

Using DEA to Estimate Potential Savings at GP units at Medical Specialty Level

Emmanuel Thanassoulis, *Aston University, UK*

Maria C. A. Silva Portela, *Portuguese Catholic of Porto, Portugal*

Mike Graveney, *University of Warwick, UK*

Abstract

The aim of this paper is to identify benchmark cost-efficient General Practitioner (GP) units at delivering health care in the Geriatric and General Medicine (GMG) specialty and estimate potential cost savings. The use of a single medical specialty makes it possible to reflect more accurately the medical condition of the List population of the Practice so as to contextualize its expenditure on care for patients. We use Data Envelopment Analysis (DEA) to estimate the potential for cost savings at GP units and to decompose these savings into those attributable to the reduction of resource use, to altering the mix of resources used and to those attributable to securing better resource 'prices'. The results reveal a considerable potential for savings of varying composition across GP units.

Keywords: DEA, primary health care, cost efficiency, allocative efficiency, price efficiency.

1. Introduction

In the UK health care is free at the point of delivery for all residents. All individuals covered by the National Health Service (NHS) are registered with a General Practitioner (GP) who in turn normally delivers his/her services within a GP unit, or Practice, comprising one or more GPs. The Practice is the gateway for the individual to health care under the NHS. A GP within a Practice is responsible for the clinical decision whether or not to refer a patient for further medical care, (normally at a hospital), the type of initial referral (e.g. inpatient or outpatient treatment) and indeed the medical specialty that would be appropriate for the referral. The (primary) medical diagnosis associated with a referral would place the patient in a 'Healthcare Resource Group' (HRG) and the care provider - generally a hospital - will be compensated then by the Primary Care Trust at the tariff applicable to that HRG. It is clear therefore that the costs

incurred by a PCT for medical care of the patients it covers depend crucially on the clinical pathways the GPs decide upon. This study in effect captures the economic consequences of varying GP choices on pathways, controlling for disease weight, under an implicit assumption that patient outcomes are not compromised albeit they may be reached by alternate pathways.

Our analysis in this paper follows the work initiated in Thanassoulis et al. (2012) where 75 GP units have been compared on referral and drug prescription costs. In that paper overall inpatient referrals, outpatient referrals and drug prescription costs were compared amongst GP units controlling for their list population characteristics. The focus on referrals relates to the fact that the way physicians approach such treatment collaboration influences several aspects of patients care, including its quality and cost. As mentioned in Barnett et al (2011), in a survey comparing referral decisions of specialist and primary care physicians, different physicians have different reasons for choosing their referral partners. (In their study they found, for example, that primary care physicians were more likely to be concerned with patient access than specialist doctors).

In Thanassoulis et al. (2012) the aim was to understand how Practices compare with each other on costs and volumes of referrals and drug prescriptions. Although differing costs amongst Practices were contextualised by the characteristics of the list population served, those characteristics reflect only indirectly their actual health needs. In order to try to account for the health needs of the people served by the Practice more directly, we need to know the type of diseases they had and the complexity and severity of the condition they presented at the point the Practice decided on a treatment pathway. However, the disease types are varied and for clinical and accounting purposes they are classified in this study into 11 clinical specialties. In addition, within each specialty patients entering the hospital are assigned a primary diagnostic (ICD –International Classification of Diseases) code. It is this code which is principally responsible for placing them within a HRG which in turn determines the financial remuneration or payment the care provider will get for the patient episode concerned. In the UK there are several hundred HRGs onto which several thousand ICDs are mapped. Base tariffs for patient (or more precisely ‘consultant’) episodes within each HRG are set nationally by the Department of Health.

In order to account for the varied health needs of the population, we undertake in this paper a comparative analysis of costs of Practices (we are modeling particularly inpatient and outpatient referrals costs and drug prescription costs) for one particular medical specialty. The

specialty analysed is the **General Medicine and Geriatrics** (GMG) combined specialty. This is amongst the specialties representing the highest percentage of the total costs of a Practice. (It represents on average about 22% of total inpatient costs of a Practice and 12% of its total outpatient referral costs.)

Controlling for the GMG related disease weight of each Practice we identify three potential components of cost savings: savings through reducing the volumes of referrals and drugs, savings through switching between types of referrals and/or drugs, and savings through securing a better unit cost profile for referrals and drugs. The latter component may appear counter-intuitive given that most payments for referrals and drugs are standardized and based on 'Payment by Results' (PbR) national tariffs. Nevertheless Practices can to an extent affect unit costs by better targeting providers at referral stage. For example, indicating accurately the disease code applicable to a referral will avoid a referred patient undergoing a second diagnosis in hospital to correct a poorly specified initial one, for which nevertheless a fee would still be paid by the PCT. Similarly, a Practice may press for prompt hospital discharge to reliable family care where applicable thus reducing length of stay and avoiding the possibility of hospital complications and surcharges on the national tariffs.

