Chapter 1
Evaluation of Performance in Health Care

1.1 Introduction

The health care industry faces new challenges every day, and comprises one-seventh of the GNP in the United States. There are new regulations, new technologies, and new organizations being created continuously as a result of public policy. Managers of health care need to respond to these challenges with sound performance evaluation and decision making. This book will offer state of the art performance evaluation methods as well as relevant and current examples to aid practicing managers and graduate students studying in this field.

Management in all industries is moving toward more objective performance evaluation and decision making. The health care industry, however, has lagged behind many other industries in this respect. When the prospective payment system first began in 1983, the health care industry had to scramble to meet the needs of their clients due to significant decreases in reimbursements for Medicare patients. The reaction to this was first to cut costs or avoid cases that would likely lose money, but later most administrators realized that the only way to keep their institutions financially viable was to improve their performance. Hence, benchmarking became the new buzz word. Unfortunately, the benchmarks established using old analytical schemes based on various multiple ratios created more dilemmas than solutions. Performance evaluation based on optimization techniques and their normative structure not only creates benchmarks, but also provides information for lacking organizations and illustrates how to improve performance. This is what is needed in the health care industry today.

This book places emphasis on the application of contemporary performance and efficiency evaluation methods, using data envelopment analysis (DEA), to create optimization-based benchmarks including, but not limited to hospitals, physician group practices, health maintenance organizations, nursing homes, and other health care delivery organizations. Hence, this book will not only be useful for graduate students to learn DEA applications in health care, but will also be an excellent reference and "how to book" for practicing administrators.

1.2 Performance Measurement

During the past few decades, parametric and non-parametric methods have been employed increasingly to measure and analyze the performance of health care services. This section reviews the issues in performance measurement for health services.

Health care managers must adapt new methods to use the resources at their disposal in order to achieve high performance, namely effective and high quality medical outcomes. Performance, as in other service industries, can be defined as an appropriate combination of efficiency and effectiveness. However, those frequently used terms, efficiency and effectiveness, are often used with a somewhat vague sense of meaning in the health care context. *Efficiency* generally refers to using the minimum number of inputs for a given number of outputs. Efficient care, therefore, means a health care facility produces a given level of care or quantity that meets an acceptable standard of quality, using the minimum combination of resources. In performance literature, *efficiency* and *productivity* are often used interchangeably. While productivity generally connotes a broader meaning, both terms are considered a component of performance. As conceptualized in Fig. 1.1, research studies suggest that improving efficiency should lead to greater health service performance, while holding constant the quality, staff skill-mix, and case-mix. *Effectiveness*, more specifically, evaluates the outcomes of medical care and can be affected by efficiency or can influence efficiency as well as have an impact on the health service performance. For instance, *effectiveness* encourages us to ask if the necessary inputs are being used in order to produce the best possible outcomes. A hospital can be efficient, but not effective; it can also be effective, but not efficient. The aim is to be both.

Health care organizations will continue to face turbulent times and more intense competition. Health care managers must face up to promoting and improving performance within their institutions if they are to survive. There is not a standard

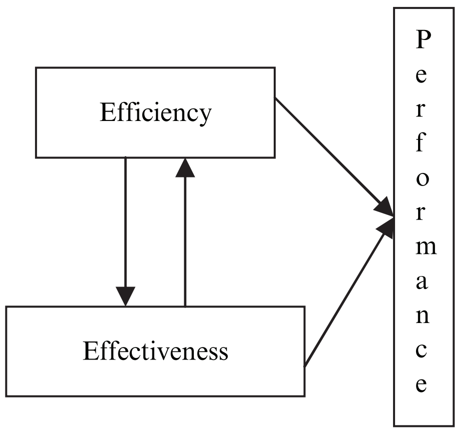

Fig. 1.1 Components of performance

formula for improving performance. Each health care organization, service and/or procedure must be examined individually. In some areas, the organization may have to increase the inputs used to improve quality. In other areas more must be done with fewer resources while holding quality constant. Health care managers will always be challenged with one of the most difficult tasks, determining the proper mix of inputs and outputs.

The relationship between efficiency and quality of care has had mixed results in prior studies. Singaroyan et al. (2006) study concluded that improving quality of health care may not always lead to efficient operations. On the other hand, Helling et al. (2006) found that increasing efficiency will result in quality. Mobley and Magnussen (2002) indicated that poor quality outcome is associated with less efficiency. Ferrando et al. (2005) mentioned that with proper guidelines, hospitals can increase efficiency without affecting the quality of care.

