WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

Measurement of Hospital Efficiency using Data Envelopment Analysis

H. David Sherman
Sloan School of Management
Massachusetts Institute of Technology

SSM Working Paper: #1317-82
March 1982

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

Measurement of Hospital Efficiency using Data Envelopment Analysis

H. David Sherman
Sloan School of Management
Massachusetts Institute of Technology

SSM Working Paper: #1317-82
March 1982

Draft for discussion purposes

This paper reports the results of a field study designed to evaluate the capabilities of DEA in locating relatively inefficient hospitals among a set of teaching hospitals. Earlier research has reported on the development of Data Envelopment Analysis (DEA), a mathematical programming based technique developed by A. Charnes, W. W. Cooper and E. Rhodes [6] [7] [8]. DEA is designed to locate and measure the technical efficiency of Decision Making Units (DMU) that use multiple inputs to produce multiple outputs on an expost basis. A key characteristic of DEA is its ability to locate inefficient DMU's based on observed behavior without specific information about the production function, i.e., the input-output relationships that characterize the technology of a set of Decision Making Units within an industry. All that is required to apply DEA is an understanding of and an ability to measure the relevant outputs and inputs. These characteristics are particularly suitable for application to health care organization because 1) considerable benefits may arise from any technique which will improve efficiency in this sector due to its size and widespread concern about health care costs; 2) the production function of these organizations is not well understood and is not currently specifiable in any detail (see for example [9] and [16]); and 3) these organizations are characterized by multiple outputs and inputs, e.g., a hospital uses labor, capital and supplies to produce many types of patient care and to provide other outputs such as research, community education, and nurse and intern training.

Other studies of DEA have found the methodology to be reliable in locating inefficiencies when applied to artificial data where the efficient and inefficient units are known with certainty. Sherman [14] found that DEA accurately

located the inefficient DMU's without knowledge of the underlying production function using only the outputs and inputs of each DMU as the observational data. Other field applications have relied on the theoretical soundness of DEA and its compatibility with economic theory as discussed in the works of the originators of this new methodology, Charnes, Cooper and Rhodes [6], [7] and [8]. Field studies have used DEA results tempered by sensitivity analysis to indicate to managers which DMU's were inefficient (see for example [3] and [4]). In these applications, the accuracy of the DEA results were not subject to tests of whether DEA was accurate except to the extent that in some cases managers found the results to be generally reasonable and more importantly management was willing to accept the DEA results as a base for taking actions to improve the efficiency of the DMU's individually and collectively.

In this study, the use of DEA in the health care environment is being considered because of the widespread interest in improving the performance of health care institutions and in reducing the costs of their services. At the same time, there is evidence that the incentives for improved health care organization efficiency are minimal due to the revenue or costs reimbursement systems, and there is also evidence that inefficiencies exist among a sizeable percentage of these institutions and particularly hospitals. Other techniques used to evaluate efficiency of heatlh care organizations including ratio analysis and regression analysis ahve been found to have serious limitation [14] so that DEA represents a new and possibly superior alternative which we consider in this study.

In this study, we specifically attempt to validate the results of using DEA to locate inefficient hospitals. In contrast to the DEA validity tests performed by Sherman [ ] using an artificial data set where the inefficient

DMU's were known with certainty, there are no absolute or even widely accepted indicators about which hospitals are inefficient. Indeed, if such indicators were available, there might be no need to pursue other efficiency measurement techniques like DEA. Consequently, we need to seek other approaches to validate the DEA results when applied to hospitals as a basis for assessing reliability of this methodology in health sector applications.

To validate the DEA results, we enlisted the help of a panel of nine health care experts familiar with the hospitals in the data set 1) to help define the relevant and ideal inputs and outputs, 2) to assess the implications of any compromises between ideal versus available input-output data and in the DEA evaluation of hospitals, and 3) to evaluate the reasonableness of the DEA results with respect to the hospitals identified as inefficient and the magnitude of these inefficiencies. These experts include hospital administrators, hospital regulators, and management consultants for hospitals. In addition to the opinion of these regulators, two other forms of validation were completed. First, where the experts questioned aspects of the DEA results, other types of data were considered to corroborate or refute the experts opinion and to determine whether such data might alter the DEA conclusions. A second step taken to validate certain results was to review them with managment of one of the hospitals (hospital J) identified as inefficient based on DEA to determine whether they agreed with the conclusions that their hospital was relatively inefficient.

The results of this study provide insight into the reliability of DEA for hospital application as well as observations about the way health care managers evaluate hospital operations and the adequacy of the data collected on hospital management information systems for purposes of improving efficiency of these organizations. Section 2 describes the procedures used to select the

hospital data set and the limitation inherent in the available data versus an ideal data set. Section 3 describes the results of the DEA evaluation and the reaction of the experts. Section 4 describes the reaction of management of one of the inefficient hospitals to the DEA results. In section 5 we summarize these results and consider the implication with respect to the types of data that are collected by hospitals, and the way managers and regulators view hospital efficiency. In addition, we suggest areas where further research may help to improve the capability and reliability of DEA in evaluating efficiency of health care organizations.

Section 2 - Hospital Data for DEA Efficiency Evaluation

The unit of study that was evaluated using DEA was the Medical-Surgical (MS) area of 22 teaching hospitals. For these 22 hospital MS areas, we collected 1976 data and used the following 3 inputs and 4 outputs to characterize their activities.

Outputs:
- y_1 - Patient days of care provided during 1976 for patients 65 years or older.
- y_2 - Patient days of care provided during 1976 for patients under 65 years old.
- y_3 - Number of nursing students in training in 1976.
- y_4 - Number of interns and residents in training in 1976.

Inputs:
- x_1 - Full time equivalents (FTE's) of non-physician medical surgical staff during 1976.
- x_2 - Medical surgical area supply dollars expenditures in 1976.
- x_3 - Bed days available in MS area in 1976.

Every component of the data selection process was considered in concert with a subgroup of the experts 1) to achieve the primary objective of evaluating the reliability of DEA in a health care setting, 2) to understand the implicit compromises in data selection and specification to allow for assessment of the generalizability of the results, and 3) to incorporate a reasonably true picture of the medical surgical area operations which would enable the researcher to derive some meaningful conclusions about the managerial relevance and usefulness of DEA for health care applications. The following discussion revie the key issues that were addressed in this process and the way these were resolved in the following order: selection of Medical Surgical area, selection of teaching hospitals, use of 1976 data, and selection of specific inputs and outputs.

Medical Surgical area - Rather than use the entire hospital as the object of the study, the Medical Surgical area was selected because 1) it represents what is often the largest department, i.e., it includes all patients that stay overnight in a hospital except for those with special treatment such as pediatrics, psychiatry, intensive care, etc.; 2) use of a subpart of a hospital

would increase the researcher's ability to determine if the inefficiencies located via DEA were indeed present because of the narrower scope of activity; 3) the range of outputs would be more contained so that case mix (the mix of patients treated) would already be somewhat narrower than that of the entire hospital-a key consideration since case mix output measures are not yet well developed; 4) there was a reasonable degree of uniformity in that all hospitals in the study acknowledged MS as a distinct area, i.e. separate MS data was reported to the state regulatory agency; and 5) a departmental assessment capabilit would represent a focus that is not heavily emphasized in research studies which might provide a basis for managerial insights not available in total hospital studies. A concern of certain experts was that all hospitals do not necessarily include the same activities in the MS area so that there was potential incomparabililty which would have to be addressed. This issue would not have been circumvented by use of the entire hospital, since hospitals may include or exclude certain activities which may be provided by outside organizations like anasthesiology services, laboratory services, etc. The comparability of the MS area did not prove to be one of the critical issues in this study but this represents an issue that needs to be carefully evaluated in other similar health care studies.

Teaching Hospital Data Base - Teaching hospitals which train nurses, interns, residents and medical students were selected because 1) they represent a more interesting challenge for DEA since these are distinct services provided in addition to patient care which are not easily addressed using other methodologies less sensitive to multiple output-multiple input situations, and 2) the data set should include roughly similar organizations with similar outputs and inputs so that either all teaching or all non-teaching hospitals would be selected and we chose the former. One issue we had to address was the

varying levels of teaching activities, as there is no uniform definition of teaching hospitals across all states (even though these data were all from hospitals within one state). Teaching hospitals, for this study, included those that reported training of nurses, interns or residents. This too, was of only moderate concern to the experts and did not prove to be problematic.

While there were more than 22 teaching hospitals within the state, certain of these were excluded because of greater specialization and/or lack of adequate or complete data.

1976 data - Use of older data was expected to be less threatening to administrators and regulators, as the results would describe past inefficiencies and would be silent on the efficiency of current operations. The fiscal year 1976 was specifically selected (instead of 1975 or 1977) because a special study on hospital case mix was completed based on 1976 operations for a subset of the hospitals in the state (on a voluntary basis) which would provide added insight to evaluate the validity of the DEA results.