We use Data Envelopment Analysis (DEA) to compute the potential cost savings at each Practice as we progressively relax the assumptions about the possibility of switching between types of referral and drug use, and introduce the notion of potentially improving on unit costs at each Practice.

This paper is structured as follows: In Section 2 we present some previous studies on primary health care efficiency, analyzing with some detail those that most closely relate to the approach presented in this paper. In Section 3 we present the models used, and in Section 4 the results for 75 GP units analysed. Section 5 concludes the paper.

2. Related Literature

Health studies can be undertaken at several levels of analysis (e.g. patient, service department, hospital, primary care units, health districts, etc.). In his most recent literature review in this area, Hollingsworth (2008) reports that 50% of health applications concern hospitals and only about 10% concern primary care (in a total of 17 studies). In Amado and Dyson (2008) a literature review on primary health care applications is undertaken and the authors report about 21 studies using the DEA methodology.

The studies on primary care have mainly looked at the efficiency of providing the health service, i.e. the efficiency of translating resources available (staff, materials and technology) into health intermediate outputs (such as consultations and treatments of various types). Examples of such studies include Szczepura et al. (1993), Zavras et al. (2002) or Amado and Santos (2009). This type of studies can be seen as addressing health care provision from a managerial perspective (as termed by Chilingirian and Sherman, 2004) as opposed to a clinical perspective, where the focus is on analyzing how a Practice utilizes the minimal quantity of clinical resources (such as consultations, referrals, treatments, and drugs) to achieve a constant quality outcome, when caring for patients with similar diagnosis complexity and severity. These perspectives embody two important concepts in health contexts: that of outputs or intermediate outputs (related to the quantity of care provided) and that of outcomes (related to the quality of care provided) (see Donabedian, 1980). Typically the transformation of inputs into intermediate outputs is seen as an efficiency assessment and the transformation of intermediate outputs into outcomes is seen as an effectiveness assessment. Amado and Dyson (2009) complement these perspectives with a third perspective of assessment related to equity. For that purpose, they define in addition to inputs, intermediate outputs, and outcomes, also local needs. The comparison between the services provided (intermediate outputs) and the local needs yields a measure of relative equity across Practices.

A number of studies, e.g. Salinas-Jimenez and Smith (1996), Luoma et al. (1996), Puig-Junoy and Ortún (2004), Staat (2003, 2011), Sørensen et al. (2009), and Thanassoulis et al. (2012) have adopted a cost perspective for the assessment of GP units, which can be seen as an additional perspective of analysis relative to the managerial and clinical perspectives noted above. All these studies have compared the costs of providing the service (like referral or medication costs), controlling for certain characteristics of the patients concerned such as age, gender, or level of deprivation. For example, Sørensen et al. (2009) used as a control variable the population served by the Practice divided into 22 categories reflecting age, gender, education and employment. These 'output' measures are not intermediate outputs or outcomes as reported above, but they work as control variables for undertaking cost comparisons between Practices.

There are not many studies on primary care that focus on the analysis of a single specialty. Specialty or disease level analysis is more frequent in hospital studies, where the complexity and heterogeneity of services offered and patients treated recommends the benchmarking of similar hospital units (like obstetrician departments in Laudicella et al., 2010, or intensive care units (ICU) in Puig-Junoy, 1998 or Dervaux et al., 2009) or groups of diseases (e.g. like diabetes

in Kristensen et al., 2010, or the DRG - Heart Failure and Shock in Dismuke and Sena, 2001). Some examples of studies in primary care focusing on a given specialty/disease are Linna et al. (2003), who analysed oral health care in Finnish health centers through a DEA method, Amado and Dyson (2009), who analysed diabetes in primary care in the UK using DEA, Thanassoulis et al. (1995) who used DEA to assess the provision of perinatal care in district health authorities in the UK, or Ozcan (1998) who used DEA to assess otitis media treatment by primary care physicians in the US. Even if these studies focus on a particular disease they still use different levels of analysis: physicians in Ozcan (1998), district health authorities in Thanassoulis et al. (1995), and health centers in Linna et al. (2003) and Amado and Dyson (2009). Given the differences in the assessed disease in each of the above studies, the input/output specifications are also different and are not indeed comparable. In some of these specialty/disease studies a managerial perspective of analysis was used, where inputs related to full time equivalent doctors and nurses dedicated to the disease (e.g. Linna et al, 2003, Amado and Dyson, 2009, Thanassoulis et al., 1995). The outputs used varied by disease or specialty. They related to the number of patients or visits to dentists in Linna et al. (2003), the number of diabetics with a complete annual review, on targeted anti-hypertensive medication and on target statins in Amado and Dyson (2009), and to the number of birth episodes, of deliveries, of abortions, of special care consultants episodes, and of intensive care consultant episodes in Thanassoulis et al. (1995). A clinical perspective of analysis was also followed in Amado and Dyson (2009) where the authors used the above mentioned outputs as inputs of a DEA model, whose outcomes were the number of diabetics with blood pressure, blood glucose and blood cholesterol under control. Thanassoulis et al. (1995) were amongst the first to consider some outcomes in the analysis, but did that in a single model rather than through two separate models as done in Amado and Dyson (2009) (and advocated in Chillingirian and Sherman, 2004). They considered service quality variables related to number of very satisfied, and number of satisfied mothers with the service, and a variable relating to the quality of the medical outcome: survival rate of babies at risk. Ozcan (1998), on the other hand, used a set of inputs that relate to a clinical perspective, but on the output side of the analysis he did not use outcome measures, but variables that reflected the severity weight of the treated patients (number of low severity, medium severity and high severity patients). Inputs of the analysis were intermediate outputs relating to the number of visits to the primary care physician and to the specialist, the number of inpatients, and the number of prescriptions and laboratory procedures. Our analysis follows more closely the approach of Ozcan (1998) in the sense that we also use some proxies for controlling for the severity of patients treated at the GMG

