A Method for Identifying Cost-Efficient Practices in the Treatment of Thoracic Empyema

Emmanuel Thanassoulis[$] Aston University, Aston Business School, Birmingham B4 7ET,

Alexander Takamura-Tweedy Aston University, Aston Business School, Birmingham B4 7ET,

Mitul Patel, Birmingham Women's and Children's NHS Foundation Trust, Department of Microbiology, Steelhouse Lane, Birmingham B4 6NH and

Satish Rao, Birmingham Women's and Children's NHS Foundation Trust, Department of Respiratory Medicine Steelhouse Lane, Birmingham B4 6NH

ABSTRACT

Background

There is a need to identify cost-efficient practices in delivering health care.

Aims

To illustrate how cost-efficient practices can be identified and disseminated in treating thoracic empyema.

Methods

Data Envelopment Analysis (DEA) was used to identify the scope for reducing Length of Stay (LOS), and therefore costs, at inpatient spell level for thoracic empyema.

Findings

LOS reduction potential was identified representing about 50% of the recorded LOS. Significant differences in potential for LOS reduction were also found between consultant teams and between sources of patient admission.

Conclusions

DEA enables multiple conditioning factors affecting costs at inpatient spell level to be taken into account simultaneously. It enables management to identify and disseminate best practice within its own hospital. The approach can be transferred to other inpatient settings beyond thoracic empyema.

Acknowledgements

The authors are grateful to Amanda Martin for retrieving the data and anonymising it.

$ **Corresponding author**, Emmanuel Thanassoulis, Department of Operations and Information Management, Aston Business School, Aston University , Birmingham B4 7ET, e.thanassoulis@aston.ac.uk, Tel +44 (0)121 204 3254

Keywords: Thoracic Empyema, Cost efficiency, Health Care, DEA

Keypoints

- Identifying cost efficient practices in treating thoracic empyema
- Assessing cost efficiency at inpatient spell level
- Controlling for patient condition in predicting Length of Stay for thoracic empyema
- Data Envelopment Analysis as an efficiency assessment method
- Identifying benchmarks for cost efficiency in treating thoracic empyema

Reflective questions

1. How can we control for co-morbidities presenting on patient admission so as to predict the expected cost of treatment during an inpatient spell?
2. How can we identify cost-efficient practices in delivering an inpatient spell for a given primary diagnosis?
3. How can we identify benchmark inpatient spells from which best practice can be drawn and disseminated?

1. INTRODUCTION

The healthcare industry is one of the largest in the world, using nearly 10% of global GDP. Due to an ageing population and increased prevalence of chronic diseases, costs and patient numbers in the healthcare sector continue to grow. The need for health care providers to operate more efficiently and optimize productivity is therefore very important. This paper reports on an approach to improving the cost-efficiency of inpatient health care delivery through a systematic identification and dissemination of best practice. The approach is demonstrated using data on inpatient spells for thoracic empyema or Pyothorax (henceforth for ease of reference **empyema**) at a major UK Children's hospital.

For clarity, an inpatient spell encompases the "continuous period of time spent as a patient within a hospital. This is as opposed to an episode, which is the "time spent under one consultant". A spell therefore could potentially contain multiple episodes. The study focuses on spells rather than episodes so as to encompass the whole time they were under the supervision of the hospital.

The approach developed will allow a hospital to assess retrospectively performance in terms of cost at the granular level of **inpatient spells**. Our approach allows the determination of a summary figure of cost efficiency, expressed here in the form of potential reduction in the Length of Stay (**LOS**) of the spell. Further investigations will then identify how factors such as admittance location, time to first operation, treatment consultant etc. affect cost efficiency. This will enable hospitals to identify and roll out operating practices conducive to improving cost efficiency while eliminating practices inimical to the efficient use of resources.

2. METHOD

We illustrate the use **Data Envelopment Analysis** (**DEA**) for the purpose at hand. We use DEA to identify benchmark inpatient spells in terms of shortest LOS, controlling for patient condition on admission. The benchmarks are in turn used to estimate the potential for LOS reduction in other spells. DEA is by now a well established method for comparative efficiency assessments. Its key advantages are that it can take into account multiple incommensurate outcome measures and set them against multiple incommensurate resource and contextual factors to derive measures of efficiency and productivity which can be used to manage performance.