Performance needs to be measured and compared across health care providers for several purposes, including:

– Detecting changes from one period to another
– Determining how organizations are functioning relative to others in a given competitive market (benchmarking or peer comparisons)
– Investigating deviations from plan

Performance in this context should be viewed as a *relative* phenomenon across health care organizations. Thus it can be compared across different providers at one point in time or it can be compared for the same provider across multiple points in time.

Table 1.1 illustrates the measurements of performance where efficiency and effectiveness are measured in time as well as across health care organizations, using efficiency and effectiveness scores (these will be explained later in Chap. 2). Performance scores range from 0.0 to 1.0, where 1.0 is the highest achievable. For the time being, let us assume that 0.90 is an acceptable performance criterion for either high efficiency or effectiveness.

In this example there is no question about the performance of Hospital 3, which held its efficiency and effectiveness score at the top for both periods. *Relative* to other hospitals, this particular hospital would be considered a *benchmark* health care organization. Conversely, the other hospitals relative to Hospital 3 had some performance issues. Hospital 4, although relatively inefficient and ineffective in Time 1,

Table 1.1 Multi-facility and multi-time performance comparison

Health care organization	Efficiency Time 1	Efficiency Time 2	Effectiveness Time 1	Effectiveness Time 2
Hospital 1	0.81	0.88	0.86	0.93
Hospital 2	1.00	0.84	0.84	0.91
Hospital 3	1.00	1.00	1.00	1.00
Hospital 4	0.78	0.94	0.86	0.96
Hospital 5	0.62	0.55	0.71	0.62

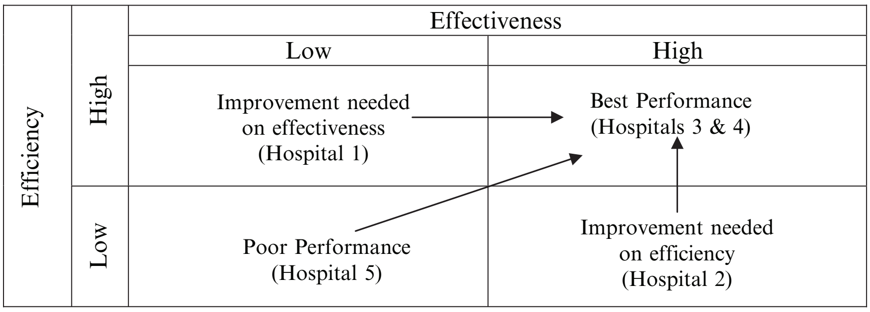

Fig. 1.2 Performance classification schema

closed this gap and became a high performer in Time 2. The situation for Hospital 1 is also promising. Both efficiency and effectiveness improved over time; however, this hospital needs more improvement on its efficiency to become as high a performer as Hospitals 3 and 4. Hospital 2 exhibits a mixed performance from Time 1 to Time 2, since its efficiency went down while effectiveness reached a relatively high standard. In the past, many health care managers argued this point, suggesting that to improve quality (effectiveness) something has to be taken away from efficiency. Of course, performance of Hospital 4 argues against this point, since both efficiency and effectiveness increased over time. Lastly, Hospital 5 was a poor performer in Time 1, and this poor performance was amplified in Time 2. Given these scenarios, one can classify the health care performance by these organizations into four groups based on their efficiency and effectiveness scores using Time 2 scores as shown in Table 1.1. Hospitals exhibiting less than high performance in either measure should aim towards the upper-right quadrant of the performance classification schema (Fig. 1.2).

The challenge of performance improvement planning is determining the values that yield efficiency and effectiveness scores, namely, what should health care managers do to improve the performance situation of the health care organization? This brings us to the methodologies that are used to calculate efficiency and effectiveness measures.

1.3 Performance Evaluation Methods

Comparative performance analysis can be undertaken by various methods, including:

- Ratio analysis,
- Least-squares regression (LSR),
- Total factor productivity (TFP),
- Stochastic frontier analysis (SFA), and
- Data envelopment analysis (DEA).