Selection of Outputs and Inputs - Table 1 reflects the ideal outputs based on a review of the formal description of the MS area included in the state reporting instructions amplified by discussion with the experts. The general approach was to ascertain what services are provided in the MS area to define its outputs and to determine what types of inputs are required to provide these services. No output included could be increased without the expectation that some inputs would also have to be increased. Numerous compromises are reported in table 1 which are primarily a result of unavailability of data and some ambiguity in specifying the inputs and outputs of this area. The latter problem reflects a degree of insensitivity or a sense that output-input relationships are not of great concern in hispitals which may be due, in part, to reimbursement systems which have emphasized

Table 1

Input and Output Variables Identified as Relevant to MS Operations based on Review of the MS Area Description, MRSC 401 Reports, and Review by Hospital "Experts".	Variables used in DEA Analysis of 22 Teaching Hospitals (N = not used) (Y = used)	Justification for Selection and/or Non-use of the Variables
INPUTS		
Labor Inputs		
• MS Physicians	N	Hospital based physicians are highly variable and a function of hospitals' policies and their chosen organization & accounting system. This variable does not reflect the true physician inputs of the hospital. No physician input measure was used because of unavailability of a comparable measure that captures hospital and independent physician inputs.
• Non-physician Staff in MS - Full Time Equivalents (FTE's)	Y Non-physician FTE's used as input in DEA Analysis	This is the common labor input measure which includes nurses and support staff on MS floors.
• Housekeeping Dept., FTE's Input to MS Area	N	These areas have FTE's which provide services directly or indirectly to MS area FTE's and patients. No measure of the proportion of time allocated to MS is available. The cost allocations used by hospitals are based on estimates which are influenced by reimbursement incentives and which are generally not based on any time or unit charge systems.
• Nursing Service Administration, FTE's Input to MS Area	N	
• RN-LPN Nursing Education Dept., FTE's Input to MS Area	N	
• In Service Nursing Education Dept., FTE's Input to MS Area	N	
• Dietary, FTE's Input to MS Area	N	
Capital Inputs		
• Number of Beds in MS Area	Y # Bed days of care available per year = # beds x 365	Reflects the aggregate level of capital resources in terms of equipment and building space utilized to provide MS patient care.
• Floor Area in MS Area	N	Both variables are available but were not used. "Floor area" is misleading because a new plant may have less floor area with a more efficient layout, or vice versa. The age of the plant was not directly available. Depreciation was the only indicator of the relative age of the plant, and it is not a reliable measure. Some hospitals have zero depreciation expense due to fully depreciated plants purchased in some
• Age of Plant in MS area	N	

Table 1 cont.

		earlier unknown year. This would also require a method of transforming age and space to a single comparable measure of space adjusted for age which would be highly subjective.
• Fixed Assets used in MS Area – Major Movable Equipment	N	Depreciation expense is the only available measure and many hospitals have fully depreciated equipment, i.e. zero depreciation, which does not reflect the actual assets utilized in the MS area.
Other Inputs		
• Supply Dollars and Purchased Service Dollars	Y Supply and Purchased Service Dollars	Information on all units of supplies or services purchased was not available. All hospitals in the study had small amounts of purchased services which were less than 10% of the supply dollars. In the absence of a unit measure of supplies, the dollar measure was used as an approximation of the related unit volume.
• Energy Usage in MS Area	N	Excluded because this relates primarily to total plant operating efficiency. No allocation based on MS usage was available.
OUTPUTS		
• Patient Days of Care in MS Area	Y Patient Days of Care in MS Area	This is a basic output measure of medical care, unadjusted for case mix as reflected in the MS area description. Reimbursements are primarily oriented around a per diem rate.
• Admissions	N	Admissions is an alternative output measure which encompasses number of patients treated and their length of stay, in contrast to patient days of care which does not reflect length of stay. It was not used because case mix surrogates available were patient-day oriented rather than admissions oriented. Length of stay is further considered in the fieldwork analysis. Breakdown of admissions by medical vs. surgical was desirable but many hospitals do not report these separately. There is also no correspondence between severity level or resource utilization levels and the medical surgical dichotomy.
• Medical Admissions	N	(see Admissions)
• Surgical Admissions	N	(see Admissions)
• Case Mix Measures to Differentiate Hospital Patient Care Levels and Resource Utilization of the Case Load		
Diagnostic Related Groups (DRG's) reflecting the number of patients or patient days devoted to treatment of each diagnosis	N	DRG's are currently the best possibility for DEA applications. This multiple output approach could not be used in this study because of lack of availability on a complete, or even adequate basis.

Table 1 cont.

Patients or Patient Days over and under 65 years of age	Y % Medicare Days, % Patient Days with age ≥ 65	Available for all hospitals and believed to be an adequate surrogate for case mix for MS in teaching hospitals, but clearly inferior to DRG's.
• Nurse Training	Y Number of nursing students	An output of MS area where most clinical training occurs requiring resources of the MS area. To some unmeasurable extent these are inputs, which suggests that they may be treated as both inputs and outputs in some future study.
• Interns and Resident Training	Y Number of Interns and Residents in Training	Training occurs in the MS area (as well as other areas) requiring added resources of the MS area. No breakdown is available as to amount of training in MS vs. other areas. Therefore, this is treated as though the same percentage of all training occurs in the MS area for each hospital.
• Medical Students	N	Breakdown is not available and there is no systematic record of the amount of student time spent in the MS area.
• Quality of Medical Care in MS Area	N	There is no available quality rating by hospital or by hospital area. All teaching hospitals are believed to be generally of high quality, possibly varying in quality by dept.
• Quality of Training of Interns, Residents and Nurses	N	An objective quality rating for teaching programs is not available.

cost reimbursement and cost containment policies that do not reward hospitals for efficiency improvements. (See for example [ ], [ ], and [ ].)

Rather than review each of the issues raised in table 1, we will attempt to provide an overview of the conclusions that were reached when we selected the 4 outputs and 3 inputs for the DEA study. In general, while the experts would have preferred that the output data correspond more directly with the ideal data in table 1, it was agreed that the input/output data selected reasonably reflected the key MS inputs and outputs. One exception was that the output measures of patient case mix that were selected only segregated patients by age and were therefore not sensitive to other possibly more important dimensions of case mix. We chose to proceed with this widely acknowledged weakness because 1) the use of 1976 data provided us with another measure of case mix for some of the hospitals, and 2) age was one acknowledged though incomplete case mix dimension which was available (see for example [ ] and [ ]), and 3) there was no case mix data collection system in the state nor is there any one accepted and valid measure of case mix. Nevertheless, use of age as a case mix indicator is clearly incomplete and any results using this measure in this or any other study would have to be carefully evaluated before any strong conclusions would be possible.

Another dimension absent from the data is quality of outputs. While it is inconceivable that all 22 hospitals provided the same quality of care, we chose to take the same position as hospital regulators; that is, teaching hospitals were all considered to provide a minimum quality of care which is generally believed to be at or above a very high standard. None of the hospitals in the study were viewed by the experts as being substandard nor were any believed to provide distinctly higher quality care or teaching than others to the point where this required higher input consumption levels. Consequently, the lack of a quality dimension was not believed to be a serious issue with this data set.

In short, the case mix data was believed to be a key limitation. The exclusion of a quality measure and compromises in other data specifications represent areas which were not believed to have a significant impact.

Understanding the data limitations described above, we proceed to describe the results of a DEA efficiency evaluation of these 22 teaching hospital MS areas.

3. Application of DEA to Evaluate the Efficiency of the Teaching Hospitals.

The version of DEA we chose for this evaluation is a form developed by CCR [ ] which is in linear program form as follows:

Objective

$$\text{Max } h_o^* = \sum_{r=1}^{4} u_r y_{ro}$$

where o is the DMU being evaluated in the set of o = 1,...15 DMU's.

Subject to

$$0 \geq \sum_{r=1}^{3} u_r y_{rj} - \sum_{i=1}^{3} v_i x_{ij}; \quad j = 1, \ldots, 15$$

$$1 = \sum_{i=1}^{3} v_i x_{io} \tag{16}$$

$$0 < u_r, v_i; \quad \begin{array}{l} r = 1, 2, 3, 4 \\ i = 1, 2, 3 \end{array}$$

Data: Outputs: y_{rj} = observed amount of r^{th} output for j^{th} DMU

Inputs: x_{ij} = observed amount of i^{th} input for j^{th} DMU

The 4 outputs and 3 inputs for the 22 hospitals are reported in exhibit 1. The DEA results are summarized in table 2 in the columns (1) and (2) for the 22 hospitals A through V. Hospitals with an efficiency rating of 1 are designated as relatively efficient while a rating of less than 1 implies relatively inefficient. Column (2) in table 2 reflects the efficiency reference set of each inefficient hospital and represents the efficient hospital against which an inefficient hospital is being most directly compared. For example, hospital D is relatively inefficient based on DEA as indicated by an efficiency ratio of .857. Hospital D's level of inefficiency is determined by comparison with its efficiency reference set, i.e., hospitals C, D, and M. These efficiency references set hospitals will always include hospitals rated as relatively efficient e.g. Hospitals C, D and M have a rating of 1.0.