The development of DEA was initiated by (Charnes et al. 1978), based on initial ideas by (Farrell 1957). Since then DEA has become the method of choice for efficiency and productivity analysis especially so in complex contexts where multiple incommensurate inputs are used to secure multiple incommensurate outputs. It is especially useful in cases where no input or output market prices exist, yet efficient resource use is important (e.g. health, education, justice, policing etc.). In a survey of DEA applications covering the period 1978 to 2010 (Liu et al. 2013) list over 3100 papers where some real life application was embedded. The five top areas in terms of application were banking, agriculture, education, health care and transportation. An outline of DEA can be found in the Appendix. For a fuller coverage of DEA as a method see (Thanassoulis 2001).

3. IMPLEMENTATION OF DEA ON THORACIC EMPYEMA DATA

3.1 Conceptualisation of the Setting for the Analysis

The assessment of cost efficiency or productivity of health service delivery has been criticized (e.g. (Hollingsworth, 2008) because at the aggregate level that it is normally conducted one cannot reflect accurately the multiplicity of differences between the individual patients which in turn bear on the costs of treatment. The best predictor of a patient's treatment costs is the medical diagnosis for which the particular inpatient admission has been made. The severity or the stage of the disease, the general health condition of the patient along with his/her gender and age can also impact the cost of treatment even after we control for primary medical diagnosis for the admission.

With the foregoing in mind our unit of analysis was set at **inpatient spell level** when one of the first 3 diagnoses was either **Pyothorax with fistula** or **Pyothorax without fistula**. It is noted that the hospital concerned used Pyothorax for thoracic empyema. Narrowing in this way the scope of the analysis to patient level and specific diagnoses enables us to control as far as possible for the exact medical condition or conditions which impact the costs of treatment during an inpatient spell. Then any variation in costs between inpatient spells would be attributable to clinical judgment and other treatment features, the efficacy of which we aim to gauge. At the time of writing reliable data on the true cost of treatment during an inpatient spell was not available to us. It is for this reason that we have opted to use **length of stay (LOS)** of the inpatient spell as a surrogate for the cost incurred during the spell, (for a similar approach see (Thanassoulis et al, 20016). This is under the implicit assumption that clinical outcomes have not been jeopardised through inappropriately short LOS. We return to this point later, after the initial assessment of cost efficiency.

Thus the resulting setting for the analysis is one where we wish to ascertain the extent to which the LOS of inpatient spells for empyema could have been shorter, if at all, controlling for the condition on admission of an empyema patient. The key question is how to reflect the condition of the inpatient on admission. This question was addressed by a combination of clinical judgment and statistical tests.

Clinical judgement, constrained by the availability of data, led to the view that the following impact variation in LOS of inpatient spells for empyema treatment:

- **"Exams" –** This is the number of radiological investigations conducted within the inpatient spell. Exams are taken as a surrogate marker for the health condition presented by the patient on admission.

- **“Proc” –** This is the number of procedures conducted. It is used as a surrogate for the severity of the empyema condition. We have excluded from “Proc” computed tomography or radiology as they are recorded within radiology exams above.
- **“Diag” -** We have used the number of separate diagnoses as comorbidities, affecting LOS;
- **Age** of the child (categorised into babies <2 years, toddlers 2-4, child 4-10, teen over 11);
- **Gender** of the patient.

A regression analysis using data on 358 spells on the foregoing variables concurs with the clinical judgement that exams and number of diagnoses do impact LOS. Number of procedures was not significant in a strict statistical sense. However clinical judgement was to retain procedures as a potential explanatory variable for LOS. Gender was not found to be statistically significant in explaining LOS. Taking baby as the base case, age is not statistically significant either except for ‘teen’ which arguably has explanatory power for LOS. However, given the relatively low number of teen cases, clinical judgement was to assess any impact of teen age post the basic initial assessment of LOS congruity with patient condition.

