Seeking interactions between patient satisfaction and efficiency in primary healthcare: cluster and DEA analysis

Panagiotis Mitropoulos*,
Nikolaos Mastrogiannis and
Ioannis Mitropoulos

Department of Business Planning and Information Systems,
Technological Education Institute of Patras,
GR-26500 Patras, Greece
E-mail: pmitro@upatras.gr
E-mail: nmastro@upatras.gr
E-mail: mitro@teipat.gr
*Corresponding author

Abstract: This study investigates efficiency and quality of care in primary healthcare centres. At first, we performed cluster analysis to group health centres by their level of quality based on user's experiences that were collected from a satisfaction survey over a representative sample of patients. Secondly, a DEA analysis was conducted by generating bootstrapped efficiency scores for each health centre, and then we regressed these scores on variables that represent clusters from certain dimensions of patient satisfaction. The results of this study indicate statistically significant relations of health centres' efficiency with several dimensions of patient satisfaction. More specifically we identified direct relations of efficiency with accessibility, lab services and facilities. On the other hand there exist inverse relationships of the health centres efficiency with patients' opinions regarding scheduling appointment and doctor services.

Keywords: clustering; multi-criteria analysis; quality; bootstrap DEA; truncated regression.

Reference to this paper should be made as follows: Mitropoulos, P., Mastrogiannis, N. and Mitropoulos, I. (2014) 'Seeking interactions between patient satisfaction and efficiency in primary healthcare: cluster and DEA analysis', *Int. J. Multicriteria Decision Making*, Vol. 4, No. 3, pp.234–251.

Biographical notes: Panagiotis Mitropoulos is an Adjunct Assistant Professor at the Department of Business Planning and Information Systems of Technological Educational Institute of Patras, Greece. He possesses BS in Mathematics and PhD in Operation Research. His research interests cover a broad area of OR/MS, focusing mainly on location analysis, multiple criteria decision making and decision support system design.

Nikolaos Mastrogiannis received his BS in Mathematics from the University of Patras, Greece in 1998 and his MS in Decision Sciences from the Athens University of Economics and Business in 2000. He also received his PhD from the Department of Business Administration, University of Patras in 2009 and currently works as a Substitute Professor at the Department of Business Planning and Information Systems, Technological Educational Institute of Patras, Greece. His research interests include data mining and operations research.

Ioannis Mitropoulos is the Coordinator of the Laboratory of Decision Making and Business Planning at the Department of Business Planning and Information Systems, Technological Educational Institute of Patras, Greece. He also teaches at the Hellenic Open University in Postgraduate Studies in the field of Health Care Management. He possesses BS in Mathematics and PhD in Operations Research. His research interests cover a broad area of OR/MS, focusing mainly on location-allocation problems, multiple criteria decision analysis, performance measurement, evaluation of efficiency and statistics.

1 Introduction

In developed countries, quality has become one of the central issues in efforts to measure and improve health system performance. However it is difficult to obtain precise measurement of quality since the complexity of quality indicators are difficult to capture by a single measure. In addition, different bases for the construction of indicators further complicate comparison between them. Thus, quality of healthcare can be operationalised in different ways.

During the last decade, healthcare managers, politicians, and other decision makers have emphasised the importance of the patient perspective as an indicator of quality of healthcare (Legido-Quigley et al., 2008). Thus, patient satisfaction is one of the major factors of certification in measuring quality of health services.

Satisfaction can be defined as the extent of an individual's experience compared with his or her expectations (Pascoe, 1983). Patients' satisfaction is related to the extent to which general healthcare needs and condition-specific needs are met. Therefore the ultimate goal of patient satisfaction assessment is to improve the quality of healthcare service delivery.

Typically, variation in patient satisfaction between different healthcare units is thought to reflect differences in efficiency and other organisational factors. However, the amount of literature investigating variability in patient satisfaction with hospital care and its association with organisational factors is limited. Reflecting these concerns, this paper focuses on the role of patient satisfaction in the efficiency variation of the primary healthcare centres.

As such, this paper attempts in a first stage to evaluate the quality of Health Centres (HCs) in Cyprus using a patient satisfaction questionnaire and organise the results obtained by means of the cluster analysis technique. Then in second stage, we examine the association between quality and efficiency in primary healthcare centres.

The measurement of satisfaction in HCs was performed using a structured questionnaire with 51 questions which evaluated various parameters of HC satisfaction and administering it to evaluate patient's views in a specific healthcare setting. The questionnaire consists of closed-ended questions measured on a 5-point Likert scale with 1 – indicating 'strongly disagree' and 5 – corresponding to 'strongly agree'.

In this paper a cluster analysis, with values obtained from the patient satisfaction questionnaire, was performed to define cluster membership from different dimensions of pertinent satisfaction. Cluster analysis is a branch in statistical multivariate analysis and unsupervised learning in pattern recognition. The essence of clustering is to partition a set of objects into disjoint and homogeneous clusters, such that objects belonging to the same cluster are more similar to each other than those belonging to different clusters (Jain et al., 1999). Objects to be clustered are represented by a set of attributes, thus an object is considered as a conjunction of attribute values. From a practical perspective, clustering plays an outstanding role in many applications, such as scientific data exploration, information retrieval and text mining, spatial database applications, web analysis, marketing, medical diagnostics, computational biology, and many others.

Conventionally, most clustering algorithms are procedures that minimise total dissimilarity, with k-means (MacQueen, 1967) being one of the most popular algorithms because of its efficiency in clustering large data sets. However, the k-means algorithm only works on numeric data, i.e., the variables are measured on a ratio scale, because it minimises a cost function by changing the means of clusters. This prohibits it from being used in applications where categorical data are involved. The k-modes algorithm (Huang, 1998), on the other hand, extends the k-means paradigm to cluster categorical data by using

1 a simple matching dissimilarity measure for categorical objects (Kaufman and Rousseeuw, 1990)

2 modes instead of means for clusters

3 a frequency-based method to update modes in the k-means fashion clustering process to minimise the clustering cost function.

