CONSTANTINE PORPHYROGENETUS INTERNATIONAL ASSOCIATION

MSRD

Journal of Management Sciences and Regional Development
Issue 4, July 2002
Correspondence: ikarkazis@aegean.gr

http://www.stt.aegean.gr/geopolab/GEOPOL%20PROFILE.htm
ISSN 1107-9819
Editor-in-Chief: Arie Reichel

MEASURING THE EFFICIENCY OF OUTPATIENT CLINICS VIA DEA: A CASE STUDY

Lea Friedman & Zilla Sinuany-Stern

Dept. of Industrial Engineering and Management
Ben-Gurion University of the Negev
Beer-Sheva, Israel

Abstract. In this paper we analyze the operational efficiency of a group of outpatient clinics in the largest sick fund in Israel. We use the Data Envelopment Analysis (DEA) to classify the clinics into two groups: efficient and inefficient. Moreover, we ranked the clinics utilizing four ranking methods: (I) the traditional discriminant analysis of two groups (DDEA). (II) the discriminant analysis of ratios (DR/DEA), (III) the cross efficiency method (CE/DEA), (IV) the canonical correlation analysis (CCA/DEA).
The goodness of fit between each method and the DEA was statistically validated. Furthermore, the compatibility among all the four ranking methods was verified to come up with an overall rank combining them into one rank (CO/DEA).
In our case study of 45 clinics, three inputs and three outputs were utilized. The potential improvements of each input and output were a by-product of DEA.

Key Words: Data Envelopment Analysis, Multicriteria Decision Analysis, Health care systems, clinics ranking.

INTRODUCTION

The Israeli health care system is composed of 4 major sick funds (see Barnoon and Pliskin, 1988); the largest is Kupat Holim Klalit (General Sick Fund—GSF).
Until 1995, as health insurance was voluntary in Israel, GSF was run by the Histadrut (The General Federation of Hebrew Workers in Israel), where the members of this workers' union were automatically members of GSF. The monthly membership fee for the union (Histadrut) included the GSF fee. The division between the union expenses and GSF was not clear. Consequently great inefficiencies were imbedded in this health system. At that time, GSF had an open door policy, while other sick funds were more selective (receiving many young and rich population in the center). Consequently, the services in GSF were of a lesser quality. Besides, the government hospitals and some private ones, only GSF had hospitals and traditionally had a large network of outpaptient clinics. Thus, GSF became over the years less and less efficient. While GSF invested mainly in the physical infrastructure, an extensive network of outpatient clinics, hospitals, etc. the other sick funds had a much limited owned building and equipment, but invested more in better private specialists.
In 1995 a new national health insurance act was passed—health care insurance became compulsory; citizens in Israel stopped paying the voluntary sick fund directly, and all were required to pay it with their social security bill. All sick funds were receiving their budget from the government according to the number of members. This policy brought a fierce competition among sick funds where each was recruiting members. Other sick funds realized that in order to increase the number of members, there is a need for more accessibility by opening a wider network of outpatient clinics nationwide. As a result of opening many competing clinics of various sick funds, many patients were tempted to transfer among them. Eventually, all sick funds found themselves in financial crisis.
(In 1998 the government decided to impose some charges on patients in specialists visits. Consequently a need for efficiency arises in all levels, especially on the network of outpatient clinics.)
GSF has conducted a Data Envelopment Analysis (DEA) study (Ben Yakov and Weidenfeld, 1997) on 45 clinics in the center of Israel, to test the possibility of utilizing DEA as a tool for measuring the relative efficiency of outpatient clinics. The variability of the weights in the DEA caused an uneasiness in receiving the results by the GSF Management. As a result we suggest here a further analysis of DEA by ranking the units via several methods. The ranking method is based on unified weights to all units.
Four ranking methods were utilized here to rank the 45 outpatient clinics, all the methods were found to be significantly compatible with the DEA classification. Furthermore, all were highly correlated with each other.
DEA was applied for health care systems mainly for hospitals, for example Sherman (1986), Byrnes and Valdmains (1994), Chilingerian and Sherman (1996), Fare, et al. (1994), Morey et al. (1995). None of these studies performs full ranking health care units in general, and specially for outpatient clinics.
In this paper we first describe the DEA model followed by the ranking methods and their validation. Afterwards the case study is described along with the results and their analysis.

