that 7.9% of inputs could have been saved in the process of producing observed outputs through best practice. Similarly, a mean efficiency for smaller units of 84.5% implies that 15.5% of inputs could have been saved. The spread of scores was wider for the small group but the Mann–Whitney U-test to compare score ranks by group size suggested no evidence to support differences in efficiency by size of screening unit ($P = 0.076$). The proportion of units operating at constant returns to scale differed by size of screening unit (42% of the large units and 21% of the small units), but this difference was not statistically significant.

Table 3 details the performance of one breast screening unit to demonstrate how managers of local services and other decision makers can use the results of DEA as an aid to finding local solutions to better performance. It shows the profile of actual inputs used by Unit A (a unit with an efficiency score of 61.6%). It compares the actual level to a target level estimated from the performance of efficient peers. For example, the following resource use could be improved in order to improve efficiency: reduce WTE radiographers from 6.7 WTE to 4.2 WTE and reduce WTE administration staff from 8.5 WTE to 4.9 WTE. The targets can be used as a starting point for local managers to identify whether improvements in input use are realisable and to identify feasible options and solutions.

6. Discussion

Information on uptake, recall and cancer-detection rates provides limited information on the performance of breast screening units as it ignores the use of inputs and the relationship between inputs and outputs. To overcome this, a measure of efficiency, that is a measure of how to minimise the use of inputs whilst maintaining output, is required. Data envelopment analysis (DEA) has been used to assess such efficiency. It is a useful technique since, as well as incorporating multiple inputs and outputs, it compares the efficiency of units compared to their peers and can also identify where efficiency gains can be made.

Table 3
Potential efficiency gain for Unit A (efficiency score of 61.6%)

Inputs	Actual level	Target level
WTE Radiologist	0.9	0.5
WTE Radiographers	6.7	4.2
WTE Administration	8.5	4.9
WTE Assessment team	2.7	0.9
Mammography machines	4	3
Assessments	1022	630

The results found no evidence to support systematic differences in efficiency by size of screening unit. This implies that there are no greater technological advantages or divisions of labour from the larger scale production process of a large screening unit. Although smaller size units may be considered preferable if they are more accessible to users, in the context of breast screening this is not the case, however, since both large and small units can use mobile units and this makes screening unit services more accessible. Although efficiency of screening units does not differ by size of screening unit, inefficiencies are present in both small and large screening units. This suggests there is room for improvement in the use of resources at inefficient screening units. These improvements can be informed by using the targets identified from the DEA. These targets can be used as a starting point for local managers to identify whether improvements in resources use are realisable as well as feasible. For example, a regional solution could be found with local screening units sharing radiologists. Currently there are no financial incentives in the provision of breast screening rewarding efficient units. One way of ensuring units meet targets would be to introduce financial incentives.

One limitation of DEA is that no account is taken of any random error in the efficiency scores and any deviations from the efficiency frontier are assumed to be due to inefficiency. This simplification arises to overcome problems of specifying functional forms. DEA is increasingly being used in a health care context, for example, to address the efficiency of acute hospitals. In that context, criticisms of the method often stem from the difficulty of taking case-mix into account. In the context of breast screening, however, case-mix is not an issue since the outputs of breast screening (invitations, screening and cancers detected) are homogeneous. Although the output measures of breast screening used in this study do not reflect final outcomes or quality of output, this limitation is common to all DEA in health care. The focus on intermediate outputs, however, does still provide useful information on efficiency. A limitation of the study is that screening units that screen fewer than 1000 women per year could not be included as this would have led to misleading cancer detection rates and have introduced bias. One way of overcoming this would have been to exclude cancers detected as an output, but this would then ignore one of the key outputs, and indeed the main objective of breast screening. In the future, an alternative might be to aggregate data over a number of years so that the number of screens goes over the 1000 threshold. Future studies could explore any changes in efficiency over time by collecting data on inputs and outputs over a number of years.