With these extensions, the k-modes algorithm enables the clustering of categorical data in a fashion similar to k-means.

In order to conduct the cluster analysis in our dataset, which includes categorical data, we have implemented the CLEKMODES clustering algorithm (Mastrogiannis et al., 2009). CLEKMODES is a novel and robust modification of the previously mentioned k-modes algorithm, that incorporates a four-step dissimilarity measure, which is based on elements of the methodological framework of the ELECTRE methods (Bouyssou, 2001; Pirlot, 1997; Roy, 1968, 1991, 1996), and in particular ELECTRE I multicriteria method (Roy, 1968) in its ELECTRE Iv version (Figueira et al., 2005). The essence of this dissimilarity measure is to compare the value of each attribute of a categorical object with the corresponding attribute value of every mode (centroid). Through these comparisons, a suitable similarity relation is developed between the object and each of the modes. This similarity relation is based on the following three principles:

1 the degree of resemblance between the object and the mode, identified according to a difference which is based on the number of matches/mismatches in their common attributes

2 the strength of the mode in terms of the attributes that positively influence the above resemblance

3 the existence of attributes in which the difference between the object and the mode is considered so significant that the object cannot be assigned to the cluster signified by the mode.

The first two principles, representing the strength of the majority of the attributes, result in choosing the best of the modes, enabling us to assign the object to the cluster of the chosen mode. On the other hand, the third principle confirms or rejects the above assignment.

Furthermore, the analysis on HCs efficiency is conducted with the use data envelopment analysis (DEA) to evaluate how the HCs manage their resources. Then we study in a second stage how these efficiencies are affected from factors that concern the clusters obtained from patient satisfaction results. Determining how these variables influence on efficiency is essential for determining performance and quality improvement strategies.

The structure of the paper is as follows. The second section describes in brief the CLEKMODES algorithm. The third section describes the DEA method and explains the DEA bootstrapping procedures. The fourth section illustrates the data and variables that were used in this paper. The fifth section reports and analyses the patient satisfaction clusters and DEA efficiency scores and describes their relation. Finally, our conclusions are given.

2 The CLEKMODES algorithm

Let $Y = \{y_1, y_2, \ldots, y_n\}$ be a dataset of categorical objects evaluated in a family $F = \{a_1, a_2, \ldots, a_d\}$ of attributes. Each attribute is different than the others, describing only a part of the clustering problem. Thus, every attribute is assigned a weight that is calculated within the main steps of the algorithm. This weight incorporates the strength and importance of the attribute in the clustering process as a whole. CLEKMODES is an iterative process, realised in four phases:

1 Selection of the k initial modes, one for each cluster.

2 Assignment of each object to the proper cluster according to a four-step dissimilarity measure. After each assignment, the mode of the cluster is updated.

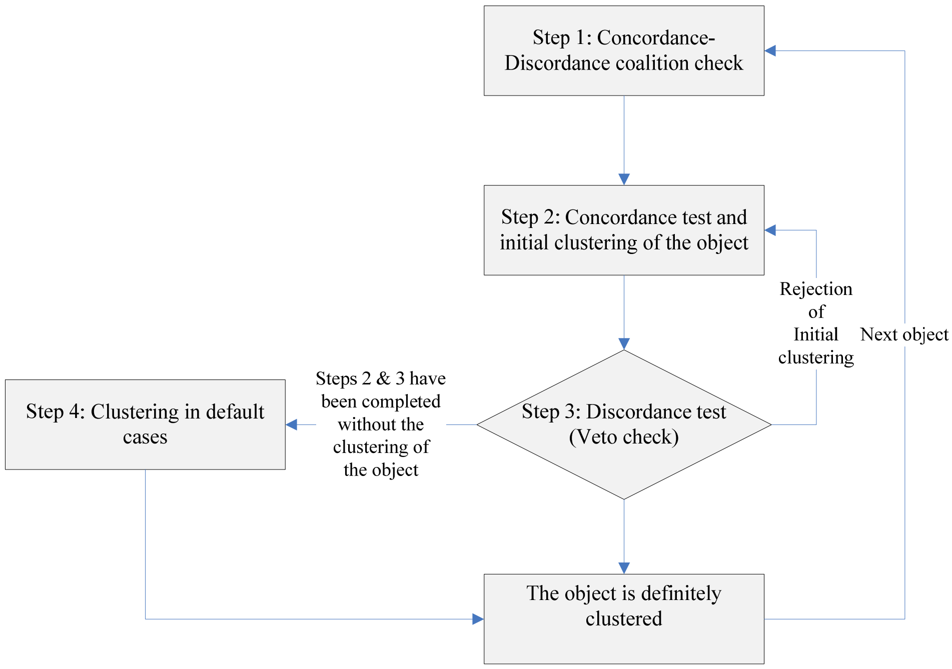
3 After all objects have been assigned to clusters, the dissimilarity of objects against the current modes is retested. If an object is found such that its most suitable mode belongs to another cluster rather than its current one, the object must be re-assigned to that cluster and the modes of both clusters must be updated.

4 Phase 3 is repeated until no object has changed clusters after a full cycle test of the whole dataset.

Let $KEN = \{x_1, x_2, \ldots, x_k\}$ be the set of the randomly selected initial modes also evaluated in the family $F = \{a_1, a_2, \ldots, a_d\}$ of attributes. These initial modes are k distinct objects from the dataset. Taking into consideration that the method is an iterative process, the clusters' modes will be re-estimated and thus *KEN* will be updated as many times as required for the algorithm to converge, according to the above four phases. The process of updating the mode of a cluster is based on Huang (1998). Note that according to this process, a mode is not necessarily an element of Y.