The results indicate that 10 of the 22 hospitals are relatively inefficient and that three of these have ratings below 0.9. As is discussed in CCR [ ], the ratings do not represent strict rank ordering of efficiency. Nevertheless, a hospital that has a low efficiency rating, e.g., below .9 is likely to require more substantial adjustments to input and output to become efficient than one which is inefficient but with a rating close to 1.0. This means that the lowest rated hospitals may have inefficiencies that are more apparent or easier to detect because of the magnitude of the inefficiencies that may be present.

At the point of the initial DEA application, we had essentially no knowledge or *a priori* expectations about which hospitals would be identified as inefficient. We presented these results to the experts, who had general familiarity with these hospitals, and asked them to evaluate the reasonableness of these results.

Table 2

Phase I - Initial DEA Evaluation for 22 Hospital MS Areas

Hospital	(1) DEA Efficiency Ratings	(2) Efficiency Reference Set (ERS) Basis for Measuring Inefficiency
A	1.0	
B	.970	A, C, D
C	1.0	
D	1.0	
E	.979	A, C
F	1.0	
G	.990	D, O, P
H	1.0	
I	1.0	
J	.857	C, D, M
K	.969	C, D, R
L	.995	A, O, R, U
M	1.0	
N	.887	D, M, R
O	1.0	
P	1.0	
Q	.957	A, C, R, U
R	1.0	
S	.993	C, D
T	1.0	
U	1.0	
V	.857	C, D, R

The experts' reactions were as follows:

-- Hospitals B, G, K, N, Q, S and V which were rated as inefficient seem to be a believable result.

-- No strong opinion could be elicited about hospital L.

-- Hospital E, also rated as inefficient, was believed to be more specialized and therefore less appropriate for comparison with the other 21 hospitals.

-- Hospital J was rated as inefficient but there was strong disagreement about the reasonableness of this designation. Three experts considered themselves particularly knowledgable about J. Two of these experts strongly concurred with its rating as inefficient and one strongly disagreed.

Experts' criticisms of the DEA results for hospital J and about the entire procedure emphasized two issues:

a. The efficiency reference sets selected by DEA did not reflect comparable hospitals, i.e., other hospitals in the group were believed more appropriate as a basis for comparison.

b. The case mix data used reflected only age which was believed to be incomplete and was believed to be a likely source of the misevaluation of hospital J.

We also found that these experts were not equally familiar with all hospitals in the group and that their knowledge of the hospitals did not let them clearly distinguish between the total entity and the MS area. In addition, their natural orientation was <u>not</u> focused on questions of technical (output-input) efficiency, but rather reflected a broad set of criteria of which some concept of efficiency was just one part. The composition of the group of experts was such that the problems encountered were not likely to be avoided by selection of other experts. Consequently, the first reactions of these experts could be construed as being generally in agreement with the DEA results except for hospital J, however, more rigorous evaluation and validation of the DEA results were required.

Table 3 is an example of the output format we found to be most understandable by the experts. It reflects the results of DEA in table 2 for hospital J which is compared with a composite of its efficiency reference set -- hospitals C, D and M. The input output matrix of C, D and M are multiplied by their shadow price (.286 for C, .608 for D and .146 for M) and summed to yield a composite that would yield the same or greater outputs than were achieved by J with fewer inputs than were used by J. Hence, a combination of actual hospital operations yields a more efficient hospital than J. Specifically, the composite produced more training outputs and the same patient day outputs with fewer FTE's, supply dollars and bed days available. In essence, the experts' concern was that J was not necessarily comparable to C, D and M and that age was not an adequate measure of case mix.

[illegible]. Emphasis on the three least efficient hospitals, V, J and N

We decided to concentrate our analysis on the three least efficient hospitals, Hospitals V, J and N, for a number of reasons:

- The input and output data selected required a number of compromises from what would be an "ideal" input/output data set because of the weak case mix data and unavailability of other data. Concentrating on the least efficient units would, it was hoped, minimize the extent to which the inefficiency was due to these necessary data compromises.
- The more inefficient the unit, the easier we thought it would be to locate the true source and possibly the true cause of the inefficiency. This could then provide a stronger basis to determine if the DEA results provided reliable information about hospital inefficiencies. If DEA was of no value for the highly inefficient units, this conclusion would probably extend to the other inefficient unuts. If DEA was useful only for highly inefficient units, then it might still be desirable for just that purpose in light of the alternative techniques available.

Table 3

DEA Evaluation of Hospital J versus Efficient Units with which J is Directly Compared

Hospital J - Efficiency Rating = .857

Hospital J Medical-Surgical area		vs. (.286) C	+ (.608) D	+ (.146) M
INPUTS	FTE	140.7	165.6	470.6
	Supp. $	60.34	131.3	564.8
	Bed days	69.54	65.52	165.3
OUTPUTS	Nurse Training	164	141	186
	Medicare Days	194.3	329.1	542
	Non-Medicare Days	409	257.7	995.2
	Intern/Resid. Training	28	26	105

	=		+		+		=
		40.24		100.68		69.03	
		17.25		79.83		82.85	
		19.88		38.01		24.25	
		46.90		85.73		27.3	
		55.57		200.09		79.51	
		116.97		156.7		145.99	
		8		15.8		15.40	

		Composite = (.286C+.608D+.146M)	vs.	Hospital J (1976 actual)	
INPUTS	FTE	210	<	250	excessive inputs
	Supp.	$180,000	<	$316,000	
	Bed days	82,000	<	94,400	
OUTPUTS	Nurse Trn.	160	=	160	
	Medicare Days	33,500	=	33,500	
	Non-Medicare Days	42,000	=	42,000	
	Int/Resid. Trn.	39	>	21	deficient output

Focusing on the three lowest rated hospitals, V, N and J, we embarked on a set of procedures to gain more confidence in these results and to resolve the conflict over hospital J. First we considered other case mix indicators to determine if they would indicate that the DEA results might have been biased by the use of only one case mix indicator -- age. This result is described in section A below. We used DEA to compare and evaluate a subset of 9 of the 22 hospitals that had a more specific case mix indicator for the MS area. This case mix indicator was available for only 9 of these hospitals solely for 1976 and would reflect the best available case mix measure for these hospitals described in section B below. We then addressed the comparability issue raised by the experts by using DEA to evaluate a subset of 7 of the 22 hospitals that were considered to be comparable "less intensive"teaching hospitals based on an "accepted" grouping system used by the states rate setting commission. The comparable group results is discussed in section C. Section D describes the reaction of the experts to the expanded tests in section A, B and C. In general, these expanded tests and experts' reactions supported the results of DEA though there remained some conflict about hospital J results which were further evaluated with management of hospital J as reported in the subsequent section 4 of this paper.

A. Introduction of alternative case mix measures in DEA evaluation

In response to the experts' concern about the adequacy of the case mix measure, we now consider the effect that alternate measures (suggested by one expert) might have on the original conclusions. Table 4 incorporates the data in Table 2 and includes additional indicators of these hospitals' operations as follows. Two other case mix surrogates included are: 1) the Washington index reflecting the weighted index of complexity of services offered (see [3]) and 2) the average length of stay (ALOS). In addition, average cost per day, total MS beds available and occupancy rate are included in Table 4 as factors which might influence the efficiency of MS. The issue we address is whether these other indicators of case mix, cost, etc. appear to compensate for the inefficient rating resulting from using DEA without incorporating these added variables.

Hospital N - Hospital N had an efficiency rating of .887. Consideration of the additional variables included in Table 4 tends to support the initial DEA results as follows:

a. The Washington index of N (53.2) is lower than the same index for its ERS hospitals upon which the inefficient rating is calculated, i.e. Washindex of D=64, M=57.4, and R=89.8. Thus, N offers less complex services and is therefore not likely to have a more severe case mix than the ERS hospitals.

Table 4

	Hospital	(1) DEA Efficiency Ratings	(2) Efficient Units - Basis for Measuring Inefficiency	(3) Washington Index	(4) Average Length of Stay (LOS)	(5) Average Cost Per Day	(6) No. Beds Available	(7) Occupancy Rate
1.	A	1.0		43.5	12	$34	317	.90
2.	B	.970	A, C, D	61.7	11	38	295	.87
3.	C	1.0		53.7	6	33	190	.87
4.	D	1.0		64.0	11	39	171	.94
5.	E	.979	A, C	28.2	11	48	239	.83
6.	F	1.0			4	29	36	.56
7.	G	.983	D, O, P	69.0	10	32	230	.87
8.	H	1.0		77.8	11	48	273	.91
9.	I	1.0		58.4		48	228	.92
10.	J	.857	C, D, M	66.0	11	37	258	.80
11.	K	.969	C, D, R	74.0	12	35	305	.89
12.	L	.991	C, O, P, U	77.6	10	38	365	.79
13.	M	1.0		57.4	12	40	454	.93
14.	N	.887	D, M, R	53.2	10	44	247	.83
15.	O	1.0		62.1	9	27	279	.86
16.	P	1.0		46.6		27	293	.78
17.	Q	.957	A, C, R, U	63.1	12	29	420	.85
18.	R	1.0		89.8	12	45	901	.92
19.	S	.993	C, D	78.2	10	55	330	.93
20.	T	1.0		90.5	13	49	270	.92
21.	U	1.0		56.0	10	36	325	.80
22.	V	.857	C, D, R	83.0	9	36	745	.83

b. Hospital N's average length of stay is 10 days which is shorter than the ERS hospitals' ALOS which is 11 days for D and 12 days for M and R. To the extent that ALOS is a case mix indicator, N has a less severe case mix than the ERS hospitals.

c. The cost per MS day in hospital N ($44) is higher than D ($39) and M ($40) but is somewhat lower than but very close to R ($45)

This small cost difference does not necessarily suggest that N is more cost efficient than R. If N had significantly lower costs than D, M and R, this might be an indication of cost efficiencies which compensate for DEA identified inefficiencies, but this is not the case.