1.3.1 Ratio Analysis

As well as an effectiveness score, this approach is the simplest of the methods for calculating performance, especially productivity/efficiency. It produces information on the relationship between *one input* and *one output*. That is, efficiency is defined as the number of output units per unit of input:

$$Efficiency\ (Productivity) = \frac{Output}{Input} \tag{1.1}$$

Many ratios often have to be calculated to capture various dimensions of performance among compatible units or a given unit over different time periods. This is especially true for the hospital sector, where organizations such as MECON Inc. provide comparative benchmark and performance statistics via *MECON-Peer Guide* (1995).

The hospital industry reports, through such publications as *MECON-Peer Guide*, many inpatient as well as outpatient statistics displaying crude and adjusted patient volume for a given facility. These reports also characterize hospital operational information from labor, supply and cost points of view across the peer groups of hospitals.

Similarly, physician group practice performance statistics are reported by departmental and group levels for subscribing groups by various organizations such as Medical Group Management Association (MGMA). Either hospital or group practices receive these thick volumes of quarterly reports containing several hundred ratios to be monitored for benchmarking by health care managers.

Using multiple ratios often produces mixed results that confuse health care managers in comparative performance analysis. To illustrate this, let us examine the situation presented in Table 1.2, where we compare ten hospitals.

For simplicity, let us assume there are two inputs, nursing hours and medical supplies; and two outputs, inpatient admissions and outpatient visits. Using this information, one can calculate four possible performance ratios as illustrated in Table 1.3.

These ratios are analogous to what is being reported in hospital performance statistics by MECON Inc. and similar organizations.

In order to identify benchmarks (i.e., best performers) one can standardize each of these performance ratios across the hospitals by identifying the best score in each

Table 1.2 Hospital inputs and outputs

	Inputs		Outputs	
Provider ID	Nursing hours	Medical supplies ($)	Inpatient admissions	Outpatient visits
H1	567	2,678	409	211
H2	350	1,200	90	85
H3	445	1,616	295	186
H4	2,200	1,450	560	71
H5	450	890	195	94
H6	399	1,660	209	100
H7	156	3,102	108	57
H8	2,314	3,456	877	252
H9	560	4,000	189	310
H10	1,669	4,500	530	390

Table 1.3 Hospital performance ratios

Provider ID	Nursing hours/inpatient admissions	Medical supplies/inpatient admissions	Nursing hours/outpatient visit	Medical supplies/outpatient visits
H1	**1.39**	6.55	2.69	12.69
H2	3.89	13.33	4.12	14.12
H3	1.51	5.48	2.39	**8.69**
H4	3.93	**2.59**	30.99	20.42
H5	2.31	4.56	4.79	9.47
H6	1.91	7.94	3.99	16.60
H7	1.44	28.72	2.74	54.42
H8	2.64	3.94	9.18	13.71
H9	2.96	21.16	**1.81**	12.90
H10	3.15	8.49	4.28	11.54

Table 1.4 Standardized efficiency ratios and ranking of the hospitals

Provider ID	Nursing hours/inpatient admissions	Medical supplies/inpatient admissions	Nursing hours/outpatient visit	Medical supplies/outpatient visits
H1	**1.00 [1]**	0.40 [5]	0.67 [3]	0.68 [4]
H2	0.36 [9]	0.19 [8]	0.44 [6]	0.62 [7]
H3	0.92 [3]	0.47 [4]	0.76 [2]	**1.00 [1]**
H4	0.35 [10]	**1.00 [1]**	0.06 [10]	0.43 [9]
H5	0.60 [5]	0.57 [3]	0.38 [8]	0.92 [2]
H6	0.73 [4]	0.33 [6]	0.45 [5]	0.52 [8]
H7	0.96 [2]	0.09 [10]	0.66 [4]	0.16 [10]
H8	0.53 [6]	0.66 [2]	0.20 [9]	0.63 [6]
H9	0.47 [7]	0.12 [9]	**1.00 [1]**	0.67 [5]
H10	0.44 [8]	0.30 [7]	0.42 [7]	0.75 [3]

ratio, then dividing this into the particular ratio of each hospital. For example, Hospital 1 (H1) has the best ratio for the Nursing Hours per Inpatient Admission, which is 1.39. Dividing this into other hospitals' nursing hours per Inpatient Admissions we can obtain a relative value compared to H1, which is considered a benchmark hospital for this particular ratio. We can label this relative benchmarking score as the standardized efficiency ratio. Table 1.4 depicts the standardized efficiency ratios for four categories. Based on the relative scores of each hospital one can rank the hospitals (shown in brackets in Table 1.4). This case further illustrates the dilemma for the health care managers that occurs when benchmark performance shown with rankings varies according to which ratio is under consideration.