The dissimilarity measure, which is applied in Phases 2 and 3, consists of four steps and is described in the flow diagram of Figure 1. In Step 1, we identify the resemblance of the object with each of the k modes, according to a difference which is based on the number of matches/mismatches in their common attributes. In Step 2, we calculate the concordance indices and the corresponding thresholds. Furthermore, through the concordance test, which is based on the weights of the attributes, we assign initially the object to the proper cluster. In Step 3, the discordance test confirms or rejects the above assignment, considering the possible objections of some of the attributes. A final Step 4 deals with possible default cases. Note that Steps 1–3, correspond to the three principles in the introduction.

Figure 1 The flow diagram of the dissimilarity measure in CLEKMODES (see online version for colours)



2.1 Description of the dissimilarity measure

2.1.1 Step 1: concordance – discordance coalition check

The purpose of Step 1 is to determine the resemblance between each object y_i, $i = 1, 2,\ldots, n$ and each of the modes x_j, $j = 1, 2,\ldots, k$, through a series of pairwise comparisons on every attribute $a_f \in F, f = 1, 2,\ldots, d$ that evaluates both y_i and x_j. In order to determine the above resemblance, a proper difference based on the concept of pairwise comparisons, is defined. This difference is based on the dissimilarity measure of Huang (1998), which is an overlap measure (Stanfill and Waltz, 1986). An overlap measure finds similarity between two categorical attributes by assigning a similarity of 1 if the values are identical and a similarity of 0 if the values are not identical. For two multivariate categorical objects, the similarity between them will be directly proportional to the number of attributes in which they match, and thus their dissimilarity to the number of attributes in which they mismatch.

In other words, for each attribute common between object y_i and mode x_j, we calculate the difference:

$$D\left(x_{j,f}, y_{i,f}\right) = \frac{\left(m_{x_{j,f}} + m_{y_{i,f}}\right)}{\left(m_{x_{j,f}} \cdot m_{y_{i,f}}\right)} \cdot \delta\left(x_{j,f}, y_{i,f}\right) \tag{1}$$

where

$x_{j,f}$ is the value of mode x_j on attribute a_f

$y_{i,f}$ is the value of object y_i on attribute a_f

$m_{x_{j,f}}$ is the number of times $x_{j,f}$ appears in the set of the modes on attribute a_f

$m_{y_{i,f}}$ is the number of times $y_{i,f}$ appears in the set of the modes on attribute a_f

$$\delta\left(x_{j,f}, y_{i,f}\right) = \begin{cases} 0, \text{ if } x_{j,f} = y_{i,f} \\ 1, \text{ if } x_{j,f} \neq y_{i,f} \end{cases},$$

According to Huang (1998), the smaller the number of mismatches between the corresponding values of the attributes of an object y_i and a mode x_j, the more similar these two are. Even though this general principle is valid for relation (1), it does not consider the size of the difference itself in the similarity process. This means that, if $D(x_{j,f}, y_{i,f}) \neq 0$ on attribute a_f, this attribute is automatically considered as a drawback in terms of how similar the object and the mode are, whether the difference is large or small. This approach considers a small difference equally inadequate, in similarity terms, to a larger one. CLEKMODES expands the number of attributes that can contribute in the development of the similarity relation between an object y_i and a mode x_j. Thus, in addition to the attributes with zero value differences, some of the attributes with small non-zero value differences could also positively influence the similarity process. In order to achieve that, we check if:

$$D\left(x_{j,f}, y_{i,f}\right) \leq q_{i,f} \tag{2}$$

Parameter $q_{i,f}$ is an indifference threshold associated with object y_i and each attribute $a_f \in F$, namely an upper limit on the difference between object y_i and any of the k modes on attribute $a_f \in F$. The attributes that satisfy relation (2), belong to a concordance coalition $Con(i,j) = \{a_f \in F: D(x_{j,f}, y_{i,f}) \leq q_{i,f}\}$, that is, the set of attributes that support a strong resemblance between the object and the mode according to relation (1). All the other attributes belong to a discordance coalition $Dis(i,j) = \{a_f \in F: D(x_{j,f}, y_{i,f}) > q_{i,f}\}$ that opposes this resemblance. Furthermore, the attributes of the concordance coalition with only zero value differences belong to $ConZero(i,j)$, which in general is a subset of $Con(i,j)$.

2.1.2 Step 2: concordance test and initial clustering of the object

In order to identify the most suitable mode to properly cluster the object, two issues must be considered, with respect to the attributes of the concordance coalition. At first, the normalised sum of the weights of the attributes that belong to the concordance coalition must exceed a certain limit, in order to ensure that the mode is strong enough to cluster the object properly. Second, some of the concordance coalition attributes, namely the ones with zero value differences, have stronger similarity capabilities than the rest. Thus, it is important to strengthen these attributes in order to ensure their superiority in the process of identifying the mode that will properly cluster the object. The consolidation of the above two issues, results in defining the concordance index and the concordance threshold.

The concordance index for each mode $x_j, j = 1, 2,\ldots, k$ and each object y_j, $i = 1, 2,\ldots,n$ is calculated as follows:

$$CI\left(x_j, y_i\right) = \left[\left(\frac{\sum_{z\in Con(i,j)} w_z}{\sum_{h\in S} w_h}\right) + bonus * \left(\frac{\sum_{c\in ConZero(i,j)} w_c}{\sum_{z\in Con(i,j)} w_z}\right)\right] \tag{3}$$

where

$$S = Con(i,j) \cup Dis(i,j)..$$

It should be noted that the weights w_z, w_c and w_h of the attributes, are automatically calculated within the clustering process, and are described in detail in Mastrogiannis et al. (2009).