Notwithstanding the compromises in the input/output data used for DEA, N continues to appear to be inefficient after considering these alternate case mix and operating variables. This of course assumes that there are no significant unknown errors in the data reported by these four hospitals. Moreover, this conclusion was further supported by the opinions of the experts as described above. Based on these results, DEA appears to have accurately identified Hospital N's MS area as relatively inefficient compared with the 22 MS areas in our data set.

Hospital V - We perform the same analysis on V as we completed above for N but find that the additional variables in Table 4 result in some questions as to whether the DEA case mix measure was adequate as suggested in the following:

a. The Washington Index for V was 83 which compared to its ERS hospitals is higher than that of C (=53.7) and D (=64) but lower than the index for R (=89.8). Thus based on the Washington Index, V may have a more severe case mix than C and D which may not be adequately reflected in the age breakdown used in the DEA evaluation. Hence, the case mix differential could compensate for the DEA inefficiency rating of .857 attributed to V.

b. The average length of stay for V was 9 days which is lower than R's ALOS (=12 days) and D's ALOS (=11 days) but higher than C's ALOS (=6 days). Again, based on ALOS, V appears to have a less severe case mix than R and D but a more severe mix than C.

c. The cost per day of V of $36 is also lower than D's cost of $39 and R's cost of $45 but higher than C's cost of $33. With respect to two hospitals D and R, V may have cost savings which compensate for technical inefficiencies.

The issue we need to address is whether N has a sufficiently more severe case mix and/or sufficiently lower costs than its ERS hospitals to compensate for the degree of inefficiency noted by DEA? There remains uncertainty about whether there are compensating

costs and case mix characteristics which might suggest initial DEA results understates V's efficiency with respect to its ERS. One expert who claimed to be well acquainted with hospital V indicated that it is clearly a less efficient hospital suggesting that the above ambiguities would not account for the inefficient rating. Nevertheless better case mix measures are needed to rerun a DEA evaluation which would be refuted less easily by hospital V. In this case, better case mix data were not available. Experts' opinion and additional insights about hospital V reported in Section C do support the conclusion reached with DEA.

Hospital J - Hospital J, with efficiency rating of .857, is the hospital with conflicting expert opinion and it was subsequently investigated in a field study described in part 4. In this analysis we find similar mixed results as were found for Hospital V above:

a. Hospital J has a Washington index of 66 which exceeds that of its three ERS hospitals C (=53.7, D (=64) and M (=57.4). Thus, J would appear to have a more severe case mix than its ERS hospitals which might have resulted in underestimating J's efficiency because of an inadequate case mix measure.

b. Length of stay for J is 11 days which is the same as C (=11 days) and D (=11 days) but is lower than M (=12 days). On this dimension, J appeared to have a case mix that is

roughly similar to its ERS case mix which would tend to support the DEA results.

c. The cost per day of J of $37 is lower than D (=$39) and M (=$40) but higher than C (=$33). Here, again, the results are inconclusive because J is not uniformly higher or lower than all the hospitals it is being compared with.

The above analysis suggests that a more refined case mix measure may be needed to resolve the uncertainty about the DEA results for hospital J which we will address in Sections C and D.

B. Application of DEA to Hospitals with Detailed Case Mix Data.

In Section B we respond to the experts' concerns about the insensitivity of the age breakdown as a case mix surrogate by using the case mix index developed using 1976 data for a subject of this hospitals in this study. A case mix severity index was developed where 3 is the higher and 1 is the lower level of severity [62]. This index and other more detailed case mix data were available only for the hospitals that voluntarily participated in the case mix study. We found that 9 of the 22 hospitals (B, C, D, I, J, M, O, S and V) had these refined case mix data. Of the original 3 highly inefficient hospitals, only J was in this set. We chose to evaluate J with respect to the other 8 hospitals to help resolve the conflict among the experts. If J is still found to be inefficient, there would be little room to qualify this conclusion based on the inadequacy of the case mix surrogate. We also were able to revise the age breakdown to specifically reflect the percentage of days for patients $\geq$ 65 years of age in the MS area (as opposed to using the entire hospital Medicare days percentage used in our other DEA evaluations reported herein). The MS percent of patient days _ 65 years of age is in Exhibit 2 column (9) and the MS case mix index is in column (10) of the same exhibit.

DEA Results

Using the same inputs and outputs we described in Section 2 except for the revised patient age breakdown for over and under 65 years old, the DEA results are presented in Table 5 for Hospital J.

Table 5

DEA Evaluations of Hospital J and A
Group of 9 Hospitals with Detailed Case Mix Data

Efficiency Rating of J	Efficiency Reference Set (ERS) of J and the factors (shadow prices) assigned by DEA to the ERS hospitals		
.849	C	D	N
	.095	.872	.121

Comparison of Case Mix Indicators and Bed Size

Hospital	J	C	D	N
MS Primary Diagnosis Case Mix Index (3 = most severe diagnosis)	2.12	2.01	2.14	2.10
Percent of patient days with age $\geq$ 65 (incorporated in DEA)	.50	.37	.56	.37
Bed Size	258	190	171	454

Hospital J is again rated as inefficient within this group and has the same ERS as in phase I using all 22 hospitals. The rating has changed slightly due to use of the more accurate MS related patient age breakdown.

This set of nine hospitals are not in the same state hospital group so there remains a comparability issue. We can, however, determine whether there is a similar case mix in these hospitals by reference to Table 5. Note that the Hospital J's case mix index (=2.12) is between Hospital D's (=2.14) and N's (=2.10) and well above Hospital C's (=2.01). Thus, one ERS Hospital, C, has a distinctly less complex case mix. The factors calculated by DEA by which the ERS hospital input/output vectors are weighted to develop the composite hospital input/output vector are listed in Table 5. These indicate the relative weight assigned by DEA to each hospital in comparing them to Hospital J. Hospital C has a lower relative importance in this comparison so that J is being compared more directly to D and N. If, for example, we weight the case mix of the ERS hospitals in proportion to these factor values, the composite case mix equals that of Hospital J ((.095)(2.01)+(.872)(2.14)+(.121)(2.10)=(2.12)).

It appears that J is being compared to hospitals which have similar case mix and that J is clearly less efficient based on the inputs and outputs we employed for the DEA evaluation.

Section C - Evaluation of a More Comparable Subset of the 22 Hospitals

We chose a subset of the 22 hospitals which would be more comparable than the entire group to determine if the DEA results would differ from phase I and whether this would resolve the issues noted about Hospital J described above. That is, more comparable hospitals are expected to have similar activities, inputs, and case mix. If we evaluate these comparable hospitals using DEA, the remaining operational differences are then likely to be due to varying degrees of efficiency because other variables will have been controlled for in the group selection process.

One set of hospital groupings, developed by the State Rate Setting Commission (RSC) were accepted as groups of "comparable" hospitals by the RSC for rate setting review purposes. Ten hospitals are included in this RSC group, which reflects the "less intensive" teaching hospitals. We noted that seven of the original 22 hospitals are in this "comparable" group which specifically includes hospitals J and V. We now compare just these 7 hospitals in a common "comparable" group using DEA to evaluate relative efficiency within this group. DEA results for this set of hospitals are believed to be less subject to criticism that these hospitals are not comparable, since these groupings have been adopted and provisionally accepted to support rate regulation,

- DEA Results

The results of DEA applied to this cross-section of comparable hospital MS areas are included in Table 6. We now consider how these results compare with the earlier results to determine whether comparability resolves the conflicts that arose in Section A. results.

Of greatest interest in Phase II results is that two of the hospitals found most inefficient J and V, are also inefficient when compared to the smaller more comparable group. The ERS for each of these hospitals has changed to include only hospitals in this group but similar results are found as indicated in Table 6.