specialty, but on the input side we are concerned with referrals and drug prescriptions volumes and costs.

3. DEA Models Used

The aim of our study was to identify potential cost savings at 75 Practices in providing care for patients with medical conditions treated within the General Medicine and Geriatrics (GMG) clinical specialty. The savings were to be identified as in Thanassoulis et al. (2012) as we progressively relax the assumptions about available potential trade-offs between type of referral and drugs and also exploiting flexibilities in securing better unit costs for referrals and drugs. We used DEA for estimating the potential savings. The identification of the input-output variables to be used is detailed in the next section. We use generic variables in this section in order to present the DEA models used.

Consider for Practice j ($j = 1, \dots, n$) a vector $\mathbf{x}_j = (x_{1j}, x_{2j}, \dots, x_{mj})$ reflecting m inputs consumed for producing a vector of s outputs $\mathbf{y}_j = (y_{1j}, y_{2j}, \dots, y_{sj})$. Consider also, that prices of inputs are known and given for Practice j by a vector $\mathbf{p}_j = (p_{1j}, p_{2j}, \dots, p_{mj})$. To estimate savings through simply reducing input quantities we used the model in (1). This yields in DEA terms the “Technical efficiency” for Practice o in the form of the optimal value θ^* of θ . Note that the model in (1) assumes variable returns to scale (VRS) which is a realistic assumption in our study as will be clear when we detail the input-output variables.

$$\text{Min}_{\theta, \lambda_j} \left\{ \theta \mid \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io}, \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0 \right\} \quad (1)$$

The optimal value θ^* in (1), indicates the degree of cost savings that Practice o can achieve by reducing input volumes radially, without compromising the level of outputs produced.

To estimate potential savings taking into account not only the quantities of inputs used but also their unit costs we solved the classical “cost efficiency” DEA model reproduced in (2). In this model for Practice o we determine the levels of the input quantities x_i that would minimize the aggregate cost C of delivering its output levels y_o . The model maintains as in (1) the assumption of variable returns to scale.

$$\text{Min}_{x_i, \lambda_j} \left\{ C = \sum_{i=1}^m p_{io} x_i \mid \sum_{j=1}^n \lambda_j x_{ij} \leq x_i, \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \sum_{j=1}^n \lambda_j = 1, \lambda_j, x_i \geq 0 \right\} \quad (2)$$

The model leads to the measure of “Cost efficiency” of Practice o as the ratio between the minimum cost (derived from (2)) (C^*) and observed cost, (C^*/C^o). Cost efficiency can be decomposed into technical and allocative efficiency. The latter is computed as: cost efficiency/technical efficiency.

Allocative efficiency measures the savings that are feasible after a Practice achieves full technical efficiency as determined in model (1). These additional savings can result by a Practice aligning the mix of inputs (and not just reducing pro rata their volumes) more in line with observed input prices.

Model (2) implicitly takes input prices as exogenously fixed. However, if we assume that a Practice can to an extent avail itself of input prices that might be more favourable for lowering the aggregate cost of its outputs than its current prices, then additional savings might be possible. To identify the potential for such savings we solved model (3). This model leads to the measure of the “*price efficiency*” of Practice o . In model (3), (introduced in Thanassoulis et al. (2012) and inspired by Tone and Tsutsui (2007)), $C_i^T = p_i x_i^*$ represent component costs, the i th relating to input i , using the level x_i^* of input i as derived from the cost minimizing model in (2).

$$\text{Min}_{C_i, \lambda_j} \left\{ \sum_i C_i \mid \sum_{j=1}^n \lambda_j C_{ij}^T \leq C_i, \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0 \right\} \quad (3)$$

Let the optimal objective function value in (3) be C^{**} . It is easily seen that $C^{**} \leq C^*$. If $C^{**} < C^*$ then model (3) has identified savings beyond those identified by model (2). These savings would be consequent on altering input prices and input mix simultaneously. There is no easy way to disentangle the effects of input price and mix alterations. Given that model (2) exhausts savings attainable through altering input volumes and mix to exploit observed prices at the Practice we will attribute any further savings identified by model (3) to input price changes though acknowledging that the savings could also be the result of input volume mix as well as input price changes. However, the volume mix changes at this stage are consequent to input price changes and so in a sense all savings identified by model (3) are the result of input price changes.