The first fraction of relation 3 is the normalised sum of weights of the concordance coalition attributes. The second fraction is the normalised sum of weights of the concordance coalition attributes that have zero value differences. Its multiplication with bonus, will result in the increase of the value of the concordance index when there are attributes with zero value differences. A large value of the concordance index implies that there are enough concordance coalition attributes with a strong total weight and some of them with zero differences in their pairwise comparisons. Hence, the corresponding mode is more similar to the object and more appropriate to cluster the object.

Following the same logic, the concordance threshold $CT(x_j, y_i)$ corresponding to the concordance index $CI(x_j, y_i)$, is calculated as follows:

$$CT\left(x_j, y_i\right) = \left[m + bonus * \left(\frac{\sum_{c\in ConZero(i,j)} w_c}{\sum_{z\in Con(i,j)} w_z}\right)\right] \tag{4}$$

where parameter m, which after extensive experimentation is valued at 70%, represents a lower limit of the normalised sum of the weights of the attributes that belong to the concordance coalition.

The concordance threshold represents the limit the concordance index must exceed in order to ensure a valid assignment of the object to the cluster of the mode that corresponds to this index. Thus, in order to identify the best mode for the clustering of the object, we must first sort the concordance indices in descending order, which means that the clustering strength of the corresponding modes decreases as the value of the indices decrease. It should be noted that parameter bonus in relations (3) and (4) constitutes an essential factor in order to properly sort the concordance indices in descending order. In particular, bonus aims at increasing the value of the concordance indices according to the number of the concordance coalition attributes with zero value differences, and its proportion in respect to the number of the concordance coalition attributes as a whole. A detailed description is given in Mastrogiannis et al. (2009).

Having set the concordance indices in descending order, the one with the larger value is compared to its corresponding concordance threshold. If the larger concordance index is at least equal to the compared threshold, the mode that corresponds to this index will assign the object to its cluster. If not, the process continues with the other indices of the descending order. This comparison represents the concordance test, given by the following relation:

$$CI_t\left(x_j, y_i\right) \geq CT_t\left(x_j, y_i\right) \tag{5}$$

Parameter $t = 1,\ldots, k$ denotes the positions of the descending order of the concordance indices and the corresponding thresholds and consequently the descending order of the modes, considering their clustering capabilities. Clearly, inequality (5) is satisfied if the normalised sum of weights of the concordance coalition attributes exceeds parameter m. As the process of finding the appropriate concordance index continues the strength of the indices decrease. Consequently the smaller the value of a concordance index, the larger the possibility of a false clustering for the object.

2.1.3 Step 3: discordance test (veto check)

After the clustering of the object is completed, we must examine the strength of the indications against the similarity of object y_i towards the selected mode x_j^*, through the discordance test. The discordance test examines the ability of the minority of the attributes (the ones that belong to the discordance coalition) to set veto to the decision of the majority. Thus, if:

$$D\left(x_{j,f}^*, y_{i,f}\right) < U_{i,f} \tag{6}$$

where $U_{i,f}$ is the veto threshold (Mastrogiannis et al., 2009), for each attribute a_f that belongs to the discordance coalition, then the initial clustering that took part in Step 2 is confirmed and the object is formally assigned to the proper cluster. It is quite possible though that at least one of the attributes does not satisfy the discordance test. In such a case the initial clustering is not valid. This means that the concordance index, and consequently the mode that lies behind it, chosen by the concordance test, is excluded from the clustering process. The method then starts again from Step 2. In other words, until a proper concordance index and its corresponding concordance threshold that will satisfy both Step 2 and Step 3 are found, the process continues by choosing and checking the next best concordance index and concordance threshold.

2.1.4 Step 4: Clustering in default cases

In some extremely rare cases, no clustering can be made through the concordance and discordance tests. This possible deficiency is resolved by introducing an alternative process, which is based on the absolute values of the differences between the concordance indices and the corresponding thresholds. The object is assigned to the cluster of the mode that corresponds to the concordance index-threshold with the largest of the above differences.

3 DEA methodology

Healthcare organisations use a mix of inputs to produce a large set of outputs. In evaluating performance on one dimension we need to take account of performance on all the other dimensions. DEA is a non-parametric (i.e., distribution-free), multi-factor, productivity analysis tool that utilises multiple input and output measurements in evaluating relative efficiency. It is best suited to the comparison or benchmarking of a number of similar operational units.

Since the introduction of DEA methodology, a considerable number of researchers have applied it in the health service sector. For a review of this literature see Hollingsworth (2003) or Worthington (2004). More recently, Amado and Dyson (2008) provided a systematic review focusing on the evolution studies that assess the performance of primary care providers.

For the first stage analysis, we employed a constant returns to scale (CRS) model that was first introduced by Charnes et al. (1978). The CRS model generates the efficiency scores by means of a linear system, maximising the ratio of outputs over inputs. A CRS model generates the input reductions or output augmentations that should be applied to the inefficiency units to become efficient. A CRS model is consistent to the peculiarities of the primary healthcare system in Cyprus. In fact the HCs are affected by CRS since the regional standards tend to align the total costs to the number of patients. The choice for a CRS model is also supported by Banker et al. (1996), who argued that for samples with less than 50 units, this model should be preferred. Thus, CRS input orientation efficiency scores of each HC can be calculated by solving the following linear problem:

$$\theta_{CRS} = \min\left\{\theta > 0 \mid y \leq \sum_{i=1}^{n} \lambda_i y_i, \theta x \geq \sum_{i=1}^{n} \lambda_i, x_i, \lambda_i \geq 0, i = 1, \ldots n\right\} \tag{7}$$

In equation (7) the efficient level of input is defined by θx, which is the projection of an observed HC (x, y) on to the efficient frontier, while θ is a scalar and λ is a non-negative vector of constants specifying the optimal weights of inputs/outputs. This problem is solved N times for each HC. The value of θ_{CRS} obtained is the Technical Efficiency score for the i^{th} HC. In order to become efficient, technical efficiency gives the decrease of inputs, which an observed hospital at location (x, y) could undertake. In the case where $\theta_{CRS} = 1$, the HC is considered fully efficient.