This study has shown that there is scope for certain individual screening units to improve their resource use substantially. Since both large and small units operate at variable returns to scale, the source of inefficiency may not relate to scale. It will be necessary for decision makers to examine the practices of individual screening units before determining feasible options for improving resource use. DEA can help to identify these feasible options and analysis of this kind should be encouraged and could, for example, be used to examine the efficiency of other national screening programmes. The implication of inefficient screening units is that they are using too many resources and consequently cannot be minimising costs.

Acknowledgements

This paper results from a research project funded by the Medical Research Council and Department of Health. We thank Jackie Brown and Martin Buxton for early advice on study design.

References

[1] NHS Breast Screening Programme. NHS Breast Screening Programme-Review 1995 (J. Patnick, ed.). Sheffield: NHS BSP, 1996.
[2] NHS Breast Screening Programme. NHS Breast Screening Programme-Review 1998 (J. Patnick, ed.). Sheffield: NHS BSP, 1996.
[3] NHS Breast Screening Programme (2000). NHS Breast Screening Programme-Review 1999 – Meeting new challenges (J. Patnick, ed.). Sheffield: NHS BSP, 1996.
[4] Blanks RG, Moss SM. Monitoring the performance of breast screening programmes: allowing for geographical variation in breast cancer incidence. Journal of Medical Screening 1996;3:82–4.
[5] Forrest APM. Breast cancer screening: Report to the health ministers of England, Wales and Scotland. London: HMSO, 1986.
[6] Gerard K, Johnston K, Brown J, Morton A. UK breast screening programme: how does it reflect the Forrest recommendations? Journal of Medical Screening 1997;4:10–5.
[7] NHS CRD. University of York., and Nuffield Institute for Health, University of Leeds., Hospital volume and health outcomes, costs and patient access. Effective Health Care 1996;2(8):1–16.
[8] Ferguson B, Sheldon T, Posnett J. Concentration and choice in healthcare. London: John Harper, FT Healthcare, 1997.
[9] Posnett J. Is big better? Concentration in the provision of secondary care. British Medical Journal 1999;319:1063–5.
[10] Charnes A, Cooper WW, Rhodes E. Measuring efficiency of decision making units. European Journal of Operations Research 1978;2(6):429–44.
[11] Hollingsworth B, Parkin D. The efficiency of Scottish acute hospitals: an application of data envelopment analysis. Journal of Mathematics Applied in Medicine and Biology 1995;12:161–73.
[12] Hollingsworth B, Dawson PJ, Maniadakis N. Efficiency measurement of health care: a review of non-parametric methods and applications. Health Care Management Science 1999;2:161–72.
[13] Johnston K. The cost and production of breast screening in the United Kingdom. Brunel University: PhD Thesis, 2000.
[14] Banxia Software. Frontier Analyst-Performance Measurement. Glasgow: Banxia Software, 1998.
[15] Nunamaker TR. Measuring routine nursing service efficiency: a comparison of cost per patient day and data envelopment analysis models. Health Services Review 1983;18:183–208.
[16] Norman N, Stoker B. Data Envelopment Analysis: the Assessment of Performance. Chichester: John Wiley and Sons, 1991.
[17] Valdmanis W. Sensitivity analysis for DEA models: an empirical example using public vs. NFP hospitals. Journal of Public Economics 1992;48:185–205.
[18] Allen R, Athanassopoulos A, Dyson AG, Thanassoulis E. Weights restrictions and value judgments in Data Envelopment Analysis: evolution, development and future directions. Annals of Operations Research 1997;73:13–34.
[19] Thanassoulis E. Simulating weights restrictions in Data Envelopment Analysis by means of unobserved DMUs. Management Science 1998;44(4):586–96.
[20] Wong YH, Beasley JE. Restricted weight flexibility in Data Envelopment Analysis. Journal of the Operation Research Society 1990;41:829–35.
[21] Parkin D, Hollingsworth B. Measuring production efficiency of acute hospitals in Scotland, 1991–94: validity issues in data envelopment analysis. Applied Economics 1997;29:1425–33.
[22] Cooper WW, Seiford LM, Tone K. Data envelopment analysis: A comprehensive text with models, applications, references and DEA-Solver software. Boston: Kluwer Academic Publishers, 2000.