For example, while H1 is considered a benchmark hospital for nursing hours per inpatient admissions, it ranks fifth on "medical supplies per inpatient admissions," third on "nursing hours per outpatient visit," and 4th on "medical supplies per outpatient visit." On the other hand, H4 displays more dramatic results: while ranking first on "medical supplies per inpatient admissions," it ranks tenth on "nursing hours per inpatient admissions" as well as "nursing hours per outpatient visits," and nineth on "medical supplies per outpatient visits." Similar mixed results can be interpreted from Table 1.4 for H9, which is a benchmark hospital, on "nursing hours per outpatient visits," and for H3, which ranks highest on "medical supplies per outpatient visits."

This illustrates the weakness of ratio-based analysis, where health care mangers often cannot pinpoint a consistent benchmark incorporating all inputs and outputs of the health care organization.

1.3.2 The Least-Squares Regression

The least-squares regression (LSR) is a very popular parametric technique, and by its formulation, it assumes that all health care organizations are efficient. While it can accommodate multiple inputs and outputs, it can also account for noise, using an error term (see "e" in (1.5)). A general formula for a least squares regression is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots \beta_n x_n + e \qquad (1.5)$$

For this model, it is further assumed that

- For any fixed value of x, y is a random variable $(y|x) = \beta_0 + \beta_1 x$,
- The y values are independent of one another,
- The mean value of y is a straight-line function of x, $y = \beta_0 + \beta_1 x_1 + e$,
- The variance of y is the same for any x, and
- y has a normal distribution for any fixed value of x.

The least-square regression has some benefits. It can be used to measure technical change if time-series data are used. In addition, scale economies can be calculated. However, its weaknesses are greater.

Using LSR in performance analysis poses many weaknesses. Firstly, the LSR uses central tendency measures (averaging techniques), which are not necessarily efficient relationships. Furthermore, LSR does not identify the individual inefficient units, and it requires a pre-specified production function due to it is parametric formulation.

Let us illustrate these weaknesses using the example developed in Sect. 1.2. Consider the first two ratios where nursing hours and medical supplies per inpatient admissions were calculated. Using these two ratios one can map the hospitals on a scatter diagram, as shown in Fig. 1.3 to analyze the hospital performance from an inpatient admissions perspective (let us label this "Hospital Performance I").

We established earlier that H1 was the best hospital considering "nursing hours per inpatient admissions," while H4 was the best based on "medical supplies per inpatient admissions." Using regression analysis, an estimate of hospital performance from the inpatient admissions perspective (Hospital Performance I), is described by line $y = 13.83 - 1.42x$, as show in Fig. 1.3. This average line best predicts efficiency relationships when observations in a scatter diagram are closer to the estimated line. Hence, H2, H6 and H10 are the closest hospitals to this line while H1 and H4 show further distance. Thus, according to regression analysis, for better performance H1