However, the standard DEA approach has come under criticism owing to the potential bias of efficiency estimates. DEA is also known to be sensitive when outliers exist. This problem becomes accentuated when sample sizes are small (Olson and Vu, 2009). Therefore the accuracy of DEA results may be affected by the sampling variation of the estimated frontier. This means that the distances to the frontier are underestimated in the case where the best performers in the population are not included in the sample. This paper addresses these inherent limitations of DEA, by applying the smoothed bootstrap approach of Simar and Wilson (1998) which by combining the DEA model with bootstrapping techniques, enables us to provide bias-corrected estimates of DEA efficiency scores, as well as confidence intervals. The bias-corrected efficiency scores are preferred over the original DEA scores since bias-corrected efficiency scores are within the lower bound (LB) and upper bound (UB) of the DEA bootstrap confidence intervals whereas the original DEA scores do not indicate biasness in the original estimates.

In the empirical literature on efficiency assessment, it has been common practice to perform a second stage analysis aimed at investigating the determinants of efficiency. The motivation of this paper responds to the fact that we have to account for the impact of environmental variables on efficiency. Within the second stage of the analysis, hypothesis tests had been performed to verify if the of the patient satisfaction obtained in the HCs may have impact on their efficiency performances. Thus we can assume that the efficiency scores θ_i can be regressed with a vector of environmental variables z_i that affect HCs efficiency, a vector of estimated parameters β_i and some statistical noise ε_i presenting the following equation:

$$\theta_i = \beta_i z_i + \varepsilon_i \quad (8)$$

For estimating equation (8), the bootstrapped truncated regression, proposed by Simar and Wilson (2007), was implemented. In this stage of analysis, efficiency scores are left truncated by 1. This approach is preferable (Simar and Wilson, 2007) to the conventional procedures of regression (Tobit estimator, OLS, etc), because the latter have reduced reliability. This is due to the fact that the DEA efficiency estimates are serially correlated with error and explanatory factors. Thus, the basic model assumption required by regression analysis, that is independence within the sample, is violated and therefore the traditional procedures of regression cannot be used. To avoid this problem, Simar and Wilson (2007) included a generated dependent variable in the second stage of the regression by using a double-bootstrapping procedure. In this approach the bootstrap DEA scores derived in the first-stage analysis are regressed against a set of environmental variables using the maximum likelihood method to explain efficiency drivers.

The step-by-step bootstrapping DEA and truncated regression procedure that adopted in that paper is described in studies such as Mitropoulos et al. (2012b) and Barros and Assaf (2009) and hence we omit details here. The analysis was performed using frontier efficiency analysis with R (FEAR, version 1.0) software (Wilson, 2008).

4 Data and variables specifications

Primary care in Cyprus is mostly provided by the NHS, to provide preventive, health promotion and curative services. A significant part of the primary care in Cyprus is the 28 HCs that now operate. These HCs are adequately staffed, well-equipped units to provide basic outpatient medical, diagnostic and pharmaceutical services as well as 24-hour on call services (see: Golna et al., 2004).

The process of data collection from 14 HCs in Cyprus was performed in 2009, under the responsibility of the Open University of Cyprus. Primary data concern the patient responds to the satisfaction questionnaire over a sample of patients. The secondary data concern indicators about the operation characteristics in each HC that are collected from their medical records.

More specifically, in order to gather information and insight on the HCs patients' view of the services they provide we assessed patient satisfaction with a structured questionnaire in each HC. The questionnaire was anonymous and the person who conducted the interview was unknown to the patients, in order not to influence their responses. The interviews were conducted when the patients left the HCs.

We employ the outpatient satisfaction questionnaire that was developed by Aletras et al. (2007). In that study the reliability of the questionnaire indicate satisfactory internal consistency and short-term repeatability for all scales. The questionnaire includes 51 questions and the answers are mostly designated in a 5-point ordinal format, with higher scores indicating more positive experiences. Initially, the questionnaire created included 64 questions measuring satisfaction with individual aspects of the HC care. From these, we have excluded the 13 questions that do not correspond to the typical ordinal format (i.e., percentages), mostly representing demographic and social questions.

The questionnaire scales describe patients' assessment of six dimensions according to their experiences with care in HC. The dimensions of patient experiences can be described as follows:

1 appointment and scheduling, describing patient experiences with communication and provision of information by HC staff

2 accessibility, describing the satisfaction about convenience and ease of access

3 waiting conditions, describing the waiting room conditions and waiting time before examination

4 doctor services, indicating patient experiences with doctors' care and competence

5 lab services, describing the conditions that concern the laboratory and pharmaceutical care

6 facilities, the patient's overall perspective and satisfaction with infrastructure and equipment.

A total of 416 questionnaires were completed out of a total of 14 state HCs. The sampling was random, and questionnaires given to patients during their departure from the centre after the completion of their visit. The frequency distribution of the questionnaires was one every four examined patients.

The selection of input/output variables to run the DEA model follows primarily previous studies in the literature (Kontodimopoulos et al., 2007; Mitropoulos et al., 2012a). Data availability was also a factor in determining the list of inputs/outputs variables. These data included staff numbers according to specialty, i.e., physicians, nurses, paramedical, administrative and other support staff, patient visits and numbers of performed medical examinations.

In light of the above the inputs identified are the three categories of staff as the main providers of services, namely physicians, nursing/paramedical (which are comprised of aggregation of all healthcare workers who have special training in the performance of supportive healthcare tasks) and administrative/support staff (which are comprised of aggregation of all other non-medical workers in the HC (administrative personnel cleaners etc.). In this study, two outputs were selected to reflect production responsibilities of HCs. Specifically, aggregated scheduled and emergency outpatient visits and aggregated laboratory-radiographic tests and pharmaceutical actions performed. These two variables are reported as 'medical care' and 'paramedical care' respectively.