Within the framework of a DEA model we set observed LOS of each spell against the number of EXAMS, PROCEDURES and DIAGNOSES pertaining to the spell. The model was then solved to determine a corresponding virtual or real benchmark inpatient spell for each real inpatient spell. Using benchmarks in this way we estimate the shortest LOS feasible for each real inpatient spell, when we control for patient condition on admission. The benchmark inpatient spells are also used ex post to compare case notes of spells to identify best practice conducive to shortening LOS and therefore to improve cost efficiency. The model solved can be found in the Appendix. It was solved using commercial DEA software (www.deasoftware.co.uk).

3.2 Data

Data on empyema inpatient spells was provided by a Children’s hospital. The data covered the period 2007 – 2016. Spells where the length of stay was 0 or 1 days were removed as relating to outpatients. Spells where the Empyema diagnosis did not fall within their first 3 diagnoses were removed as Empyema would likely not have been a main reason for their spell in hospital. Finally, spells where the patient died were also removed, as their LOS was right-censored while also spells with LOS> 1.96 standard deviations from the mean were removed as outliers. This process left 358 spells, with mean 11.12, standard deviation of 5.78 and median of 9 days. The bulk of the spells last between 8 and 10 days. However, there is also a strong positive skew with some spells lasting over a month even after excluding 8 outlier spells that lasted even longer than that.

Figure 1 shows the numbers of exams, procures and diagnoses per spell. Some 75% of spells have 7 or fewer exams and this number drops to 5 in the case of procedures and diagnoses. However, the distribution is skewed and a minority of cases have as many as 34, 17 and 14 Exams, Procedures and diagnoses respectively. Thus a significant number of spells (top quartile) differ substantially from the rest of the spells on patient condition on admission.

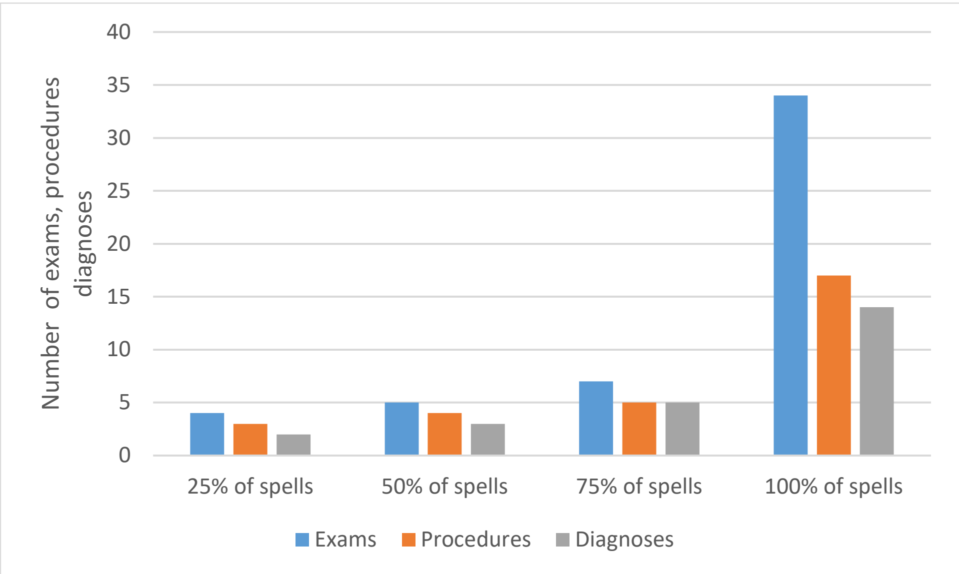

Figure 1: Distribution of variables relating to patient condition on admission

4. RESULTS

Figure 2 shows the estimated scope for LOS reductions estimated.