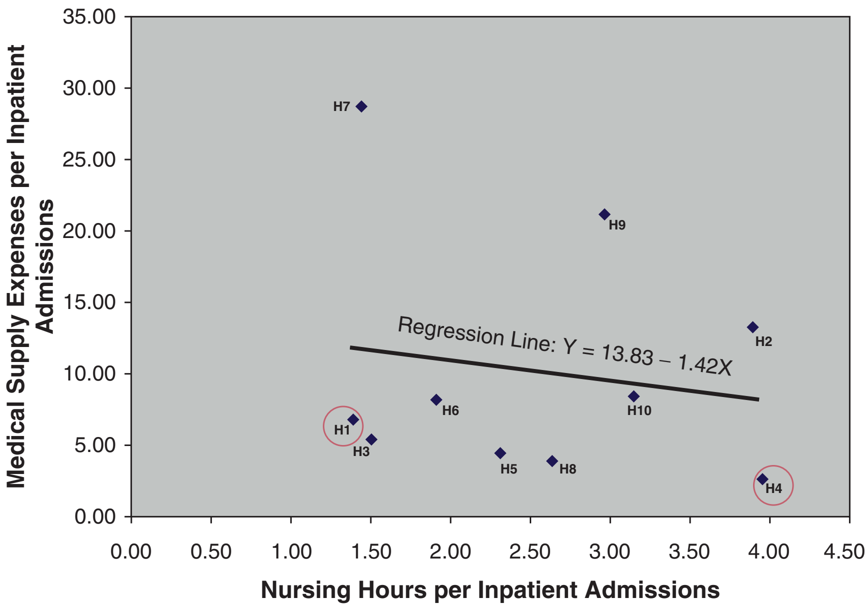

Fig. 1.3 Hospital Performance I

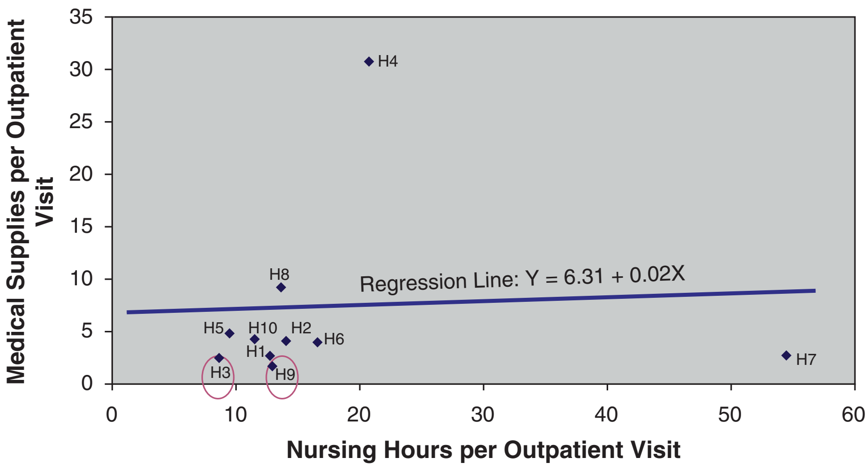

Fig. 1.4 Hospital Performance II

and H4 should further move toward the average, as illustrated by the regression line. In reality, this means H1 and H4 should give up their benchmark status with respect to these ratios and actually become inefficient.

We can replicate the same evaluation for the second dimension of the performance using a regression estimate of hospital performance from the outpatient visits perspective (Hospital Performance II). This case is described by line $y = 6.31 - 0.02x$, as show in Fig. 1.4. As it can be interpreted from this figure, H3 and H9, which were considered benchmark hospitals for this dimension of the hospital performance, are not good examples of performance based on the regression line because they are further away from the average performance with respect to H2, H5, H6, H8, and H10.

As these two examples illustrate, the regression analysis does not necessarily predict the best performance or the most efficient relationships. Hence, we must explore other methodologies that would describe more robust performance measures.

1.3.3 Total Factor Productivity (TFP)

TFP overcomes the weakness of single ratio analysis and incorporates multiple inputs/outputs into a single performance ratio. More specifically, TFP is measured using index numbers. Index numbers can be used to measure price and quantity changes over time, and also measure differences across health care organizations.

$$TFP_{ab} = \frac{\sum_{i=1}^{N} p_{ib} q_{ib}}{\sum_{i=1}^{N} p_{ia} q_{ia}} \tag{1.6}$$

In formula (1.2) TFP_{ab} index measures the change in the value of selected quantities of N outputs from period "a" to "b", where *p* represents the prices of these outputs. The most commonly used indices are: Laspeyres index, Pasche index, Fisher index, Tornqvist index, and Malmquist index. The difference between the Laspeyres and Pasche indices is whether the base period or current period quantities are used as weights. To overcome this difference, the Fisher index uses a geometric mean of the Laspeyres and Pasche indices. Similarly, Tornqvist index uses various geometric averages for price and quantity.

The Laspeyres, Pasche, Fisher and Tornqvist indices are non-parametric techniques that can be used with panel or cross-sectional data to measure the performance of two health care organizations in one time period or performance of one health care organization in two time periods. However, when more than two health care organizations needed to be compared at the same time or over time, these methodologies are not useful. Since TFP is not commonly used by the health care industry, we will not elaborate on these four indices any further. Of the TFP measures, the most frequently used method in health care is the Malmquist index.

The Malmquist index overcomes some of the shortcomings of the other indices discussed above. With the Malmquist index, health care managers can compare many organizations across two time periods. The Malmquist index can be obtained through frontier approaches such as DEA or SFA. The Malmquist index does not assume that all firms are efficient nor require price data.