DATA ENVELOPMENT ANALYSIS (DEA)

DEA was developed by Charnes et al. (1978) for measuring the efficiency of organizational units Basically, DEA classifies n organizational units into two groups—efficient and inefficient—based on given multiple inputs and multiple outputs.
Let x_{ij} be a given level of the i-th input of unit j (i=1...m, j=1...n) and y_{rj} a given level of the r-th output of unit j (r=1...s). For each unit k, the model calculates the optimal weights for the inputs (v_i^k) and the outputs (u_r^k) which maximizes the ratio between the weighted output and the weighted input.

$$h_{kj} = \sum_{r=1}^{s} u_r^k h_{rj} \Big/ \sum_{i=1}^{m} v_i^k x_{ij} \qquad \begin{array}{l} k = 1\ldots n \\ j = 1\ldots n \end{array}$$

DEA determines the optimal weights v_i^k and u_r^k by solving the following problem:

$$\begin{array}{ll} \text{Max} & h_{kk} \\ \text{s.t.} & h_{kj} \le 1 \\ & u_r^k \ge 0, \quad v_i^k \ge 0 \end{array}$$

These ratios for all the n units are bounded from above by one (to prevent trivial solutions) and the weights are all positive. Each unit k is assigned the highest possible efficiency score (ratio) by choosing the most favorable weights. Therefore, if a unit does not reach the maximum possible value (1) it is inefficient, otherwise it is efficient.
Obviously, in the DEA the values o the weights would differ from unit to unit, (v_i^k, u_r^k). This variability in the choice of weights characterizes the DEA. Therefore, we cannot perform a full rank of all the units based on the DEA scores since they come from non-unified weights.
The above ratio model can be formulated as a linear programming problem:

$$\begin{array}{ll} \text{Max} & \sum_{r=1}^{s} u_r^k y_{rj} \\ \text{s.t.} & \sum_{i=1}^{m} v_i^k x_{ik} = 1 \\ & \sum_{r=1}^{s} u_r^k y_{rk} - \sum_{i=1}^{m} v_i^k x_{ij} \le 0 \\ & u_r^k \ge 0 \ , \quad v_i^k \ge 0. \end{array}$$

The dual problem for unit k is:

$$\text{Min} \quad \theta_k$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^{+} = y_{rk} \qquad r=1,2..s$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_r^{-} + \theta_k = 0 \qquad i=1,\dots,m$$

$$\lambda_j \geq 0 \qquad j=1,\dots,n$$

where θ_k is the dual variable of the first constraint and λ_j are the duals of the rest of the constraints. The slack variables of the first set of dual constraints, S_r^+ r=1...s, represent the improvement needed for each output r of unit k. Similarly S_i^- represent the reduction needed for each input i of unit k.

RANKING MODELS

DEA provides the full efficient frontier, i.e., the pareto optimum solution. Obviously in DEA, the values of the weights differ from unit to unit (v_i^k, u_r^k).This variability in the choice of weights characterizes the DEA. Therefore, we cannot perform a full rank of all the units based on the DEA scores since they are based on non-unified weights.

We utilized here 4 ranking methods to rank the units.The properties of each method are given in Table 1.

The first method, DDEA, is based on the traditional discriminant analysis of two groups (see Sinuany-Stern et al., 1994). This method is based on the pregiven classification of DEA to where the common weights are found, which optimally post divide the units into two groups. there is no distrinction between inputs and outputs; thus we expect that the signs of the weights of the inputs and outputs wil be opposite. Basically this is a linear model.