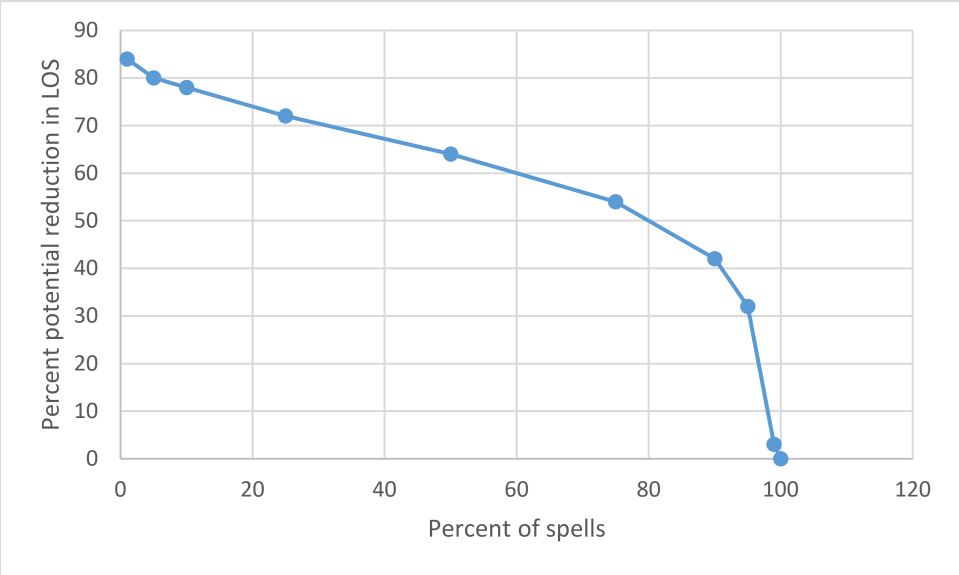

Figure 2: Scope for LOS reductions.

Some 50% of spells can shave off at least 64% of their spell duration. The quarter of least 'efficient' spells can reduce their durations by at least 72% and three quarters of spells can reduce their LOS by at least 54% of their original LOS. On the whole, the results taken at face value suggest there is a remarkably large scope for the reduction of LOS. Aggregating potential savings in terms of LOS days across all 358 spells it is found that a total of 1952 days could have been saved which is just under 50% of the aggregate observed LOS days. Clearly the findings cannot be taken at face value as there may be factors affecting patient condition that it has not been possible to capture appropriately in the model, including the specific impact each type of diagnosis, exam or preocedure may have had on the duration of an in-patient spell. Nevertheless the results do give a preliminary indication that there is a great deal of potentially unnecessary stay in hospital that is worth investigating further with a view to eliminating it.

One of the strengths of the DEA method used is that it yields the potential savings in LOS at spell level, identifying at the same time the comparator spells on which the estimated savings are based.

This makes it possible for hospital management to reassess the findings at spell level for any shortcomings in the model. Where, however, no explanation in clinical terms can be found for LOS in excess of the shortest estimated feasible, lessons can be learned about treatment protocols conducive to LOS savings without detriment to clinical outcomes. The reassessment of findings at 'inefficient' spell level againts its 'efficient' peers (benchmarks) can be demonstrated using an example.

Spell 648 is an example of an 'inefficient' spell. It had LOS of 19 days which the DEA model estimated could be reduced to 5 days. This is a drastic potential saving and so it calls for a further clinical review of the case. The model estimated the potential reduction to a LOS of 5 days with reference to two 'efficient peer' benchmark spells: 581 and 499.

Table 1 compares the data of spell 648 with those of its benchmark efficient spells.

Table 1: Comparing inefficient spell 648 with benchmark spells 581 and 499

	LOS (Days)	EXAMS	PROCEDURES	DIAGNOSES
Inefficient spell 648	19	11	- *Unspecified opening of chest, - Insertion of tube drain into pleural cavity and* - *second Insertion of tube drain into pleural cavity*	*'Pyothorax without fistula'*
Benchmark spell 581	6	17	- *Insertion of tube drain into pleural cavity,* - *Insertion of central venous catheter NEC,* - *Approach to organ under ultrasonic control,* - *Invasive ventilation and* - *Non-invasive ventilation NEC*	*pyothorax without fistula plus another 6 diagnoses*
Bench mark spell 499	4	3	- *Thoracoscopic approach to thoracic cavity* - *NEC and Right sided operation*	*pyothorax without fistula plus one more diagnosis*

As can be seen from the data in Table 1, on the face of it, the benchmark spells 581 and 499 relate to patients who were more ill than the patient relating to spell 648. Yet LOS at the benchmarks was far shorter than it was in the case of spell 648. So the analysis does indicate the need for a clinical review of the case notes to see how the significant differences in expected LOS at the 3 spells on Table 1 have occurred. Clearly in reality this cross examination of 'inefficient' and corresponding 'efficient peer' spells needs to be extended to a significant number of cases to draw reliable conclusions for treatment protocols conducive to shorter LOS.