Note that overall cost efficiency is C^{**} / C^o , the denominator being the observed cost at Practice o . This ratio can be seen as the product of cost efficiency defined earlier as C^* / C^o ,

and “price efficiency” defined as C^{**} / C^* , since we have: $\frac{C^{**}}{C^o} = \frac{C^*}{C^o} \times \frac{C^{**}}{C^*}$.

4. Identifying inputs and outputs at the GMG Specialty Level

The analysis at the specialty level revealed some challenges in identifying inputs and outputs, partly because the data needed is not generally available at specialty level. Table 1 shows the input-output variables that were ultimately identified for assessing the GMG specialty of Practices. We detail here their definition and derivation.

On the left column under ‘Inputs’ we show the modelled costs. These are broken down to the three main types of cost a GMG patient can incur. That is, as an inpatient at some hospital, as an outpatient at some hospital and through drug treatment outside hospital, though this may be as a follow up of a spell as an in or out patient. We use these variables in turn as volumes or costs of referrals depending on whether technical or cost efficiency is being assessed. We have used the actual rather than tariff unit cost because we want to identify any potential savings in actual rather than tariff expenditure. Despite the PbR tariff system, contractual arrangements with provider hospitals can mean that a PCT pays in some instances a different amount to what the PbR tariff would dictate.

Table 1: Input and output variables: the GMG specialty

Inputs	Outputs
- Inpatient referrals (cost and volume)	- Nature and Volume of Diagnoses driving inpatient referrals (SI);
- Outpatient referrals (cost and volume)	- Nature and Volume of Diagnoses driving outpatient referrals (SO);
- Drug prescriptions (cost and volume)	- Practice list population above 48 years of age

The variables listed as Outputs in Table 1 can be seen as control variables so that Practices can be compared on the Inputs. Collectively the three variables are intended to reflect the disease and age weight driving the GMG specialty costs being modeled. Of the variables listed in Table 1 only inpatient referrals, outpatient referrals and the aggregated List population

above 48 years of age are directly available for the specialty being modelled. The remaining variables were derived indirectly, since they were available for the practice as a whole but not for the GMG specialty. The remainder of this section describes how the variables concerned were derived.

4.1 Variables to reflect the need for GMG related drug prescriptions

Drug costs and volumes were available at Practice level and not at the specialty level as needed. The input variable relating to drug volumes and costs is therefore a constructed variable, which is in turn the sum of the volumes and costs of a number of drug categories that were deemed to relate primarily, if not exclusively, to the GMG specialty.

The approach we used for identifying the drug categories that were deemed most related to the specialty under assessment was to regress each drug cost (we had 59 drug categories and respective costs) on three variables: Practice list population, inpatient volumes for the GMG specialty and outpatient volumes for the GMG specialty. If the cost of a certain category of drug was explained mainly by list size and not by the volume of inpatient and outpatient referrals within the GMG specialty, then this category of drug was deemed not to be related much with this specialty; on the other hand, if the cost of a certain category of drug was explained mainly by the specialty inpatient and outpatient referrals and not by list size, and the model provides a good fit (R-squared above 70%), then that category of drug was deemed to be related primarily with the GMG specialty. As a result of this analysis we identified 21 out of the 59 drug categories as relating primarily to the GMG specialty. These ranged from Lipid regulating drugs to drugs used in diabetes, antidepressant drugs and so on. The list of drugs was discussed with clinical doctors and health care managers in a specially convened workshop and was deemed a good proxy for the drugs that are mostly prescribed to patients diagnosed with conditions covered by the GMG specialty.

The volumes and costs of the GMG related categories of drugs identified were summed in turn to yield the variable for the volume and cost of drugs used as an input in Table 1.

4.2 Diagnoses driving referrals and drug prescriptions of the GMG specialty

One key way to capture the medical condition of patients, as it pertains to the need for referrals and drugs of the GMG specialty, is with reference to diagnoses mapped to the GMG specialty. In the year of analysis (2007/8) ICD codes were available only for those who had had at least one episode as an inpatient. The primary diagnosis (ICD) of each inpatient leads to a Healthcare Resource Group (HRG) tariff. We identified 52 HRGs that related very directly with the GMG specialty with the help of a Physician. These varied from the most frequent HRG like