The second method, DR/DEA, Discriminant Analysis of Ratios (Sinuany-Stern and Friedman, 1998) is also based on the pregiven classification of DEA. The DR/DEA scales the units by using the common weights for all the units in a ration form $T_j = \sum_{r=1}^{s} u_r y_{rj} \Big/ \sum_{i=1}^{m} v_i x_{ij}$, $j = 1,2,\dots,n$. DR/DEA calculates the best common weights using the discrimninant criterion which maximizes the ratio of the between group variance $[SS_B(T)]$ and the within-group variance$[SS_W(T)]$ of T.

The third method, CE/DEA, is based on the cross efficiency matrix h_{kj} of the DEA. We rank the units according to $\bar{h}_j = \sum_{k=1}^{n} h_{kj} \Big/ n$—the average ratio given to unit j by all other units (see Sinuany-Stern and Friedman, 1998).

The fourth method, CCA/DEA, is based on the canonical correlation analysis (Friedman and Sinuary-Stern, 1997). The canonical correlation analysis (CCA) calculates common weights, u_r, v_i, for all the units by maximizing the correlation between linear combinations of two sets. The composite input of unit j is $Z_j = \sum_{i=1}^{m} v_i x_{ij}$ and the composite output of unit j is $W_j = \sum_{r=1}^{s} u_r y_{rj}$. In order to fully rank the units, we utilize the ratio $T_j = W_j / Z_j$ for ranking.

Table 1. The properties of the Ranking Methods

Properties Rank Method	Closed form solution	Based on DEA results	Based on common weights
DDEA	yes	yes	yes
DR/DEA	no	yes	yes
CE/DEA	no	yes	no
CCA/DEA	yes	no	yes

VALIDATION

The basic model is the DEA, which provides the full efficient frontier, i.e. the pareto optimum solution. Thus we validate all the suggested ranking methods against the DEA classification. In order to test the fitness between the DEA classification and the ranking method there are two non parametric tests. The first test is the Fisher Exact Probability test (see Siegel and Castellan, 1988). The second is the Mann-Whitney rank sum test of two independent groups (efficient and inefficient).

According to the Fisher Exact Probability test one can validate if the ranks of the efficient units are all over a cut-off point, against the null hypothesis that the ranks of the efficient units are being spread randomly over and below this point. The test is performed by calculating the exact probability of a random separation of the efficient units over and below the cut-off point. If this probability is less than some value ($\alpha=0.05$), there will be a fitness between the DEA classification and any ranking method. For the Fisher Exact Probability test we need to determine, on the ranking scale, a cut-off point between the efficient and inefficient sets (see Friedman and Sinuany-Stern, 1997), which minimizes the number of misclassifications of the members of the two groups in relation to the cut-off point.

The Mann-Whitney rank sum test verifies whether the efficient units sum of ranks is significantly lower than the sum of ranks of the inefficient units. The null hypothesis states that the average sum of ranks of the efficient units represents random ranking versus the alternative that the sum of ranks of the efficient units is equal to the ideal case where they receive all the lowest ranks.

The Fisher Exact Probability test takes into account only the number of misclassification but not the ranks. Therefore the Mann-Whitney test is a higher level test, it is more accurate since it considers the ranks, not only the number of misclassified units as done by the Fisher test. Thus we can first utilize the Mann-Whitney test, and if we accept the null hypothesis (i.e. random ranking) we turn to the lower test—the Fisher exact probability. Nevertheless, looking at the classification, the first test can give some insight to the extent of the fitness between the DEA classification, and the full ranking order of classification given by the cut-off point. In each ranking method the cut-off point provides a new classification of the units into "efficient" and "inefficient" for the ranking model.

In order to compare more than two rankings, we used the Friedman nonparametric two-way analysis of variance test, where the null hypothesis states that the ranks are spread randomly, versus the alternative, which claims that the units' ranks are compatible in the various ranking methods. The statistic in the Friedman test is the sum of Ranks (R_j) of each unit j; this statistic is used for the combined ranking (CO/DEA) (see Friedman and Sinuany-Stern, 1998).