5. DISCUSSION

Our aim in this paper is to illustrate how DEA can be a useful tool in identifying operating practices and treatment protocols conducive to shorter LOS and ultimately to cost savings. One aspect which facilitates this process is that DEA assessments are at spell level. This enables us to cluster the findings by a number of categorical variables such as primary diagnosis, origin of admission, treatment team etc in order to investigate their impact on LOS and thereby any lessons that could be learned for improving cost efficiency. We illustrate the approach here using the following categorical variables for clustering the findings.

5.1 Primary diagnosis

We begin by clustering the efficiencies (i.e. fractions to which observed LOS levels can be reduced) by the patient's primary diagnosis, '*Pyothorax without fistula' (J869*), '*Pyothorax with Fistula' (J860*) or '*Other*'. This leads to subsets of 266 'J869' spells, 30 'J860' spells and 62 'Other' spells. It can be seen in Table 2 that the spells where the patient was initially diagnosed with J860 were on average less efficient.

Table 2: Mean fractions to which observed LOS can be reduced		
Primary Diagnosis		
J869	J860	Other
0.56	0.48	0.56

Statistical tests confirmed that spells where the primary diagnosis is *'Pyothorax with Fistula'* (J860) have indeed larger scope to lower LOS than the rest of the spells. It is noted that J860 is the more serious diagnosis and, all else being equal, would be expected to lead to a longer LOS. This does suggest that perhaps these spells should be assessed as a separate subset. However, on assessing Pyothorax with fistula spells separately only a handful of spells were found to have a significantly different efficiency rating to that obtained when all spells were assessed together. We would recommend that management retain as a basic case the original pooled assessment but in the case of inefficient spells with fistula also examine their efficiency score from the second, within subset assessment, in case the with fistula diagnosis justifies the longer than expected LOS.

5.2 Time to first operation

Clustering the findings by time to first operation (TTFO) enables us to check whether it impacts the scope for reducing LOS. TTFO varied between 1 hour and 3 minutes at its shortest and 17 days, 21 hours and 27 minutes at its longest, though this range is based on only 70% of the spells for which the requisite TTFO existed. On average TTFO in this subset was a little under 2 days. There appears to be some indication that the longer TTFO the larger the scope for saving on LOS. However, the relationship is relatively weak. (The correlation between TTFO and scope for LOS reduction (as a factor of observed LOS) was 0.138). Thus a reduction in TTFO would in principle be conducive to shorter LOS though the scope for reduction from TTFO alone seems limited.

5.3 Admittance location

A further investigation was conducted to ascertain whether the source of admission of the patient had an effect on the scope for reduction in LOS. The spells are separated into those admitted from

‘Residence’ (184 spells) and ‘Other Hospital’. It is found that spells relating to those admitted from other hospital have scope to reduce observed LOS by about 43% compared to 47% for those admitted from Residence. This difference is statistically significant. Moreover, it was found that there is also a statistically significant difference in the LOS itself between the two groups. Those admitted from residence have mean LOS of approximately 12 days compared to just over 10 days for those admitted from other hospitals. Thus for patients admitted from residence we have both longer LOS and higher scope for reduction of LOS. Closer investigation revealed that the shorter LOS for those admitted from another hospital is probably because of shorter TTFO. Those coming from a residence take nearly an extra day before having their first operation.

This finding about TTFO led us to also investigate whether admitance location impacts other treatments within a spell. Interestingly, the variation seen in LOS is replicated for exams, procedures and diagnoses as well. Those admitted from a residence have higher values for all three variables but not in a statistically significant manner. This may be as a result of some of the exams, procedures etc. occurring at the previous hospital, meaning they are no longer required in the hospital under investigation.

From the managerial perspective therefore one finding to take away is to see how those admitted from their own residence can have their TTFO reduced so as to ultimately reduce LOS.

5.4 Variation between treatment teams

Each treatment team is led by a consultant. Clustering the spell efficiencies by treatment team it is possible to see whether there are significant differences between teams in scope for LOS reduction. Such information can help management disseminate best practice across treatment teams. The last column of Table 3 shows for each consultant team the mean efficiency of the spells handled by that team. It is recalled the efficiency of a spell is the proportion to which its observed LOS level can be reduced.