F06 - *Diagnostic Procedures, Oesophagus and Stomach*, or E36 - *Chest Pain <70 with or without out complication* to some less frequent HRGs like A20 - *Transient Ischaemic Attack >69 or with complications*. These diagnoses have been suitably aggregated to construct the first two proxy variables appearing under ‘outputs’ in Table 1. The proxy variable termed SI (*Severity of health conditions of Inpatients*) is derived from the GMG specialty inpatient diagnoses. It is the aggregate quantity computed as:

$$0.15 \left(\sum_{i=1}^{53} \text{elective tariff}_i \times \text{volume HRG}_i \right) + .85 \left(\sum_{i=1}^{53} \text{non - elective tariff}_i \times \text{volume HRG}_i \right) \quad (4)$$

In (4) each one of the first 52 HRGs is weighted by the tariff for that HRG category. This tariff reflects the cost for treating the diagnosis concerned as set by the Department of Health. The tariff reflects among other things the severity of the disease concerned, any comorbidities, age and gender of the patient. HRG tariffs are different depending on whether a patient is an elective or non-elective (emergency) admission. The above computation of the proxy variable reflects the fact that on average 15% of inpatient admissions to the GMG specialty were elective and the rest non-elective. For the first 52 HRGs, those accounting for most inpatients in the GMG specialty by volume, we have used the actual tariff. The remaining HRGs were grouped into HRG category 53 and its tariff was taken as the average tariff of all HRGs included in this category labelled as ‘other’.

The foregoing proxy variable reflects primarily the volume and cost of medical conditions within the list population driving inpatient referrals and indirectly the use of drugs. A second proxy variable was sought (SO – *Severity of health conditions of Outpatients*), which would more directly relate to medical conditions driving outpatient referrals to the GMG specialty and drug use within that specialty. Here we faced the problem that for outpatients there is no diagnosis or HRG associated with each outpatient referral. We proceeded nevertheless on the assumption that inpatient referrals, whether or not to the GMG specialty, do reflect medical conditions in the aggregate across the list population of a Practice and some of them, depending on inpatient diagnosis, could have had outpatient referrals to the GMG specialty. In addition, list patients without any inpatient referral in the year under consideration could have also had outpatient referrals to the GMG specialty. A preliminary analysis of outpatient referrals found that 53% of outpatient referrals to the GMG specialty had not had an inpatient referral of any kind in the year of analysis while the rest (47%) of outpatient referrals to this specialty did have at least one inpatient referral that year, but not necessarily to the GMG specialty.

To identify the inpatient health conditions that most closely relate to outpatient referrals in the GMG specialty we proceeded as follows. Using patient level data we examined all categories of inpatient diagnosis, and for each category we aggregated the number and cost of outpatient referrals to the GMG specialty. In this way we could identify two-digit ICD categories which tended to also generate volume and cost of outpatient referrals to the GMG specialty. The following percentages, were then used to select ICDs that generate referrals to the GMG specialty:

- PC1: Percentage of all outpatient referrals to the GMG specialty accounted for by outpatients who had a given ICD;
- PC2: Percentage of total cost of outpatient referrals to the GMG specialty accounted for by outpatients who had a given ICD.

These related percentages capture the ICDs that drive outpatient volumes and costs of the GMG specialty from those persons in a Practice's list who also have had some type of inpatient episode in the year of analysis. The values of PC1 range from 0% to 6.55% and the values of PC2 range from 0% to 5.94%. The highest percentages pertain to ICD K5 - *non-infective enteritis and colitis; other diseases of intestines*, meaning that about 6.6% of outpatient referrals to the GMG specialty pertained to patients that have also been referred as inpatients under the ICD K5. We selected those ICDs that showed PC1 or PC2 above 1%. 33 ICD categories satisfied these criteria and their costs were summed to construct the second proxy variable (SO) appearing under outputs in Table 1.

The first two proxy variables under outputs just described do capture the impact on referrals and indirectly on drugs for the GMG specialty but only in respect of patients who had had some inpatient referral. Clearly the outpatient referrals and drugs of the GMG specialty are also driven by list patients who did not have any inpatient referral in the year under consideration. In order to capture such additional demand for the GMG specialty outpatient referrals and drugs we regressed each one of the three variables under Inputs in Table 1, both in cost and volume format, on the two proxy output variables and additional variables reflecting the list size of the practice by age and gender as defined in Table 2. The age/gender groupings control, in very broad terms, for the effect of age and gender on health services needs faced by a Practice. The best regressions identified in terms of explanatory power were those that involved, in addition to the two proxy output variables constructed above, **the number of list patients a practice had that were aged above 48 years**. This constitutes the third proxy variable listed under Outputs in Table 1.

