European Journal of Operational Research 255 (2016) 570–582

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Identifying the scope for savings at inpatient episode level: An illustration applying DEA to chronic obstructive pulmonary disease

Emmanuel Thanassoulis[a], Maria Silva Portela[b,*], Mike Graveney[c]

[a] Aston University, Birmingham, UK
[b] CEGE, Católica Porto Business School, Rua Diogo Botelho 1327, 4169-005 Porto, Portugal
[c] University of Warwick, UK

ARTICLE INFO

Article history:
Received 19 July 2015
Accepted 13 May 2016
Available online 20 May 2016

Keywords:
Health care
COPD
Efficiency analysis

ABSTRACT

Chronic obstructive pulmonary disease (COPD) is characterised by a largely irreversible obstruction of the airways, and is one of the leading causes of chronic morbidity and mortality worldwide. This paper illustrates the use of Data Envelopment Analysis (DEA) to assess the potential for cost savings at COPD inpatient episode level. The analysis uses the length of stay of each episode as a surrogate for expenditure on that episode while allowing for the medical condition of the patient and the quality of care received. We find substantial possible reductions in length of stay which would translate to cost savings. The paper also explores differences both between hospitals and between care teams within hospitals so that cost efficient protocols of treatment can be identified and disseminated.

1. Introduction

Chronic obstructive pulmonary disease (COPD) is characterised by a largely irreversible obstruction of the airways, and encompasses both emphysema and chronic bronchitis (Afonso, Verhamme, Sturkenboom, & Brusselle, 2011). According to the World Health Organization (WHO) COPD is one of the leading causes of chronic morbidity and mortality worldwide. This paper investigates the comparative *efficiency* of certain UK hospitals in handling this disease using data at inpatient episode level. It sets the length of stay (LOS) of an episode against the medical condition of the patient, including co-morbidities with COPD, age and gender with the view to identifying the potential for reduction of LOS at no detriment to the quality of treatment the patient has received. While performance is assessed at inpatient episode level, the results are then aggregated at higher levels, including hospital level, with a view to identifying systematic best practice at hospital or constituent unit level. We assess the scope for savings using Data Envelopment Analysis (DEA). This is because of its flexibility in estimating a piece-wise-linear cost frontier which envelops the observed data rather than requiring the specification of an assumed functional form for that frontier. We thus avoid the risk of mis-specifying a frontier functional form and of making assumptions about the probability distribution of any inefficiencies. Further, DEA has the advantage of identifying benchmark episodes of treatment which are most suitable for each episode which has scope for savings. Using episode-specific benchmarks enables non-technical hospital managers to gain confidence in the findings when comparing instances of treatment within their hospital on costs and quality of service. Episode – specific benchmarks also are likely to be the most suitable examples from which best practice for adoption can be drawn.

In the UK there has been a growing interest for analyses at patient level since the introduction of patient-level costing systems in the National Health System (NHS) in the mid 2000s. However, “in 2011 less than half of NHS trusts had patient-level information and costing systems (PLICS)” (Blunt & Bardsley, 2012). NHS hospitals are paid on the basis of a fixed price per diagnosis-related admission, designated ‘tariff’ under the Payment by Results (PbR) framework. The tariff is computed from national reference costs for each Healthcare Resource Group (HRG). Each medical diagnosis maps to an HRG. The tariff can be complemented in cases where the patient requires specialist care or longer length of stay. As will be seen in Section 3, the cost data returned by hospitals for treating each inpatient episode revealed serious problems of reliability. This made it impossible to use cost data in our analysis. However, we could use Length of Stay (LOS) data, which is far more reliable. LOS can be seen as a surrogate for the expenditure by the hospital in treating a patient, given that we have controlled for primary diagnosis and use surrogates for any co-morbidities the patient has.

To the best of our knowledge this is the first paper to analyse the potential savings in the treatment of COPD patients. The DEA models used for this purpose, though not totally new, possess

* Corresponding author. Tel.: +351 226196200.
E-mail addresses: e.thanassoulis@aston.ac.uk (E. Thanassoulis), csilva@porto.ucp.pt (M. Silva Portela), mikegraveney09@gmail.com (M. Graveney).

http://dx.doi.org/10.1016/j.ejor.2016.05.028
0377-2217/

some methodological novelties necessitated by the fact that the input variable used was truncated (as only episodes with LOS of 2 or more days were considered). Hospitals are compared on their ability to reduce costs through reductions in LOS. Consultants within each hospital are also compared on their ability to reduce costs given the characteristics of the episodes each one handles. This study contributes to the scarce literature on applying non-parametric frontier models to patient level data.

We find substantial possible reductions in length of stay of COPD patients, which would translate in cost savings. This is in line with previous literature suggesting that hospital capacity can be increased through reduction of LOS without compromising quality (see Berk and Moinzadeh, 1998). We find, however, in some cases that there is a trade-off between LOS and time to re-admission, with too short lengths of stay generating sooner readmissions for patients with a high disease weight. We also find that hospitals and care teams within hospitals can have differing performance on length of stay, sometimes by patient gender, not entirely justified by the disease weight of the patients treated. We suggest such differences be investigated with a view to identifying treatment protocols, which are economical without detriment to quality of care.