Table 3 Spell efficiencies by initial consultant team

Initial Consultant	No. of spells	Avg LOS	Avg Exams	Avg Proc	Avg Diag	Mean Efficient LOS
C6064661	17	9.941	5.471	4.059	5.176	0.620
C5188516	60	8.95	5.333	3.500	2.933	0.577
C3260582	28	12.536	8.500	4.536	4.321	0.573
C3402959	16	10.813	5.625	4.188	3.375	0.557
C4319047	13	12.000	7.154	5.077	4.231	0.550
C4567778	13	10.308	6.231	4.231	3.615	0.541
C2822363	108	10.769	6.046	3.843	3.454	0.539
C7018160	13	11.538	6.077	4.077	3.154	0.509
C3290662	12	10.000	4.583	3.333	2.333	0.497

There does appear to be a fairly substantial variation between consultants on mean efficient LOS fraction, with the worst performer in terms of low efficient LOS being C3290662. This is despite dealing with spells with some of the lowest exams, procedures and diagnoses among the teams listed in Table 3. Closer look at the cases showed they are mainly patients with Empyema without fistula as their primary diagnosis. So, on the face of it there is no obvious explanation for some of the long LOS levels of this team both relative to other teams but also relative to their own shorter LOS spells. From the management perspective the team may be able to benefit by re-examining their own practices across less and more efficient spells in terms of LOS as well as those of their efficient peer consultant teams.

6. PRACTICE IMPLICATIONS

The primary aim of this paper has been to demonstrate how DEA, a general purpose method for comparative efficiency assessments, can aid health care delivery managers to identify and disseminate good practice in order to improve cost efficiency. The key advantage of DEA is that it rests on a simple comparison of inpatient spells obviating the need for more sophisticated models where an algebraic relationship needs to be postulated between LOS and factors driving it, a relationship which runs the risk of being miss-specified. DEA can compare health delivery activities at a very granular level, including that of an inpatient spell as illustrated here. This in turn makes it possible to reflect very closely the condition of the patient, focusing as in this case, on one primary or a very few closely related primary diagnoses pertaining to the admission. Additional data on patient condition can be taken into account such as secondary diagnoses, perhaps prior admissions to hospital, durations of stay and any other factors that clinicians can identify as pertinent to the condition of the patient on admission. Once the condition of the patient is captured the methodology enables the user to compare inpatient spells on costs (where available) or LOS (as a surrogate for cost). The comparison would identify any systemic inefficient practices since any variation in costs or LOS could no longer be attributed to patient condition which would have in large measure been controlled for.

In our case the aim was to identify practices conducive to cost efficiency in the treatment of children under 16 for empyema. Pyothorax with fistula was found to have lower efficiency than Pyothorax without fistula. However, this lower efficiency may simply be the result of our model being unable to account fully for patient condition on admission. Further investigation by management is needed. Another finding of interest was that admissions from residence tend to have more scope for LOS reduction than those from other hospitals. This merits investigation by management as to how the finding may be exploited to lead to shorter LOS.

The granular level of the analysis at patient-spell level makes it possible for management and clinicians to verify the robustness of the findings by directly contrasting each case with a small number of efficient peer spells. Contrasting LOS inefficient with comparable LOS efficient spells has the added advantage that it enables management to identify operating practices conducive to greater cost efficiency.

There are, of course limitations to the analysis which further research should address:

- The use of simply the number of diagnoses/procedures/lab tests, is a crude measure for, respectively, complexity/seriousness/general health condition of the patient. This shortcoming can be addressed by clinical input, coupled with data analysis aimed at identifying perhaps a weighting structure so that secondary diagnoses accompanying a primary diagnosis can be reflected more appropriately for the condition of a patient on admission.

- The analysis has tacitly assumed that clinical outcomes were 'appropriate' for patient condition on admission. This assumption can be relaxed in one or both of the following ways. Prior to the assessment spells where clinical outcomes may have been below expectation can be removed from the data set to enable comparisons only of spells with appropriate clinical outcomes. In addition, post the assessment inefficient spells can be compared with efficient peers which have at least as good clinical outcomes as the inefficient spell concerned.