Table 2: Age/gender groups considered

Age/gender group	Average % of Practice list patients
age < 3	3.09%
age >= 3 and age < 18	17.58%
age >= 18 and age < 35 and sex = Male	10.23%
age >= 18 and age < 35 and sex = Female	9.84%
age >= 35 and age < 48 and sex = Male	10.48%
age >= 35 and age < 48 and sex = Female	9.92%
age >= 48 and age < 70	27.11%
age >= 70	11.74%

Regressions were also run in which each age/gender band in Table 2 was divided into patients above and below the 80th percentile on “index of Multiple Deprivations” (see Thanassoulis et al., 2012). However, the additional variables of ‘deprivation’ did not prove significant in this context, partially because patients on each age band, above and below the 80th national percentile of deprivation, was considered at the Practice level rather than at the Specialty level.

Table 3 shows the explanatory power the three proxy output variables have for each one of the three inputs variables in turn when regressed in volume format (VI, VO and VD) or cost format (CI, CO and CD). I, O and D stand respectively for in, and out patient referrals and D for drugs, all relating to the GMG specialty.

Table 3: Results from regressing each input variable on the three proxy output variables of Table 1.

	CI	VI	CO	VO	CD	VD
Proxy for inpatient referrals (SI)	462(.000)	.0481(.000)	-55.5(.000)	-.415(.000)	-146(.214)	9.762(.613)
Proxy for outpatient referrals (SO)	571(.000)	-.025(.316)	251(.000)	2.14(.000)	1387(.000)	93.16(.016)
List Age 48+	9.15(.409)	-.004(.170)	15.29(.000)	.12(.000)	189.7(.000)	18.75(.000)
R ²	97.0%	99.4%	93.5%	93.9%	92.9%	83.7%
St error	48094.74	13.80	11594.05	96.33	1.27009E5	20911.23
(% average)	(9.85%)	(4.7%)	(14.99%)	(14.37%)	(15.71%)	(23.8%)

Values in bold correspond to the coefficients that are statistically significant in each regression. As can be seen the 3 proxy output variables explain well over 90% of the volume or cost of each input variable, the exception being the volume of the GMG specialty prescriptions which are explained only at about the 84% level. These levels of explanatory power are

considered satisfactory and we proceeded with the analysis using the 'input' - 'output' set in Table 1, with the proxy variables constructed as outlined above.

One additional issue was that of returns to scale that should be maintained in setting up the DEA models. We opted for the least restrictive assumption, that of variable returns to scale. That is, given the variables in Table 1 we cannot assume that a proportional rise or fall of referrals or drugs will match precisely a proportional rise or fall in the levels of the output variables.

5. Results

Our analysis relates to data for the financial year 1/4/2007 to 31/3/2008. In terms of contributions to overall costs, drugs represent the highest share at 58.85% of the total costs (see Figure 1). Inpatient referral costs account for 35.52% of total costs and the balance of 5.63% is accounted for by outpatient referrals. It should be noted that the contribution of drug costs to overall costs in the GMG specialty may be exaggerated here in that the drugs we have identified as relating in the main to the GMG specialty may also include patient prescriptions which are not related to the GMG specialty. The converse is of course also possible that the drugs we have excluded may include prescriptions for patients who might otherwise have had or indeed did have the GMG specialty referrals. This ambiguity cannot be resolved without attribution of drugs to specialties at the prescription stage, data not available to us. The percentage of 58.85% of total costs being accounted for by drugs is not, however, too far from the 51% of costs that drugs account for when costs are seen at Practice rather than specialty level.

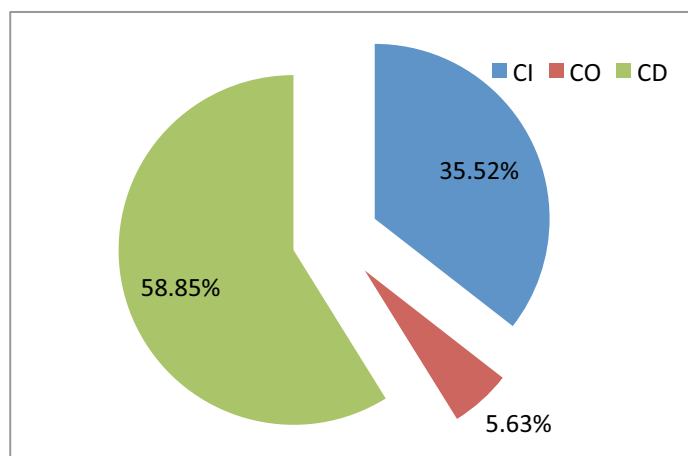


Figure 1: Break down of total costs of the GMG specialty

5.1. Estimating the scope for savings

We estimate the overall savings progressively in three phases, gradually relaxing the assumptions about the alternatives open to a GP for managing clinical conditions of patients presenting with conditions relating to the GMG specialty. The three phases in terms of the assumptions underpinning them are respectively as follows:

There is **no** scope for switching between inpatient, outpatient referrals and drugs at the GMG specialty level;

There **is** scope for switching between inpatient, outpatient referrals and drugs at the GMG specialty level but unit costs are fixed;

There **is** scope for switching between inpatient, outpatient referrals and drugs at the GMG specialty level and unit costs can be varied.