2. Motivation for analysing health care costs at patient level

The assessment of cost efficiency or productivity of health service entities at the aggregate level has been criticised by many authors (e.g. Hollingsworth, 2008; Bradford, Andrew, Kleit, Krousel-Wood, & Richard, 2001). This is primarily because at the aggregate level one cannot reflect accurately the multiplicity of differences between treated patients. The best predictor of a patient's treatment costs is the medical diagnosis for which the particular instance of treatment is intended and the accompanying general health condition, including age and gender of the patient. This has long been recognised, and most cost analyses in the health care field use one or other 'case-mix' measures as an explanatory factor for treatment costs when analyses are at aggregate level. However, such a 'case mix' is still too broad a category to reflect accurately the treatment needs of a cohort of patients.

In this paper we narrow down case-mix to cases where the 'primary' diagnosis relates to just 4 Health Resource Groups (HRGs), all related to the COPD condition. This is the first step in controlling for the disease weight each patient presents, which drives their treatment cost. However, it is well known that even within each HRG a large variability of costs exist due to various factors. This can be because both the HRG does not fully capture the medical condition of a patient, and therefore his or her clinical needs, and also because of the efficacy of the operating practices both clinical and managerial of a hospital, which can impact costs. The aim in this paper is to isolate the latter costs by controlling for the clinical needs of the patient. One way to control for the clinical needs of the patient is to use patient and indeed episode level data. These reflect more accurately the disease weight of the patient at the point of admission. Moving from aggregate data (e.g. hospital or General Practice level) to patient episode level data, makes it possible to derive findings specific to a particular health condition (diagnosis).

Most studies in the health literature using data at patient/episode level have relied on multilevel regression models, since patients are nested into services and services are nested into hospitals (see e.g. the literature review of Diez-Roux, 2000). These multilevel studies have been mainly concerned with identifying the cost drivers of patients with a certain disease, and assess the impact of each cost driver and the variability of costs across hospitals. For example, Dormont and Milcent (2004), Kristensen, Laudicella, Ejersted, and Street (2010), Laudicella, Olsen, and Street (2010), Olsen and Street (2008), and Daidone and Street (2011) used multilevel models to analyse the costs of treating patients with various diseases. Dormont and Milcent (2004) analysed cost variability of French hospitals based on 7314 stays for acute myocardial infarction during 1994–1997; Olsen and Street (2008) were interested in patients attending vascular departments in 8 Danish hospitals in 2004; Kristensen et al. (2010) analysed all diabetic patients in the UK in 2005/2006, in a total of 31,371 patients, and Laudicella et al. (2010) analysed patients from 136 obstetrician departments in the UK in 2005/2006. Daidone and Street (2011), on the other hand, analysed UK patient data from 2008/2009 and assessed 12 m patients not restricting to a single disease or specialty, but considering all patients receiving diverse types of treatment. As it would be impossible to include in the regression models dummy variables for about 1400 HRGs, Daidone and Street (2011) divided each patient cost by the mean cost of all patients within the same HRG. They used this standardised cost variable as the dependent variable in a regression model. The use of this variable implied that the authors were investigating the drivers of cost deviations from average and the effect of the hospital on these deviations.

All the above multilevel studies attempted to explain costs of treatment, where a number of control variables (relating mostly to patients) have been taken into account. Examples of such types of variable are the age, gender, race, socio-economic status, number of operations, number of diagnoses, existence of other conditions or co-morbidities in addition to the disease being analysed, emergency admissions, transfers, death, the HRG group and, diagnostic markers based on ICD-10 (the international classification of disease, 10th revision). In multilevel models, it is possible to isolate the proportion of total variation in cost that is due to patient characteristics and that due to hospital/department characteristics. The latter is usually called the hospital/department effect. This effect can also be explained by additional variables (this time measured at the hospital or department level). DEA models too can consider the multilevel nature of data, but through a different approach where all units are analysed together (with reference to a pooled frontier) and also within a certain hierarchical level (e.g. the hospital).

DEA has been used before for patient/episode level analyses. For example, Puig-Junoy (1998) used DEA to assess 16 Intensive Care Units in Catalonia, Spain using data on 993 patients, and Dervaux et al. (2009) used a directional distance function in a Free Disposal Hull (FDH) technology using data on 15,029 patients treated in 25 Parisian ICUs during 2000. Bradford et al. (2001) used Stochastic Frontier Analysis (SFA) to compare the costs of two technologies for treating coronary artery disease in 645 patients observed in the year of 1994, while Dismuke and Sena (2001) used a Malmquist–Luenberger framework to analyse quality-adjusted measures of productivity change for three diagnostic technologies commonly used in the main Portuguese hospitals over the period 1992–1994.

Bradford et al. (2001) model was very similar to the multilevel models described above in that the costs of treating patients for cardiac revascularisation were explained through patient characteristics (such as co-morbidities, age, gender, etc.), and a dummy for the doctors treating the patient.

Regarding the use of DEA models, Puig-Junoy (1998) used patient level data on 2 outputs (number of days surviving in the hospital, and the discharge status - death or surviving at hospital discharge) and some inputs reflecting patient characteristics (e.g. survival probability at admission, mortality risk level, weighted ICU days, non-ICU hospital days), and clinical practice (e.g. available nurse days per patient, available physician days per patient, and technological availability.) Puig-Junoy (1998) also used a second stage regression analysis to explain the efficiency yielded by the DEA analysis, where explanatory variables included the type