5 Results and discussion

5.1 Clustering results

We have conducted cluster analysis in seven datasets, that correspond to the initial 14 × 51 dataset and the six subsequent datasets derived by each of the previously mentioned dimensions. In these datasets, an object represents a HC, each attribute represents an item/question, and every attribute value corresponds to a patient's response. It should be noted, that every attribute value in these datasets, is a categorical value estimated as the average of several responses for a particular object/HC, approximated to the nearest integer. The seven datasets are shown in Table 1.

Table 1 Description of the datasets

Datasets	*No. of objects/HC*	*No. of attributes*
Dataset 1 (initial dataset)	14	51
Dataset 2 (1st dimension)	14	5
Dataset 3 (2nd dimension)	14	6
Dataset 4 (3rd dimension)	14	7
Dataset 5 (4th dimension)	14	10
Dataset 6 (5th dimension)	14	12
Dataset 7 (6th dimension)	14	11

In Table 2 we can see the clustering results of the above seven datasets, separated in two clusters (named 0 and 1, respectively). In order to strengthen the validity of our results, we have run each dataset, 25 times, using 25 randomly selected discrete sets of initial modes, and obtained as our results, the ones that appeared more frequently. In Table 3 we can see the percentages of the most frequent and second more frequent clustering results, and the apparent superiority of the first, which supports their validity and robustness.

Table 2 Clustering results

HC	*Dataset 1*	*Dataset 2*	*Dataset 3*	*Dataset 4*	*Dataset 5*	*Dataset 6*	*Dataset 7*
HC1	0	0	1	0	0	0	0
HC2	1	1	1	1	1	1	1
HC3	0	0	1	0	0	0	0
HC4	0	1	1	0	0	1	0
HC5	0	1	1	0	0	1	0
HC6	0	1	1	0	0	1	0
HC7	1	1	1	1	1	1	0
HC8	0	0	0	0	0	0	1
HC9	0	0	1	0	0	0	0
HC10	0	1	0	1	1	1	1
HC11	1	1	1	1	0	1	0
HC12	0	0	1	0	0	0	0
HC13	0	0	1	0	0	0	0
HC14	0	1	1	1	0	0	0

Table 3 Most frequent and second most frequent results

Datasets	*% appearance of most frequent clustering result*	*% appearance of second most frequent clustering result*
Dataset 1 (initial dataset)	64%	28%
Dataset 2 (1st dimension)	60%	28%
Dataset 3 (2nd dimension)	64%	24%
Dataset 4 (3rd dimension)	72%	12%
Dataset 5 (4th dimension)	44%	24%
Dataset 6 (5th dimension)	76%	16%
Dataset 7 (6th dimension)	68%	24%

Note: Number of runs = 25

Finally, we have conducted an empirical evaluation of each of the two clusters, based on the pairwise comparison of their final modes (after the completion of the clustering process). In particular, some of the attributes/questions of a mode are valued higher than the ones of the other mode. In such a case, we can empirically designate the mode with the higher average values as ‘better’ than the other mode. Given the above described ‘superiority’ of the final modes of one cluster in comparison to the other, in Table 2 we can also see that value ‘1’ corresponds to the superior cluster, and value ‘0’ corresponds to the other cluster, in all seven datasets.

Table 4 Bootstrapped efficiency results

HC	*Original DEA*	*Bias corrected*	*Bias*	*Upper bound*	*Lower bound*
HC 1	0.755	0.682	0.073	0.635	0.749
HC 2	1.000	0.845	0.155	0.799	0.992
HC 3	0.659	0.588	0.071	0.532	0.655
HC 4	0.873	0.789	0.084	0.742	0.866
HC 5	0.938	0.854	0.084	0.784	0.933
HC 6	1.000	0.879	0.121	0.839	0.990
HC 7	0.856	0.759	0.097	0.702	0.850
HC 8	0.734	0.658	0.076	0.617	0.727
HC 9	1.000	0.744	0.256	0.687	0.992
HC 10	0.670	0.623	0.047	0.580	0.668
HC 11	0.909	0.846	0.063	0.771	0.905
HC 12	1.000	0.856	0.144	0.805	0.992
HC 13	0.958	0.874	0.084	0.823	0.951
HC 14	0.537	0.497	0.040	0.461	0.535
Average	0.849	0.749	0.099	0.698	0.843

5.2 DEA results

The efficiency results are presented in Table 4. The first column provides the original CRS DEA-efficiency scores, the second column provides the DEA bootstrapped efficiency scores, the third column provides the bias of the original DEA, and the fourth and fifth columns represent the LB and UB of the DEA-bootstrap confidence intervals. A first look at the results indicates that the original DEA estimates lie for every block outside the estimated confidence intervals, while the bootstrapped DEA estimates lie for every block inside the confidence interval. This problem is due to the bias in the original estimates, and it is a main reason why the bootstrapped DEA scores are preferred to the original estimates.

Recall, that as a first step, we measure the efficiency score via DEA. We find that, on average, the efficiency measure for HCs is 84.9%, ranging between a low efficiency score of 53.7% and the highest efficient score of 100%. These outcomes are very common. Ozcan (2008) summarises the efficiency scores of a number of evaluation studies in healthcare. Most of these studies report scores near 90% depending on the DEA-variant chosen the distinct services and resources and sample.

It is evident from the first column in Table 4 that there are four efficient HCs on the frontier of best practices with a technical efficiency score equal to 1. However, when considering the bootstrapping results (column 2 of Table 4) none of the HCs appear to be close to the frontier. Since the bias is large relative to the variance in every case, the bootstrap estimates are preferred to the original estimates (Simar and Wilson, 1998). The original efficiency estimates lie also outside the estimated confidence intervals in the last two columns of Table 4 in every instance. This is due to the bias in the original estimates, and the fact that the confidence interval estimates correct for the bias. These results therefore reinforce the fact that the DEA bootstrap model is more superior to the traditional DEA model in estimating the efficiency scores of HCs.