Phase 1: Assuming there is no scope for switching between inpatient, outpatient referrals and drugs pertaining to the GMG specialty.

Under this assumption the variables in Table 1 are used in volume terms and costs of referrals and drugs are not taken into account. The assumption is that the only option open to the Practice for savings is to reduce the volumes of referrals and drugs to levels more compatible with those of benchmark Practices facing a similar mix of health conditions of inpatients and outpatients and number of list patients aged over 48 years. The generic DEA model compatible with this assumption is Model (1). Under the foregoing assumption we identify aggregate cost savings across the 75 Practices on the order of 3% of observed total expenditure.

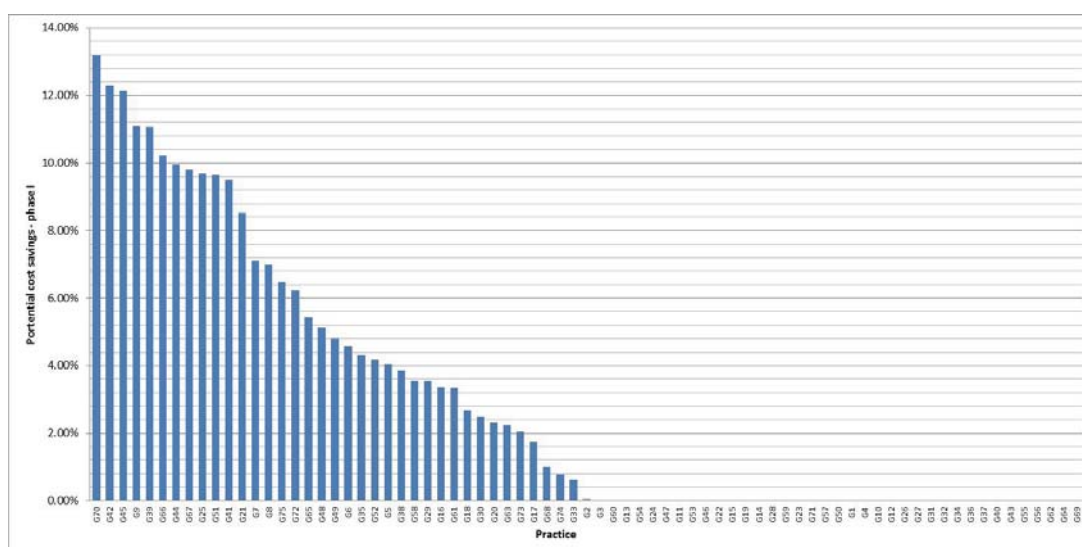


Figure 2: Potential savings as percentage of aggregate actual expenditure

As can be seen in Figure 2 there is a substantial difference between Practices on their potential for savings. More than half the Practices have no scope for reducing the volumes of referrals and prescriptions but some 12 Practices have potential to save in excess of 8% of their costs by simply reducing the volume of referrals and drugs. Overall, however, we find the potential for reducing referrals to the GMG specialty and prescriptions in pure volume terms is relatively low.

Phase 2: Assuming there is scope for switching between inpatient, outpatient referrals and drugs pertaining to the GMG specialty maintaining unit costs

Under this assumption the DEA model is configured so that we estimate the minimum aggregate referrals and drug costs a Practice could have incurred in relation to the GMG specialty, controlling for the diagnoses and list population above 48 years of age that drive those costs. Thus we relax the Phase 1 assumption that these drivers pretty well determine the mix of treatment pathways that are needed in the GMG specialty at the Practice. Rather, through the model we enquire whether a given Practice could have adopted some combination of referrals and drugs modelled on benchmarks, which would minimise its aggregate cost of referrals and drugs pertaining to the GMG specialty, given the health profile of its patients and the unit costs (or prices) it faces.

The model solved in this phase is the classical cost efficiency DEA model (2). We identify during this phase considerable scope for savings (see Table 4). Clearly most cost savings can be obtained from switching between inpatient, outpatient and drug pathways than from pro rata changes of their volumes.

Table 4: Scope for savings by switching treatment pathways at observed unit costs of referrals and prescriptions

	% of actual GMG specialty referrals and drugs expenditure
Total phases 1 and 2	10.97%
Phase 1	2.59%
Phase 2	8.38%

Table 4 clearly shows the enhanced scope for savings that opens up when we allow for switching between treatment pathways rather than maintaining the mix of referral and drug

pathways and simply lowering their volumes. Figure 3 shows the range in the potential for savings in the GMG specialty in aggregate through volume decreases (Phase 1) and referral and drugs volume mix adjustments (Phase 2).