Hospital J Phase II Results - Inefficient when Compared with Comparable Hospitals

Table 7 indicates that Hospital J, compared with a composite of its three ERS hospitals, produces the same amount or fewer outputs with more inputs. If we accept the 1B grouping as valid, then we can assert that these four hospitals are comparable and that J is technically less efficient in producing the four outputs with the specified inputs. At this point it would be reasonable to accept the DEA results that point to J as inefficient because we have already explicitly considered case mix with respect to age in DEA and the comparable groupings suggest that there is further similarity in case mix and in other operating characteristics within this group. The data in Table 6 then support the comparability of J and its ERS. For example, the size of J's and its ERS hospital's MS areas is similar (between 171 and 317 beds with J at 258 beds). J may have a more severe case mix based on the Washington index, but its average length of stay is in the same range as the ERS hospitals A, D and O. Thus, This appears to provide additional evidence that Hospital J is inefficient. We re-submit this result to the experts, as will be described in Section .

Hospital V - Phase II Results

Similar conclusions to those of Hospital J apply to Hospital V, (see table 6) except that in this instance, one expert with close knowledge of that hospital's operations indicated in Phase I that this hospital is believed to be less efficiently operated.

Hence, Hospitals J and V appear to be inefficient when compared to these more comparable hospitals.

Table 6

Comparison of "less intensive" teaching Hospitals' MS Areas.

Hospital**	DEA Efficiency Rating	Efficiency Reference Set (ERS)	MS Cost Per Patient Day	MS Cost Per Patient
A	1.0	-	$34	$408
B	1.0	-	38	418
D	1.0	-	39	429
J	0.88	A, D, O	32	407
O	1.0	-	27	243
Q	1.0	-	29	348
V	0.93	O	36	324
		Average Cost	$34.29	$368.14
		Standard Deviation	4.27	62.45

** Same letter designation as original DEA test.

Table 7.[illegible]

Comparison of Hospital J with
MRSC Group 1B Hospitals

<u>Hospital J</u> - <u>DEA Efficiency Rating = .880</u>

Hospital J Medical-Surgical area	vs.	A		D		O	
INPUTS: FTE		310		165.6		206.4	
INPUTS: Supp. $		134.6		131.3		151.2	
INPUTS: Bed days		116		65.52		102.1	
OUTPUTS: Nurse Training	(.138)	291	+ (.296)	141	+ (.498)	157	=
OUTPUTS: Medicare Days		553.1		329.1		324.8	
OUTPUTS: Non-Medicare Days		495.2		257.7		553	
OUTPUTS: Intern/Resid. Training		47		26		82	

	43		48.97		102.79
	18.55		38.82		75.3
	15.98		18.49		50.84
=	40.1	+	41.69	+	78.19
	76.2		97.31		161.75
	68.24		76.2		275.39
	6.48		7.98		40.84

		Composite = (.138A+.296D+.498O)	vs.	Hospital J (1976 actual)	
FTE		195	<	250	excessive inputs
Supp.		$133,670	<	$316,000	excessive inputs
Bed days		85,310	<	94,400	excessive inputs
Nurse Trn.	=	160	=	160	
Medicare Days		33,530	=	33,500	
Non-Medicare Days		41,983	=	42,000	
Int/Resid. Trn.		55.3	>	21	deficient output

D. Response of Experts to DEA Results in Sections A, B and C.

The results of Sections A, B, and C were presented to the experts. Their reactions to DEA results and other insights about the overall analysis are discussed in this section.

General Reactions to DEA Results.

Overall the experts strongly support the objectives and need for the new DEA type analysis. Three experts specifically noted that the application of DEA to hospitals would be useful in identifying where operations could be improved and they expressed an interest in promoting further study of this approach to efficiency evaluation of hospitals.

Generally, the experts agreed that the specific hospitals identified as highly inefficient with DEA was reasonable. The conflict over the hospital J is effectively resolved in the subsequent Part 4. This conflict was the result of a strong dissenting opinion of only one of the experts. Two other experts strongly concurred with the DEA evaluation of J. Overall, the results of the DEA analysis were believed to be accurate with respect to the three highly inefficient MS areas identified by DEA with respect to the selected inputs and outputs.

Feasibility of DEA Application

a) Data. The experts were sympathetic with data compromises needed to complete this DEA evaluation. While the data used were generally comparable, there are undoubtedly some errors which are due to hospital reporting systems or human error. In addition, one expert

provided an example of one way the DEA results might be biased. There is a nursing shortage which is affecting certain hospitals now more than in 1976. One hospital in the data set is known to be understaffed with respect to nurses. Their inputs are consequently lower than they should be based on management's judgment and this may be reflected in somewhat lower quality even though this quality differential is not objectively identifiable. This hospital would appear to be more efficient than it would be under normal fully-staffed circumstances. Understaffing could also make other hospitals appear to be inefficient with DEA if these other hospitals are compared to the understaffed hospitals. Some of these biases in the DEA results can be identified by further analyzing those results which appear incongruous with experts' expectations. For example, a data correction is located in reviewing DEA results in the subsequent Section 4. Other problems of this nature may require specific adjustments and some will be difficult to locate without extensive study. This data related problem would affect DEA results as well as results using other evaluation techniques.

Improved case mix data are likely to be available in the near future to allow for more precise hospital comparisons. The ability of DEA to incorporate case mix data in specific detail was welcomed by several of the experts.

b) Selection of Inputs and Outputs. Although there are not believed to be substantial differences in departments in hospitals, the MS boundary problems noted in Section 2 will have some impact. The DEA results would have been more readily endorsed by the experts if the potential MS differences were specifically reflected in the input and output measures used. At the same time, current reporting systems do not provide these data. If DEA was formally adopted for hospital evaluations, the related reporting requirement to provide accurate and comparable data would have to be specified. Comparisons of entire hospitals would also reduce the department boundary problem somewhat. Nevertheless, within the MS area of the 22 hospitals, no expert was aware of differences which were known or believed to be so substantial as to alter the overall results with respect to the highly inefficient hospitals.

The following analysis further supports the experts' reactions and resolves the conflict about Hospital J. Specifically, it indicates how the DEA methodology can be helpful in diagnosing inefficiencies within a hospital identified as inefficient by DEA.

Section 4. - Field Study of Hospital J - Verification of Inefficiencies Identified by DEA

The DEA results appeared to be reasonable to the experts familiar with the hospitals in the data set, the inputs and outputs employed, and the meaning of the DEA evaluation. The only exception was with respect to hospital J, where our expert strongly disagreed.

In this phase V, we address the finer question of whether the DEA information can help locate technical inefficiencies within an inefficient hospital and whether this information can be translated into managerial action to improve efficiency of hospitals.

The DEA results for hospital J were presented to the executive director and the director of finance of that hospital. Our objective was to verify the existence of the inefficiencies noted by DEA determine if these inefficiencies can be reduced via managerial action. Based on Table 7 Hospital J has inputs in excess of the composite of the hospitals in its ERS as follows:

	Composite Inputs of Hospitals A, D, and O (from Table)	Actual Hospital J Input Level	Excess Inputs of J compared to its ERS Hospitals (from Table)
Full Time Equivalents	195	250	55
Supply Dollars	$133,670	$316,000	$182,330
Bed Days	85,310	94,400	9,090

The ERS hospitals also had 34 more Intern/Resident Training outputs than Hospital J (see Table 7). The data for these four hospitals' MS areas were again verified with the rate setting commission reports to be sure that the extremely high supply dollar amounts in Hospital and other variables accurately reflected the data reported to that agency.

A ratio analysis was also completed to determine if this seemingly high inefficiency level of Hospital J found in phases I, II and III would also be evident with simple ratios and because ratio analysis

is a more familiar technique to hospital managers. The results of this analysis are reflected below in Table 8. These ratios suggest that J has the highest FTE, supply dollar and bed day inputs (lowest occupancy rate) in relation to patient day outputs without any adjustment for case mix or teaching outputs. Hence, J appears to be using an excessive amount of these three inputs per unadjusted patient day of care.

Note that the ratio analysis approach used by the rate setting commission would not have identified J as inefficient because its costs did not exceed one standard deviation above the group mean. Ratio analysis may, however, be used in tandem with DEA to the extent that it can examine individual relationships between variables when inefficiencies are believed to be present. Ratio analysis cannot be easily adjusted, however, to incorporate the multiple inputs and outputs and it is not so adjusted by the RSC or in Table 8. Based on the DEA results and with no prior knowledge that hospital J was inefficient, we have been alerted to the existence of inefficiencies and we are now able to focus on Hospital J's use of 3 inputs to produce 4 types of outputs. It is not clear that this same result could have been as directly achieved by evaluating all possible ratios.

Table 8

Comparative 1976 Medical-Surgical Area Statistics

Hospital J compared with its ERS hospitals in the comparable group of teaching hospitals

	Efficiency Reference Set Hospitals for hospital J leading to a DEA efficiency rating of .88 for hospital J in phase II			
Hospital:	J	A	D	O
FTE/Patient Day	.00331	.0029	.0028	.00235
Supp/Patient Day	$4.18	$1.28	$2.23	$1.72
Occupancy Rate	.80	.90	.94	.86
Cost Per Day	$37	$34	$39	$27
Avg. Length of Stay	11	12	11	9

The executive director and financial manager of Hospital J were given an overview of the DEA methodology and results. They had already been briefed by one of our experts about this study and its conclusions with respect to Hospital J. We explained that the data are 4 years old, that they relate only to the MS area, and that the inefficiency did not necessarily reflect managerial weaknesses. Management indicated that they were surprised at the finding because they monitor their activity closely and felt their operating ratios were at least in line with Massachusetts teaching hospitals.