Malmquist index numbers can be defined using either the output-oriented approach or the input-oriented approach. An important feature of the DEA Malmquist index is that it can decompose the overall efficiency measure into two mutually exclusive components, one measuring change in technical efficiency (catching-up effect) and the other measuring change in technology (innovation). In Chap. 6, we will illustrate the use of Malmquist index for hospital performance in multi-periods.

1.3.4 Stochastic Frontier Analysis (SFA)

SFA is also a parametric technique. SFA assumes that all firms are not efficient (this is improvement over LSR) and accounts for noise.

A general stochastic frontier model can be formulated as:

$$\mathrm{TC} = \mathrm{TC}(\mathrm{Y}, \mathrm{W}) + \mathrm{V} + \mathrm{U} \tag{1.7}$$

where TC=total cost

Y = output
W = input prices
V = random error assumed normally distributed with zero mean and variance
U = the inefficiency residual.

SFA can be used to conduct tests of hypotheses. It can also be used to measure technical efficiency, scale economies, allocative efficiencies, technical change, and TFP change (if panel data are available). However, SFA requires input and output quantities for empirical estimation of production functions. It can also be used to analyze panel or cross-sectional data.

SFA comes with certain shortcomings as well. For example, it requires specification of functional form and specification of a distributional form for the inefficiency term, U in (1.7). With the use of price information as well as quantity information, additional measurement errors may be added to the results (Kooreman, 1994). The resulting inefficiency may be due to technical or allocative inefficiency or combination of both. These two sources of inefficiencies cannot be separated, which is prudent since such knowledge might illustrate the need for different policy actions (Kooreman, 1994).

1.3.5 Data Envelopment Analysis (DEA)

DEA is non-parametric technique. DEA assumes that not all firms are efficient. It allows multiple inputs and outputs to be used in a linear programming model that develops a single score of efficiency for *each* observation used to measure technical efficiency, scale efficiency, allocative efficiency, congestion efficiency, technical change and TFP change (if panel data available and Malmquist indices calculated).

DEA requires input and output quantities if production efficiency is examined and can be used with both cross-sectional and panel data.

DEA does not account for noise due to its deterministic nature (deviation from the frontier is a result of inefficient operations). However, researchers are currently developing stochastic and other variants of DEA models that incorporate a random error component.

Since the DEA is considered as the main performance evaluation methodology considered in this book, the remaining chapters will illustrate the various DEA models and their applications.

1.4 Measurement Difficulties in Health Care

Measurement of the variables that describe the true nature of service production is an important prerequisite for performance measurement. In health care, due to the nature of the services provided, it is often difficult to find the appropriate variables and their measurements. Of course this depends on the level of analysis and whether it is carried out at the hospital level or the departmental level. Often, the departmental level measurements cannot be aggregated to the hospital level. For example, unit measures in a laboratory are different than in radiology or in nursing units. Thus, when hospital level measures are considered, what has been included in service production measures might be considerably different if these evaluations are carried out at the departmental level. For instance, performance of laboratories or radiology services across hospitals can be carried out as long as the measurements are consistent for each department.

Defining and measuring the output at the hospital level varies considerably across providers by the volume and scope of services provided, and also by patients' severity. Thus appropriate adjustments, such as case-mix adjustment, should be undertaken. In addition, outputs such as education, research, and certain community services may not be available in all hospitals. Lack of homogeneity in outputs produced and scale of operations may force one to conduct the performance analysis on those facilities considered peer-group organizations. Similarly, defining and measuring the inputs may pose difficulties as well. For example, differences may arise in pricing of input units, supply and materials or labor costs across facilities depending upon region. Similarly, capital assets valuation, depending upon when these are acquired and what type of depreciation rates are used, may render great variations in inputs. These issues will be further visited as various performance measurement applications are presented in ensuing chapters.

1.5 Summary

This chapter has introduced concepts of performance measurement in health care organizations. These included two dimensions of performance; efficiency and effectiveness (quality). To evaluate the performance, a survey of methods was provided including their strengths and weaknesses. These methods include: ratio analysis, the least squares regression, total productivity indices including Malmquist index, SFA and DEA. In what follows, we describe the various DEA models and their extensive use for performance evaluation in health care.