- For the comparisons between consultant teams, the sample size for most teams was fairly small in our case. This makes any definitive conclusions about consultant teams hard to justify. Larger samples per consultant team are needed to make findings at consultant team more reliable. Indeed given large enough numbers of spells per consultant team would enable assessments within consultant team and across all consultants to better isolate good and poor consultant team effects.

By using methods such as that in this paper, management at a hospital would be able to identify spells where there is inefficiency. This could then help them to improve any processes followed, by providing a solid basis for any changes made.

7. REFERENCES

1. Charnes A, Cooper W W, and Rhodes E L, (1978) Measuring the efficiency of decision making units , European Journal of Operational Research, 2:6:429-444.

2. Farrell M J, (1957) The Measurement of Productive Efficiency. Journal of the Royal Statistical Society. Series A (General): 120:3: 253-290.

3. Hollingsworth B, (2008) The measurement of efficiency and productivity of health care delivery. Health Economics, 17:1107-28.

4. Liu John S., Lu Louis Y.Y., Lu Wen-Min, Lin Bruce J.Y. (2013) A survey of DEA applications, Omega:Vol 41, Issue 5, pp. 893-902

5. Thanassoulis E, (2001) Introduction to the Theory and Application of Data Envelopment Analysis: A foundation text with integrated software, Kluwer Academic Publishers, Boston, Hardbound, ISBN 0-7923-7429-0, 2001:1-312

6. Thanassoulis E, Portela M and Graveney M, (2016) Identifying the Scope for Savings At Inpatient Episode Level: An Illustration Applying DEA to Chronic Obstructive Pulmonary Disease European Journal of Operational Research, Volume 255, Issue 2:570–582,

APPENDIX

A1: A Graphical illustration of the principles of Data Envelopment Analysis (DEA)

Figure A1, drawn on a simple production context where one resource is used to produce a single output, demonstrates the principles underlying DEA. We assume that we do not necessarily have constant returns to scale and so we cannot use a simple ratio of output to input to measure performance. To measure the *efficiency* of a unit we need to estimate either the minimum resource its output level would justify, or alternatively, the maximum output level its resource level could support. Then the *input efficiency* of the unit will be the ratio of the minimum input needed for its output level to its actual observed input level. The *output efficiency* of a unit is defined in an analogous manner. The two efficiency measures are not necessarily equal under non-constant returns to scale but they are so under an assumption of constant returns to scale.

Figure A1 can now be used to illustrate the concepts underpinning a DEA assessment. At the centre of the approach is the assumption that averages of observed 'units' (in our case inpatient spells) are feasible in principle even if not actually observed in practice. For example all resource – output level combinations on the linear segment U_1U_2 in Figure A1 represent averages of the resource-output levels at U_1 and U_2 using different combinations of weights that add up to 1 for at U_1 and U_2 respectively. The averages can be of any or all the observed units (spells), computed using any set of weights attaching to the units which are non-negative but do add up to 1. If we proceed in this manner with the units in Figure A1 we can construct the 'production space' shaded. This contains all 'units' feasible in principle even if not observed in reality. The boundary $U_1U_2U_3U_6$ of the production space corresponds to maximum output for any given level of resource and so it is referred to as the *efficient* boundary.