The above measures provide an overview of the general efficiency of HCs in Cyprus. However, in order to account for the sources of efficiency changes we have also estimated the relationship between the DEA efficiency scores and patient satisfaction clusters. Relative to the second stage regression, we used the bootstrapping procedure discussed in Section 3 to overcome the serial correlation problem of the DEA-efficiency estimates. In this stage two models are considered. The first model represents as environmental variables the clusters from the seven dimensions which were addressed in the patient satisfaction survey. And the second model represents as environmental variables the overall patient satisfaction clusters. These models can be expressed as follows:

Model 1:

$$\begin{aligned}\theta_i = \beta_0 + \beta_1 cluster_1 + \beta_2 cluster_2 + \beta_3 cluster_3 + \beta_4 cluster_4 \\ + \beta_5 cluster_5 + \beta_6 cluster_6\end{aligned} \tag{9}$$

Model 2:

$$\theta_i = \beta_0 + \beta cluster_{(overall)} \tag{10}$$

where θ_i is the technical efficiency scores, $cluster_1$ to $cluster_6$ are the six clusters according to the six dimensions that addressed from the patient satisfaction survey. In model 2 the $cluster_{(overall)}$ represent the overall patient satisfaction clusters.

All clusters are represented in the regression analysis by dummy variables that take the value of one for the cluster with better patient satisfaction, and zero for cluster with lower patient satisfaction.

Table 5 Sources of CRS technical efficiency

Model 1					*Model 2*				
Variable	*Coefficient*	*z-statistic (p-value)*	*95% confidence interval*		*Variable*	*Coefficient*	*z-statistic (p-value)*	*95% confidence interval*	
			LB	*UB*				*LB*	*UB*
Constant	0.541	11.45 (0.000)	0.448	0.634	Constant	0.731	58.88 (0.000)	0.933	0.997
Appointment/scheduling	–0.260	–4.28 (0.000)	–0.379	0.141	Overall	0.085	1.84 (0.065)	–0.005	0.176
Accessibility	0.210	5.58 (0.000)	0.136	0.283					
Waiting conditions	0.006	0.22 (0.828)	–0.045	0.057					
Doctor services	–0.103	–3.01 (0.003)	–0.170	0.035					
Lab services	0.351	188.41 (0.000)	0.347	0.355					
Facilities	0.102	2.98 (0.003)	0.035	0.169					

Note: Number of iterations = 2,000

The estimation results of the truncated regression models presented in equation (9) and equation (10) are shown in Table 5. This table includes the value of each estimated coefficient, the z-statistic and the corresponding p-value as well as the LB and the UB of each estimated coefficient.

Our estimation in model 1 shows that the coefficients for appointment/scheduling and for doctor services are negative and statistically significant. Therefore the technical efficiency has inverse relationships in clusters that represent these two dimensions of patient satisfaction. In other words more satisfaction in these dimensions appears in less efficient units. A possible explanation for the effect that appeared in doctor services is that the patients turn to primary services seeking for a stable and more personal care. More satisfaction in the less efficient HCs may imply that doctors spend more time and have better personal acquaintance with their patients. Conversely, increased patient dissatisfaction and many complaints are due to breakdown in the doctor-patient relationship because of the busy environment that the efficient HCs operate. Similarly the procedure of getting an appointment seems more difficult and complicated in efficient but busy HCs.

On the other hand the coefficients of accessibility, lab services and facilities are positive and statically significant. A possible explanation for these results is that the more efficient HCs may have been supported with better equipment and infrastructure than the less efficient ones, in order to service the increased number of patients. Finally, the coefficient of waiting conditions has a positive sign but is not statistically significant.

In model 2, the coefficient of the overall patient satisfaction cluster is positive and not statically significant. This may be considered as an expected result since in model 2 the overall clusters are more general than the clusters inserted in model 1. Thus the clusters that correspond to the six dimensions are more specific and focused in analysing the patients' satisfaction.

6 Conclusions

Low quality within the healthcare industry has been a major concern; however, this situation is unlikely to improve without a general change in the way efficiency is managed. Patients' opinion and their satisfaction, is essential for the quality standards of the provided quality care. Identifying the aspect of care that influences patient satisfaction may be useful to design changes in health delivery system.

According to our results the satisfaction of users for health services provided, is determined by factors relating to organisational and functional characteristics, but also by the actual interpersonal relationship and communication with health professionals and particularly the doctor-patient relationship.

This study prioritises satisfaction dimensions in the HCs, in order to design changes that will improve the quality in primary healthcare. However, it is difficult to propose some changes in an institution, after a satisfaction survey, because dissatisfaction may have underlying causes (Hutchison, 1993). Therefore satisfaction studies are necessary (for feedback), in order to promote changes in the structural and operational frame of health organisations (Georgousi, 1999).

Consequently, additional research will be required to identify the appropriate strategies that could systematically address the different expectations and satisfaction of outcomes that we observed in this study. Moreover, it might be helpful to convince groups of patients to elicit their views about their concerns and what suggestions they may have to improve the HCs service quality.

Acknowledgements

This research has been co-funded by the European Union (European Social Fund) and Greek national resources under the framework of the 'Archimedes III' project of the 'Education and Lifelong Learning' Operational Programme.

References

Aletras, V., Zacharaki, F. and Niakas, D. (2007) 'Questionnaire for the measurement of outpatient satisfaction in the ophthalmology clinic of a Greek public hospital (in Greek)', *Archives of Hellenic Medicine*, Vol. 24, No. 1, pp.89–96.

Amado, C. and Dyson, R. (2008) 'On comparing the performance of primary care providers', *European Journal of Operational Research*, Vol. 185, No. 3, pp.915–932.