THE CASE STUDY

In this case study, 45 outpatient clinics of the largest sickfund in Israel—GSF—were studied in the central region of Israel (see Ben-Yakov and Weidenfeld, 1997). Ben-Yakov and Weidenfeld (1997) run the DEA for two separate groups: the small and big clinics in one group (16 clinics), and the rest in a second group (29); oveall, 19 of the 45 clinics were efficient. Since in DEA the units need to be similar in size, we run the model, in our study, for all the units in one group (45 units).

Four inputs were given:

X_1. Clinic built area.

X_2. Number of physician hours.

X_3. Number of nurse hours.

X_4. Number of administrative personnel hours.

Five outputs were given:

Y_1. Number of members insured.

Y_2. Number of specialists visits.

Y_3. Net number of transferees.

Y_4. Number of home care patients.

Y_5. Percentage of members who receive treatment.

After running the CCA, X_2, Y_3 and Y_5 received a negative weight in the CCA (see Friedman and Sinuany-Stern, 1997). Since X_2 and X_3 are highly correlated, and since the number of physician hours (X_3) is a leading factor, when we excluded the number of nurses hours (X_2) the new weight of X_3 became positive, therefore we excluded X_2. Thus we based our ranking on 3 inputs (X_1, X_3 and X_4) and 3 outputs (Y_1, Y_2 and Y_4) .

RESULTS

Based on 3 inputs and 3 outputs, 9 clinics were DEA efficient (see Table 2). Obviously there are differences among the various ranking methods, although each ranking was compatible with DEA (see Table 3) in all cases (p-value <0.06, in most cases p-value <0.0001). Using the non-parametric two-way analysis of the Friedman variance test, utilizing the sum of ranks, we found that there is a high compatibility among them (p-value = 0.000001), thus we can combine the sum of ranks as an overall ranking. In the combined ranking all 9 DEA efficient clinics were ranked very highly (in the first 13 places)
In summary, looking at the improvements needed in each input and output for the inefficient units, it was interesting to note that there was no need to improve (increase) the number of members in any clinic. Together with the negative weight received in the CCA of the number of transferees (Y_3), it means that although GSF receives its budget from the government on the basis of the number of members, on the clinic level it is not reflected in the short-run. Overall the number of home care patients (Y_{34}) needs 7% increase in 13 clinics, which means that there is over-capacity in those clinics, and that they can serve more home patients. The number of specialists visits (Y_2) requires a 12.4% increase. In summary, the outputs do not require big improvements, while the inputs in this specific solution of the DEA (the optimal solution in not unique) require much higher improvements. The clinic built area (X_1) requires the maximal overall improvement (about 30% cut). Indeed it is known that GSF has significantly more built area in relation to other sick funds The number of physician hours (X_2) requires an overall 26% cut, while the number of administrative personnel hours (X_4) needs a 24% cut.
Although Ben-Yakov and Weidenfeld (1997) run the model in two groups and for 4 inputs and 5 outputs, their results were significantly compatible with our results (with p-value of 0.0029 for the Mann-Whitney test, and 0.0037 for the Fisher Exact Probability test).

Acknowledgement

This paper was partially supported by the Paul Ivanier Center for Robotics and Production Management of Ben-Gurion University of the Negev.

Table 2. Summary of Rankings

UNIT	DEA	CE/DEA	CCA/DEA	DR/DEA	DDEA
20*	4	1	1	1	1
37*	22	6	7	6	3
13*	24	3	4	10	7
1*	25	8	10	3	4
3*	33	10	5	16	2
7*	34	2	3	24	5
26*	37	5	2	20	10
5*	44	4	16	15	9
23	46	13	6	14	13
43*	48	16	13	4	15
10	50	14	12	7	17
29*	56	11	14	9	22
36*	65	7	8	44	6
45*	66	12	19	17	18
35	68	22	9	26	11
2	70	9	11	42	8
18	73	15	15	29	14
16	75	18	21	11	25
30*	79	24	31	5	19
24	84	27	20	25	12
27	93	19	17	33	24