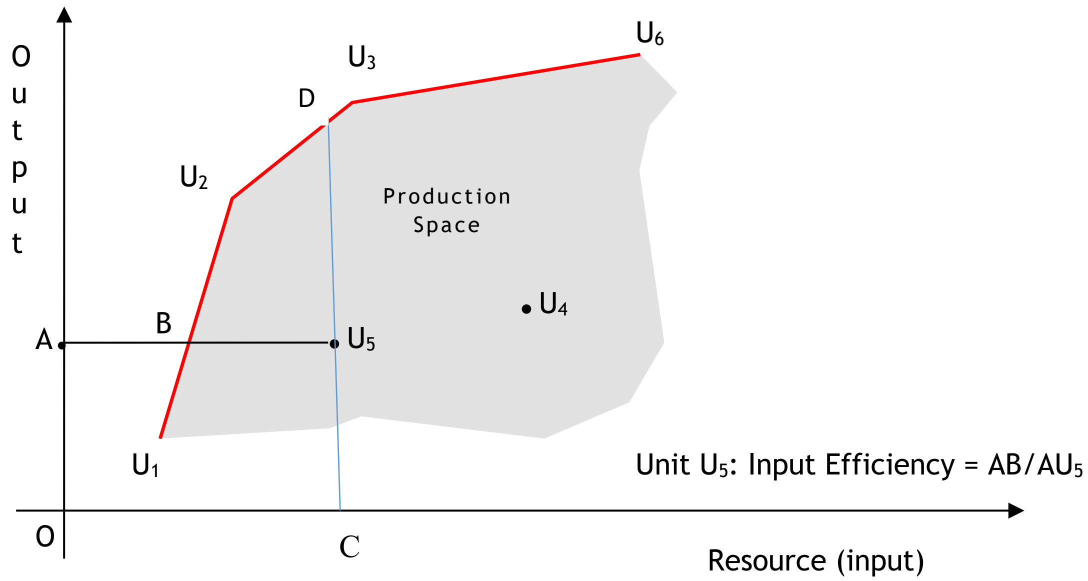

Figure A1: Measuring Efficiency by DEA- Graphical illustration

Using the efficient boundary as a reference, information about the performance of all units can be derived. For example:

1. The boundary units U_1, U_2, U_3, and U_6 are relatively efficient in the sense that no other unit can offer better performance (e.g. lower level of resource) given the output level of each;
2. Units U_4 and U_5 are relatively inefficient in the sense that in each case there exist other production points that offer better performance in terms of output per unit input. Measures of their input efficiency can be deduced by estimating the reduced resource level their output level would justify. For example, in the case of unit U_5 the virtual unit at B shows the level of resource its output level would justify if it had been operating as efficiently as can be deduced from the performance of units U_1 and U_2 whose interpolation is used to create the virtual benchmark unit at B. Thus:

- The input level B in practice is seen as a target for unit U5 to reduce its input while maintaining its output, if it is to be efficient;
- The units U_1 and U_2 are the closest units to its scale size (in terms of output level) and they are the benchmarks it could emulate in terms of operating practices in order to improve its performance. They are referred to as its *efficient peers*;
- A measure of the efficiency of U_5 is given by expressing the minimum resource level justified by its output level as a fraction of its actual resource level. Thus the efficiency of U_5 is AB/AU_5; (In an analogous manner the output efficiency of unit U_5 can be obtained as CU_5/CD by using as target output level at D justified by its resource level OC.)

A.2. The DEA model used

Using the principles outlined above we constructed a **Data Envelopment Analysis** (**DEA**) to estimate the '**efficient**' (i.e. shortest possible) level of LOS for each spell. Let LOS_j , $EXAMS_j$, $PROC_j$ and $DIAG_j$ be the observed levels at spell j for exams, procedures and diagnoses respectively. The estimated lowest level for the LOS of some spell j_0 is derived by solving the model in (M1) to determine the value for θ* which is the fraction to which its observed level LOS_{j0} can be reduced controlling for its $EXAMS_{j0}$, $PROC_{j0}$ and $DIAG_{j0}$.

$$\theta^* = \text{Min}\ \theta$$

Subject to:

$$\sum_{j=1}^{n} \lambda_j LOS_j \leq \theta \quad LOS_{jo}$$

$$\sum_{j=1}^{n} \lambda_j EXAMS_j \geq EXAMS_{j0} \qquad \text{(M1)}$$

$$\sum_{j=1}^{n} \lambda_j PROC_j \geq PROC_{j0}$$

$$\sum_{j=1}^{n} \lambda_j DIAG_j \geq DIAG_{j0}$$

$$\sum_{j=1}^{n} \lambda_j = 1 \qquad \lambda_j \geq 0$$

Model (M1) determines values of the weights λ, adding up to 1, which can be used to create a virtual benchmark inpatient spell which will exhibit as low a proportion θ of the observed LOS_{j0} of spell j_0 as possible, while at the same time relating to at least the levels of $EXAMS_{j0}$, $PROC_{j0}$ and $DIAG_{j0}$ found at spell j_0 .

The model in (M1) was solved in respect of each empyema inpatient spell on which we had data.