Hospital J's Management's Response to DEA Evaluation

Hospital J's management's response to the DEA results are described below.

1) No substantial differences between these hospitals' MS area data were believed to exist. Although hospitals may be motivated to adjust data to maximize their reimbursements in different ways, this was not believed to be a serious issue with respect to the MS area data submitted to the RSC and used in this study.

2) Data reported by Hospital J were traced back to their internal records and were found to be accurately reflected in the input and output variables used in the DEA evaluation.

3) The MS area was believed to be similarly structured among all teaching hospitals that J was aware of. When we raised the issue of whether other tasks like housekeeping and dietary services were assigned to the MS area, it was dismissed by J's management as unlikely to account for the DEA results with respect to J and the ERS hospitals with which J was directly compared.

4) A comprehensive case mix measure was noted as being absent, although this topic was not probed in any depth.

5) Explanations for inefficiencies located in Hospital J required more in-depth analysis. Three distinct explanations were found and verified as possible explanations for the degree of inefficiency identified in Hospital J as follows.

a) Excess Supplies. The Hospital J accounting system charges supplies used on the MS floor and supplies chargeable directly to patients to the MS floor during the year. At the end of the year, the supplies chargeable directly to patients are transferred out of the MS area. We verified this transfer which amounted to $141,000 (transferred to Central Service Supplies (CSS)). Hospitals A, D and O had no such transfer on their 401 reports suggesting that this was unique to Hospital J, at least amoung these four hospitals. Thus, Hospital J's net supplies, which

never appear as such in our data, actually were $174,900 (315,900-141,000). (Revised DEA results with this adjustment are described below.)

b) Excess Hospital Beds. The occupancy rate of Hospital J was 80% in 1976 and among the lowest for teaching hospital MS areas in our study. Subsequent to 1976, Hospital J's management acknowledged the need for a bed capacity adjustment and their number of beds in MS (and hospital-wide) was reduced after 1976. Certain of these beds attributed to J in 1976 were described by management as less than full service beds which were actually available only for emergency use. Thus, J's capacity was overstated in 1976. MS area beds in J decreased from 258 in 1976 to 236 beds in 1980. This was verified by reference to subsequent years'reports. (We did not determine if other hospitals also decreased their number of MS beds.)

c) Excess MS Staff. Management of J are aware that they have a "richer" (higher) staff to patient ratio than many other hospitals. They claim this is done to maintain a higher quality of life for patients and that this is believed to lead to other benefits such as greater contributions from the community J serves. The staffing level was compared with about seven hospitals against which J was believed by its management to be comparable.

Based on these data, the director of finance estimated that J had an extra 0.15 paid hour per patient day which translates to about 5.4 extra FTE's per year in the MS area.

With this information, a revised set of DEA evaluations was completed to determine the effects of this new information on J's efficiency rating.

<u>Revised DEA evaluation of Hospital J based on field diagnosis</u>

In Table 9 we compare the adjustments that our DEA evaluation would prescribe to make J efficient compared with the adjustments suggested by Hospital J's management.

Table 9

	Excessive Inputs based on DEA evaluation of Hospital J (Table 5.8)	Input Adjustments and Excesses Noted by Hospital J Management	Cause of Excess Inputs
FTE's	55	5.4	Intentionally richer staffing .
Bed days Available	9090	6935	Reduction of 19 beds to compensate for low occupancy rate
Supply $	$182,230	141,000	Transfer to CSS area due to unique accounting system

There are several possible paths to make hospital J more efficient and it is management's prerogative to select the

most sensible path. The question we now consider is whether the path described in Column B of Table 9 would make J efficient.

The supply dollar excess in column B of Table 9 is effectively a correction of the data. Bed day excess in column B reflects an inefficiency acknowledged by management which was subsequently adjusted. The excess FTE inputs are intentional and still subject to potential review by regulators because of the potential cost savings and improvement in technical efficiency that might result from a reduction of FTE inputs.

We first considered what J's efficiency rating would be with the supply data correction and the reduced bed days available level reflecting the level of efficiency that would have existed if the beds were eliminated in 1976. The revised efficiency rating for J was 0.96 which indicates that J is still inefficient after these adjustments, although substantially more efficient than it originally appeared in the initial results in Section 2. (This assumes other hospitals did not reduce their bed size.)

We now include an adjustment to reflect the acknowledged relatively high FTE level and rerun DEA to reflect a reduction of this input by 5.4 FTEs, This is the amount of reduction needed to bring J into line with other hospitals according to J's management. This resulted in an efficiency rating of 1.0. If J reduced its personnel by 5.4, the MS area would no longer be technically inefficient in comparison with this group of hospitals. No further explanations were considered necessary as the set of adjustments noted by management

were adequate to increase the efficiency to 1 and were verifiable as reasonable in light of the hospital's circumstances. In essence, this resolves the conflict over J, as its management has pointed to areas on the DEA evaluation where they admit to excess inputs in the 1976 data.

The results of this study support the validity of DEA results in locating the highly inefficient units and the potential feasibility of using DEA as a tool to locate such inefficient units. With no a priori knowledge about which, if any, of the 22 hospitals were inefficient, and with existing groupings and ratio analysis approach pointing to hospitals other than J, V and N as highly inefficient, these hospitals were accurately identified by DEA as inefficient based on judgment of the experts. Beyond the experts input, the field study of hospital J indicated that the apparent inefficiency was real and identifiable by management of the hospital. Hospital J's inefficient DEA rating was due to both data error and managerial inefficiency, further reflecting the potential analytic capabilities of this methodology. Thus, DEA results, when properly interpreted, could identify inefficient hospitals, and also provide meaningful information to help locate technical inefficiencies within hospital units.

5. DEA in Hospital Settings - A Management and Regulatory Tool

The reactions of the experts described in this chapter were directed toward assessing the validity of DEA results. We also asked

the experts to address a related issue. If DEA is found to be reliable, is it a tool which would be of value to hospital management or regulators. Based on expert judgment, the DEA results appear to be reliable at least with respect to the more inefficient hospitals. We now review the responses of experts on the feasibility of using this technique in practice.

a) Use of DEA as a Management Tool. The most favorable response from our experts was with respect to use of DEA for comparing hospitals by their administrators. This type of analysis appears to be of great interest to hospital managers. They already conduct comparative analyses to some extent through other less comprehensive techniques like ratio analysis. DEA evaluations would allow administrators of hospitals identified as inefficient to review these results in comparison with the relatively efficient hospitals. Administrators might first consider whether data problems or inconsistencies contribute to the result. Once the data issues were resolved, the DEA results might point to the presence of technical inefficiency as occurred with Hospital J. Management could then investigate the procedures used by other more efficient hospitals to determine how these higher efficiencies are achieved and to determine if those techniques can be used to improve the inefficient hospitals' operation.

Applying DEA on a departmental basis would allow the general administrators to hold their departments accountable for any inefficiencies located in comparison to other similar departments in

other hospitals. Such inefficiencies, which are somewhat like production variances, would either have to be justified based on other factors excluded in the DEA analysis or investigated to determine how their input-output relationship can be improved to the level of the more efficient hospitals' departments.

Such management application would, of course, have to be tempered by behavioral consideration and by controlling the amount of gaming that develops around any measurement technique.

b) Use of DEA as a Regulatory Tool. Several experts believed that DEA was potentially useful for regulatory purposes. They preferred the use of DEA as a tool to help regulatory agency auditors locate inefficient hospitals for purposes of encouraging those hospitals to improve operations rather than to penalize them for their inefficiencies. DEA results can be used by regulators to educate inefficient hospitals about where they appear to be relatively inefficient.

DEA might also be useful as an audit tool for regulatory purposes as an additional technique for screening out inefficient hospitals. Here, hospitals found to be inefficient might be penalized by reducing their reimbursement rates, preventing them from applying for

certificates of need until they improve efficiency, etc.

Use of DEA as a regulatory tool raises questions about the political and legal feasibility of penalizing hospitals' inefficiencies via rate setting or other methods. Two experts, one of whom was the chief executive of a health regulatory agency, described this aspect of the regulatory process as follows. Politically, DEA would be subject to the same pressures that are exerted on any such methodology. Hospital administrators would assess the impact of using a technique like DEA. Those administrators who expect to be penalized by the results would be motivated to attempt to limit its usage or oppose its adoption. If a large number of hospitals or a small group of politically powerful hospitals expected DEA to be damaging to their institution, there would be much, and possibly successful, resistance to the methodology.

The regulatory agency would have to be convinced of the reliability of DEA in light of these hospitals' objections, and would then negotiate the way it was to be used with those opposing its adoption.