Banker, R.D., Chang, H. and Cooper, W.W. (1996) 'Simulation studies of efficiency, return to scale and misspecification with nonlinear functions in DEA', *Annals of Operational Research*, Vol. 66, No. 4, pp.233–253.

Barros, C.P. and Assaf, A. (2009) 'Bootstrapped efficiency measures of oil blocks in Angola', *Energy Policy*, Vol. 37, pp.4098–4103.

Bouyssou, D. (2001) 'Outranking methods', in Floudas, C.A. and Pardalos, P.M. (Eds.): *Encyclopedia of Optimization*, Vol. 4, pp.289–255, Kluwer.

Charnes, A., Cooper, W. and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, Vol. 2, No. 4, pp.429–444.

Figueira, J., Mousseau, V. and Roy, B. (2005) 'Electre methods', in Figueira, J., Greco, S. and Ehrogott, M. (Eds.): *Multiple Criteria Decision Analysis: State of the Art Surveys*, pp.133–153, Springer, New York.

Georgousi, E. (1999) 'Features of GP services in Greece in relation to other European countries', *Primary Health Care*, Vol. 11, No. 4, pp.193–202.

Golna, C., Pashardes, P., Allin, S., Theodorou, M., Merkur, S. and Mossialos, E. (2004) *Health care systems in transition: Cyprus*, WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies, Copenhagen, Denmark.

Hollingsworth, B. (2003) 'Non-parametric and parametric applications measuring efficiency in health care', *Health Care Management Science*, Vol. 6, No. 4, pp.203–218.

Huang, Z. (1998) 'Extensions to the k-means algorithms for clustering large data sets with categorical values', *Data Mining and Knowledge Discovery*, Vol. 2, No. 3, pp.283–304.

Hutchison, A. (1993) 'The role of patient satisfaction assessment in medical audit', *Scandinavian Journal of Primary Health Care*, Vol. 11, No. 1, pp.19–22.

Jain, A., Murty, N. and Flynn, J. (1999) 'Data clustering: a review', *ACM Computing Surveys*, Vol. 31, No. 3, pp.264–323.

Kaufman, L. and Rousseeuw, P. (1990) *Finding Groups in Data*, Wiley, New York.

Kontodimopoulos, N., Moschovakis, G., Aletras, V.H. and Niakas, D. (2007) 'The effect of environmental factors on technical and scale efficiency of primary health care providers in Greece', *Cost Effectiveness and Resource Allocation*, Vol. 5, No. 1, p.14, doi:10.1186/1478-7547-5-14.

Legido-Quigley, H., McKee, M., Nolte, E. and Glinos, I.A. (2008) *Assuring the Quality of Health Care in the European Union. A Case for Action*, p.5, WHO Regional Office for Europe: Observatory Studies Series No 12, Copenhagen.

MacQueen, J.B. (1967) 'Some methods for classification and analysis of multivariate observations', *Proceeding of the 5th Berkley Symposium on Mathematics, Statistics and Probability*, pp.281–297.

Mastrogiannis, N., Giannikos, I., Boutsinas, B. and Antzoulatos, G. (2009) 'CLEKMODES: a modified k-modes clustering algorithm', *Journal of the Operational Research Society*, Vol. 60, No. 8, pp.1085–1095.

Mitropoulos, P., Mitropoulos, I. and Giannikos, I. (2012a) 'Combining DEA with location analysis for the effective consolidation of services in the health sector', *Computers and Operations Research*, Article in press, doi: 10.1016/j.cor.2012.01.008.

Mitropoulos, P., Mitropoulos, I. and Sissouras, A. (2012b) 'Managing for efficiency in health care: the case of Greek public hospitals', *European Journal of Health Economics*, Article in press, doi: 10.1007/s10198-012-0437-0.

Olson, K. and Vu, L. (2009) 'Economic efficiency in farm households: trends, explanatory factors, and estimation methods', *Agricultural Economics*, Vol. 40, pp.87–99.

Ozcan, Y.A. (2008) *Health Care Benchmarking and Performance Evaluation: An Assessment Using Data Envelopment Analysis*, International Series in Operations Research & Management Science, Springer.

Pascoe, G.C. (1983) 'Patient satisfaction in primary health care: a literature review and analysis', *Evaluation and Program Planning*, Vol. 6, No. 3, pp.185–210.

Pirlot, M. (1997) 'A common framework for describing some outranking methods', *Journal of Multi-Criteria Decision Analysis*, Vol. 6, No. 2, pp.86–92.

Roy, B. (1968) *Classement et choix en présence de points de vue muptiples: La méthode ELECTRE*, Vol. 8, pp.57–75, R.I.R.O.

Roy, B. (1991) 'The outranking approach and the foundations of ELECTRE methods', *Theory and Decision*, Vol. 31, pp.49–73.

Roy, B. (1996) *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, Dordrecht.

Simar, L. and Wilson, P. (1998) 'Sensitive analysis of efficiency scores: how to bootstrap in nonparametric frontier models', *Management Science*, Vol. 44, No. 1, pp.49–61.

Simar, L. and Wilson, P. (2007) 'Estimation and inference in two-stage. Semi-parametric models of production processes', *Journal of Econometrics*, Vol. 136, No. 1, pp.31–64.

Stanfill, C. and Waltz, D. (1986) 'Toward memory-based reasoning', *Communications of the ACM*, Vol. 29, No. 12, pp.1213–1228.

Wilson, P. (2008) 'FEAR: a software package for frontier efficiency analysis with R', *Socio-Economic Planning Sciences*, Vol. 42, No. 4, pp.247–254.

Worthington, A. (2004) 'Frontier measurement efficiency measurement in health care: a review of empirical techniques and selected applications', *Medical Care Research and Review*, Vol. 61, No. 2, pp.135–170.