con.t

Table 2 continued

UNIT	DEA	CE/DEA	CCA/DEA	DR/DEA	DDEA
8*	95	21	34	2	38
15	100	31	25	23	21
33	100	20	32	32	16
41*	100	40	22	8	30
12	102	17	24	38	23
32	110	33	27	22	28
6	111	25	36	30	20
44*	114	39	23	12	40
4	115	34	33	19	29
14	115	23	30	35	27
40	120	42	18	28	32
11	121	32	28	27	34
42	121	28	26	41	26
34*	122	35	41	13	33
19	131	26	29	45	31
38	142	44	43	18	37
25	147	36	37	39	35
17*	148	30	38	37	43
31*	148	43	45	21	39
9	149	29	39	40	41
22	152	38	35	43	36
21	153	37	40	31	45
28	161	41	42	34	44
39	167	45	44	36	42

* These units are efficient according to the DEA model.

Table 2

Summary of Rankings

Test Ranking method	Mann-Whitney p-value	Fisher test p-value
DDEA	0.000039	0.000000
DR/DEA	0.05	0.06
CE/DEA	0.000014	0.000000
CCA/DEA	0.000063	0.001
Combined	0.000024	0.000000

References

Ben-Yakov, T. and Weidenfeld, S. "Applying DEA to clinics in Dan-Petach-Tikva region". Technical report, Mifne Mehandesim, 1997.

Barnoon, S. and Pliskin, J. "Israel" in "The International Handbook of Health-Care Systems", Edited by R.B. Saltman, 1988.

Byrnes, P. and Valdmains, V., "Analysing Technical and Allocative Efficiency of Hospitals" in "Data Envelopment Analysis: Theory, Methodology and Applications", A. Charnes, W.W. Cooper, A.Y. Lewin and L.M.Seiford (editors), Kluwer Academic Publishers, Boston (1994).

Charnes, A., Cooper, W.W. and Rhodes, E., "Measuring the Efficiency of Decision Making Units", European Journal of Operational Research, 1978, 2, 429–444.

Chilingerian, J.A. and Sherman, D.H., "DEA and Physician Report Cards: Using Assurance Regions to Benchmark Clinical Bet Practice in an HMO", Annals of Operations Research (1996).

Fare,R., Grosskopf, S., Lindgren, B. and Roos, P. "Productivity Developments in Swedish Hopsitals: A Malmquist Output Index Approach", in "Data Envelopment Analysis: Theory, Methodology and Applications", A. Charnes, W.W. Cooper, A.Y. Lewin and L.M.Seiford (editors), Kluwer Academic Publishers, Boston (1994).

Friedman, L. and Sinuany-Stern, Z. "Scaling units via the Canonical Correlation Analysis and the DEA", European Journal of Operational Rsearch, 1997, 100 (3), 629–639.

Friedman, L. and Sinuany-Stern, Z. "Combining Ranking Scales and Selecting Variables in their DEA context: The Case of Industrial Branches", Computers and Ops. Res. Vol. 25, No. 9, 781–791 (1998).

Morey, R.C., Retzlaff-Roberts, D.L and Fine, D.L. (1994) "Getting Something for Nothing: Estimating Service Level Improvements Possible in Hospitals", International Transactions in Operational Research, Vol. 1, No. 3 (1994), pp. 285–292.

Sherman, D., "Managing Productivity of Health Care Organizations" In "Measuring Efficiency: An Assessment of Data Envelopment Analysis", R.H. Silkman (Editor), New Directions for Program Evaluation, No. 32, American Evaluation Association, San Francisco, Jossey Bass, Inc., (1986), pp. 31–46.

Sinuany-Stern, Z., Mehrez, A. and Barboy, A. "Academic Departments Efficiency via DEA", Computers and Operations Research 1994, 21 (5), 543–556.

Sinuary-Stern, Z. and Friedman, L., Rank-scaling units using DEA and Discriminant Analysis of Ratios", European Journal of Operational Research, 1998.

Sinuary-Stern, Z. and Friedman, L., Rank-scaling in the DEA Context", Studies in Regional and Urban Planning, 1998 (6), 135–144.