Once such a technique is adopted, the legal feasibility issues which may arise will generally relate more to the legislative powers of the agency than to the methodology itself. Courts have overruled health regulatory agencies where their actions were interpreted to exceed their statutory authority. If this were a limiting factor in the use of DEA, it would relate only to the anticipated use of DEA under the agency charter rather than to the

question of whether it is an appropriate methodology for that purpose. If the DEA approach were adopted by such an agency after they studied it and found it to be reliable, the burden of proof as to the methodological weaknesses would rest with the regulated parties, i.e., the hospitals. Generally, the regulatory agency is viewed by the court as the expert with a staff capable of assessing the regulatory methodologies. Consequently the court will tend to support the agency unless the opposing side can produce convincing evidence to discredit the methodology. The test generally applied is whether the methodology is good, and whether there is a better approach. One expert who specializes in issues of health law was aware of no cases where the courts found a regulatory agency's selection of a methodology to be unsupportable, and thus illegal.

As the quality of data improves, the inputs and outputs will more closely reflect the true activity levels of hospitals. This should allow for much more rigorous testing of DEA. If expanded tests further support the findings of this study, this proven reliability of DEA would decrease the likelihood of successful political and legal opposition to its use in a regulatory setting.

Qualification - Weaknesses in the "Expert Benchmark"

DEA was applied to a set of teaching hospitals' medical-surgical areas. This application required several compromises in data specification from an "ideal" data base. The DEA results appeared reasonable to our set of experts knowledgeable in many aspects of hospital management. On the other hand, from experience (or otherwise) the experts tended to rely on ratio analyses and past experience without adequate appreciation (we think)

of the possibilities that DEA offers for handling many more variables than these (or other) hospital experts are accustomed to considering.

Evidence from this quarter must also be qualified because each expert judgment is more likely to be qualitative and subjective than quantitatively oriented and objectively verifiable. Furthermore, there are other limitations such as:

- The experts could not compare all of the hospitals within a large sample in a relatively even manner but tend rather to have detailed knowledge based on experience with only a subset of these hospitals. For example, no expert consulted was willing even to rank all the teaching hospitals in our study in part because of the multiple dimensions involved and possibly for other reasons as well.
- Expert opinions also tend to reflect a mixture of considerations such as hospital prestige, effectiveness, economic efficiency, staff quality, teaching quality, etc., and it is not likely that such experts could segregate issues of technical cy from other issues, as is done via DEA. Alternatively this may be regarded as a strength in the use of such experts to guard against deficiencies of DEA in its focus on issues of efficiency while ignoring other performance dimensions.
- Experts may not be aware of specific hospital department efficiency levels and may therefore project their overall hospital evaluations to the department being evaluated even when asked specifically to assess only one department--the medical-surgical area.

- Experts may also confuse different types of efficiency including scale and allocative efficiencies rather than confining themselves only to technical efficiency in assessing hospitals. This may be further reinforced because of the heavy emphasis placed on hospital cost containment in recent state and federal legislation which emphasize issues of "best" price or need for a service (effectiveness) very prominently while leaving other issues in the background to some extent. Consequently, experts may need to be trained or oriented to independently assess technical efficiency--the object of this study--and then to go on to other types of efficiency, too, such and allocative efficiencies.

- Current hospital incentive systems do not reward administrators for improving their hospital's technical efficiency even though this is one avenue to cost reduction. (This condition arises because the hospital rate setting systems have generally reimbursed hospitals for their costs, regardless of whether they reflect efficient or inefficient costs). Consequently, these administrators may not be adequately trained or oriented toward any types of efficiency evaluation.

Qualification - Completeness of DEA in Locating Inefficient Hospitals

The field evaluation of one of the hospitals with a low DEA efficiency rating verified that there were distinct areas of inefficiency as well as data problems which were accurately reflected in the DEA results. Furthermore, this hospital would not have been

identified as inefficient by the existing evaluation mechanisms used in Massachusetts. Nevertheless, all of the hospitals in the study could theoretically be technically inefficient. Even if DEA were operating perfectly in its identifications, it can only locate <u>relative</u> inefficiencies when applied to data sets, as we used them, without external knowledge about absolutely attainable levels of efficiency. Thus the efficient units are not necessarily technically efficient in an absolute sense.

Verification of such efficiency might be undertaken, of course, but with our limited resources we concentrated on the least efficient hospitals. Thus, even though DEA was found to be accurate in these dimensions, it remains to be seen whether the efficient units were also correctly identified by site visits and studies instead of relying only on the concurrence of a body of experts, as we did.

<u>Strengths and Limitations of DEA in Use as a Management Tool in the Health Care Sector</u>

We now consider the use of DEA as a management tool for the health care industry and discuss the relative strengths and limitations of such applications. There are at least two potential user groups who may benefit from use of DEA in the health care sector: (1) health care managers within hospitals, and (2) regulators of hospitals. The regulators may complete hospital efficiency assessments either to determine which hospitals are inefficient or to determine which hospitals require additional management services to improve their operations. They may then need further aid to determine how best to penalize the inefficient institutions and to provide incentives

for improving operations and reallocating resources. Beyond this, a regulator might use DEA to cull out inefficient hospitals from the rate setting base, thereby establishing reimbursement rates based only on the more efficient hospitals.

Health care managers, such as hospital administrators might also utilize DEA to evaluate the technical efficiency of their hospital and its departments compared with other hospitals.

While DEA appears most directly useful as a tool to identify relatively inefficient hospitals within a set of hospitals, these user groups may find the additional analytic capabilities provided by DEA to be of value as well. After a hospital or hospital department is identified as relatively inefficient, DEA may be used to perform various sensitivity and/or "what if" analyses. While the variety and number of possible sensitivity analyses may be endless, depending on the imagination of the user, we note a few examples to illustrate the types of issues that might be analyzed via DEA as follows:

- What changes in operations would make the inefficient unit efficient? Here management might propose various possible ways of reducing inputs or increasing outputs. Each of these alternatives might be compared with other hospitals to determine which of these alternatives will make the hospital or department technically efficient.

- How sensitive is the hospital's efficiency to alternative input/output specifications? Where alternative sets of input and output measure are used, DEA can be reapplied to determine if the conclusions about one hospital's inefficiency is heavily influenced by the choice of input and output measures. For example, the use of patient versus patient day as an output measure may lead to different conclusions about a hospital's efficiency. Unless one of these is determined to be the "proper" measure, both sets of results may have to be considered. This capability is particularly important for hospitals, since there are currently numerous but no authoritative taxonomies of outputs and inputs and ways to measure them.

• Are inefficiencies due to current management practice or to characteristics of a hospital that cannot be varied at the discretion of management? Certain inputs, like the age and quality of the plant, may not be controllable over the short run but may contribute to overall technical inefficiency as calculated by DEA. In such situations, a DEA evaluation can be redone with inputs altered to reflect these nondiscretionary inputs at an efficient level. See [23] The resulting analysis would suggest whether such a hospital was still inefficient after nondiscretionary input are effectively filtered out of the efficiency assessment.

Use of DEA in Concert with Other Evaluation Tools

As we have already noted, DEA assessments will help identify inefficient units and the potential location of the inefficiency, e.g. which department is relatively inefficient and which inputs or outputs appears to contribute to the inefficiency. The specific location and cause of the inefficiency and the remedy to improve efficiency would require use of other Operation Research techniques or Industrial Engineering methodologies. In addition, other techniques, like ratio analysis may help locate operating relationships which relate to other inefficiencies. For example, the analysis in section 4 of the hospitals in section 3-C indicated that one hospital appeared to be cost inefficient by virtue of an excessive cost per day ration, a statistic which is of key interest to regulators. At the same time, DEA found other hospitals in the group to be technically inefficient, a result that was not available from the ratio analysis but which could also lead to the desired reduction of cost of hospital care.

DEA is a new methodology which potentially can strengthen the ability of a manager to assess efficiency and it appears to be a technique that could be added to the analytic tools now used by managers. Once a

data base is developed to capture the relevant inputs and outputs for a DEA analysis, the incremental cost of running a DEA analysis is comparable to that of running any linear programming package, and should therefore be quite modest. Hence, we do not believe the use of DEA will be reduced to a question of which analytic resources the institution can afford. Rather, it is likely to be a question of whether DEA should also be used, and in which ways.

LINEAR:
16-JUL-80 13:48 .CMP

Exhibit 5.5
Example of DEA LP tableau
Evaluation of Hospital C Compared to the 22 hospitals in Phase I

OBJECTIVES	MSFTE	SUPPS	BEDDAYS	NURSETRN	MEDCARDAY	PATDAY	INTRESID		
		000's	000's		00's.	00's			
HO	0	0	0	164	194.3	409	28		ε =.0001
Hospital CONSTRAINTS	MSFTE	SUPPS	BEDDAYS	NURSETRN	MEDCARDAY	PATDAY	INTRESID	RELATION	IC-4
Denom.	140.7	60.34	69.54	0	0	0	0	EQ	1
A	-310	-134.6	-116	291	553.1	495.2	47	LE	0
B	-278.5	-114.3	-105.8	155	376.4	556.3	3	LE	0
C	-140.7	-60.34	-69.54	164	194.3	409	28	LE	0
D	-165.6	-131.3	-62.52	141	329.1	257.7	26	LE	0
E	-326.8	-41.89	-87.47	209	267.7	461.8	13	LE	0
F	-24.4	-11.74	-13.14	0	35	35	19	LE	0
G	-178.5	-171.7	-64.18	0	307.6	428.3	14	LE	0
H	-391.9	-575.5	-99.37	0	350.2	558	153	LE	0
I	-323.9	-59.43	-62.3	0	318.8	434.9	29	LE	0
J	-219.9	-315.9	-94.43	160	335.3	419.9	21	LE	0
K	-284	-292.5	-111	0	353.5	638.6	0	LE	0
L	-233.7	-291.7	-96.99	154	322.9	437	85	LE	0
M	-170.6	-584.8	-105.3	186	512	995.2	105	LE	0
N	-219.2	-251.3	-93.4	92	323.5	436.6	16	LE	0
O	-266.4	-151.2	-102.1	157	324.8	553	82	LE	0
P	-218	-.154	-107.2	127	470.6	366.7	100	LE	0
Q	-354.5	-217	-153.7	285	437.8	819.2	92	LE	0
R	-1052	-1926	-329.8	316	991.8	2047	434	LE	0
S	-352.9	-249.1	-112.7	0	367.9	663.1	0	LE	0
T	-306	-948.1	-100.2	0	198.9	723.3	187	LE	0
U	-803	-80.46	-118.6	110	749.3	699.6	266	LE	0
V	-530.4	-770.8	-215	144	584.1	1197	89	LE	0
V1 (VRT)	1	0	0	0	0	0	0	GE	1E-4
V4 (VRT)	0	1	0	0	0	0	0	GE	1E-4
V6 (VRT)	0	0	1	0	0	0	0	GE	1E-4
U2 (VRT)	0	0	0	1	0	0	0	GE	1E-4
U5 (VRT)	0	0	0	0	1	0	0	GE	1E-4
U6 (VRT)	0	0	0	0	0	1	0	GE	1E-4
U1 (VRT)	0	0	0	0	0	0	1	GE	1E-4

Exhibit 1

Data on Hospitals for DEA and Related Analyses (n - not available)

Hospital	Inputs to MS Area			Outputs of MS Area			Case Mix Related Surrogates and Indices					Other Descriptive Variables		
	No. non-physician FTE's	Supply $'s and Purchased Services	Bed days Available	Patient Days of Care Provided	Nurses in Training	Interns and Residents in Training	% Medicare Days in Entire Hospital	% Patient Days with Age ≥ 65 in MS	Weighted MRSC Case Mix Index for MS	Washington Index	Average Length of Stay in MS	Number of Beds in MS	Occupancy Rate in MS	Average MS Direct Cost Per Day ($'s)
A	310	134,543	116,022	104,827	221	47	52.76	n	n	43.5	12	317	.90	34
B	278.6	114,280	106,832	93,272	135	3	40.36	50	2.03	61.7	11	295	.87	38
C	140.75	60,340	69,540	60,330	164	23	32.21	37	2.01	53.7	6	190	.87	33
D	145.6	131,349	62,516	58,683	141	26	56.05	56	2.14	64.0	11	171	.94	39
E	326.8	61,992	87,474	72,949	209	13	36.70	n	n	28.2	11	239	.83	43
F	24.4	14,735	13,140	7,317	-	19	50	n	n	N/A	4	36	.56	36
G	179.5	171,682	84,180	73,590	-	14	41.80	n	n	69.0	10	230	.87	32
H	334.9	675,492	99,372	90,820	-	163	38.56	n	n	77.3	11	273	.91	45
I	323.9	59,832	82,300	75,371	-	29	42.30	44	2.08	58.4	n	228	.92	45
J	249.9	315,851	94,428	75,522	160	21	44.40	50	2.12	66.0	11	258	.80	37
K	234	207,481	191,020	99,212	-	-	35.63	n	n	74.0	12	305	.59	35
L	233.2	291,703	96,990	76,098	134	86	43.02	n	n	77.6	10	265	.79	33
M	470.6	564,834	165,256	153,717	186	105	35.26	37	2.10	57.4	12	454	.93	40
N	259.2	234,261	90,040	75,413	92	16	42.90	n	n	53.2	10	247	.83	44
O	206.4	151,175	102,114	87,776	157	82	37.0	44	1.97	62.1	9	279	.86	27
P	218	**	107,238	83,728	127	100	56.20	n	n	46.6	n	293	.73	27
Q	344.6	236,994	153,720	130,696	285	92	37.32	n	n	63.1	12	420	.85	29
R	1052	1,026,003	329,766	303,728	316	434	32.62	n	n	89.8	12	901	.92	45
S	342.9	249,094	112,728	105,098	-	-	36.91	34	2.13	78.2	10	330	.93	55
T	356	948,094	100,170	92,221	-	187	21.57	n	n	90.5	13	270	.92	49
U	808	80,457	113,625	94,890	110	266	26.27	34	2.24	56.0	10	365	.80	35
V	530.4	770,757	214,959	178,087	144	89	32.80	n	n	83.0	9	745	.83	36

** Data available was found to be inaccurate and corrected data was not available. This unit is effectively excluded from the study in as much as this hospital is not in the efficiency reference set (ERS) of other hospitals and therefore does not affect their efficiency ratings.

References

[1] Banker, R., A. Charnes, W. W. Cooper, A. Schinnar, "A Bi-Extremal Principle for Frontier Estimation and Efficiency Evaluation," Management Science (forthcoming).

[2] Bentley, J. D. and P. W. Butler, Describing and Paying Hospitals - Developments in Patient Case Mix (Washington, D.C.: Association of American Medical Colleges, 1980).

[3] Bessent, A. and W. Bessent, "Determining the Comparative Efficiency of Schools Through Data Envelopment Analysis," Research Report CCS 361 (Austin, Texas: Center for Cybernetic Studies, University of Texas at Austin, December 1979).

[4] Bessent, A., W. Bessent, J. Kennington, B. Regan, "An Application of Mathematical Programming to Asses Productivity in the Houston Independent School District," Technical Report OR 80011 (Dallas, Texas: Southern Methodist University, June 1980).

[5] Charnes, A., W. W. Cooper, A. Lewin, R. Morey and J. Rousseau, "Efficiency Analysis with Non-discretionary Resources," Research Report 379 (Austin, Texas: Center for Cybernetic Research, University of Texas at Austin, September 1980).

[6] Charnes, A., W. W. Cooper, and E. Rhodes, "Measuring Efficiency of Decision Making Units," European Journal of Operational Research, Vol. 2, No. 6, November 1978, pp. 429-444.

[7] Charnes, A., W. W. Cooper, and E. Rhodes, "Short Communication: Measuring Efficiency of Decision Making Units," European Journal of Operations Research, Vol. 3, No. 4, July 1979, p. 339.

[8] Charnes, A., W. W. Cooper, and E. Rhodes, "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through," Management Science, Vol. 27, No. 6, June 1981, pp. 668-697.

[9] Feldstein, M., Economic Analysis for Health Service Efficiency (Amsterdam, Holland: North-Holland, 1968).

[10] Hamilton, D. and G. Jensen, Health Care Cost Variation: Departmental Analysis Incorporating Case Mix (A Study of Massachusetts Hospitals) Final Report, October 1978, Report AAI No. 78-114 (Cambridge, Massachusetts: Abt Associates, 1978).

[11] Massachusetts Rate Setting Commission, "Executive Summary - FY 1981 Groups of Hospitals," (internal unpublished document of MRSC, 1980).

[12] Massachusetts Rate Setting Commission, Hospital Bureau, "Grouping of Hospitals for Purposes of Fetermining Reasonable Cost Levels," Document No. 114.1 CMR, Rate Setting Commission, 1980.

[13] Massachusetts Rate Setting Commission, Commonwealth of Massachusetts Hospital Uniform Reporting Manual (Comm. of Massachusetts Publication: #10865-202-130-9-78-C.R., 1978).

[14] Sherman, H.D., Measurement of Hospital Technical Efficiency: A Comparative evaluation of Data Envelopment Analysis and Locating Inefficiency in Health Care Organizations," unpublished Doctoral Dissertation Harvard Graduate School of Business Administration, 19

[15] Trivedi, V. M., "Classification of Short-Term General Hospitals for Control and Equity," Inquiry, Vol. 15, No. 3, September 1978, pp. 255-264.

[16] Weiner, S. M., "External Incentives for Institutional Productivity - The Role of Planning, Regulation, and Private Initiatives in the Hospital Setting," Paper presented to the Conference on Incentives for Improving Productivity in Health Care, The National Council on Health Planning and Development, January 31-February 1, 1980.

Date Due