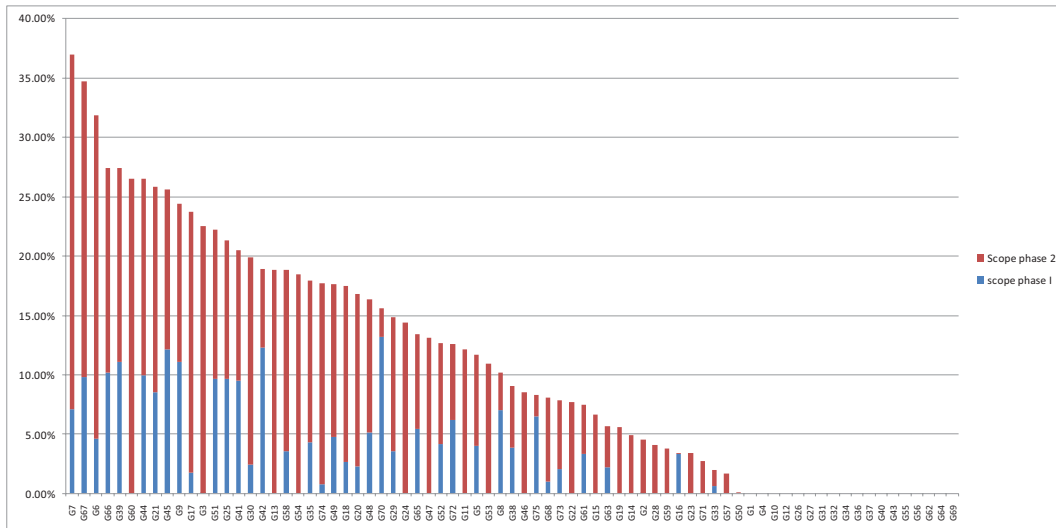


Figure 3: Scope for savings as percentage of actual aggregate referrals and prescriptions expenditure

The scope for savings, as we might expect, increases when we allow for adjustments to the mix of referrals and drugs. The median practice can now save about 10% of its aggregate cost, and 20% of the practices can save above 20% of their aggregate cost. That is, certain Practices may prescribe a higher number of doses per prescription than others for certain drugs. Information at that level of detail was not available to us.

It is noteworthy that in Table 5 for no Practice do we find that it would be cost-effective to increase the volume of drug prescriptions.

Table 5 summarises the adjustments to the volume of inpatient (VI) and outpatient (VO) referrals and drug volumes (VD) that Practices need to make to effect the savings identified in Phase 2. The columns show the ratio of *Target* volumes (named TVI, TVO, TVD) to observed volumes, the former having been identified via model (2). For example, 30 practices need to raise their inpatient referrals, 18 to keep them as they are, and 27 to lower them. The average change is a rise in inpatient referrals of 1.4%, and a reduction in outpatient referrals to 97.4% of observed levels. The main overall change the table suggests is a drop of over 18% in prescriptions, though it should be noted that this does not take into account the differences in volumes of drug prescriptions by Practice. That is, certain Practices may prescribe a higher

number of doses per prescription than others for certain drugs. Information at that level of detail was not available to us.

It is noteworthy that in Table 5 for no Practice do we find that it would be cost-effective to increase the volume of drug prescriptions.

Table 5: Number of Practices needing to change referral and drug volumes to minimise costs at observed unit costs

	TVI/VI	TVO/VO	TVD/VD
# >1	30	23	0
# =1	18	18	18
# <1	27	34	57
Average	1.014	0.974	0.817
Maximum	1.360	1.346	1.000
Minimum	0.873	0.481	0.486

Phase 3: Estimating the scope for savings when we permit adjustments to volumes and mix of referrals and drugs, and to their unit costs

This phase represents the ultimate relaxation of assumptions. We assume that for managing medical conditions of the GMG patients not only can Practices opt for a mix of referrals and prescriptions by volume that best suit their unit costs, but that those unit costs themselves may be capable of adjustment. The model solved at this stage was model (3). This model applies the traditional DEA assumption that a virtual Practice can be constructed as an average of those observed, with the only difference being that the average is over the optimal (for observed prices) inpatient, outpatient and drug cost components as defined in that model rather than the volumes of those variables, used traditionally. This in turn implies that controlling for disease weight as proxied in the outputs of the model a Practice should in principle be capable of attaining the optimal mix of outputs and cost components of other benchmark practices, and the unit costs implicit in those cost components.

Under these relaxed assumptions we identify additional cost savings in the GMG specialty across the 75 practices amounting to nearly 7% of observed expenditure. See Table 6.

Table 6: Scope for savings by switching treatment pathways and adjusting unit costs of referrals and prescriptions in the GMG Specialty

	% of actual GMG specialty referrals and drugs expenditure
Total across 3 phases	17.39%
Phases 1 and 2	10.97%
Phase 3	6.42%

Figure 4 shows the range of the potential for savings in the GMG specialty at Practice level. The scope for savings, as we might expect, increases when we allow for adjustments not only to the mix of referrals and drugs but to joint adjustments of unit costs and volumes. The median practice can now save about 18% of its aggregate drugs and referrals costs compared to 10% when only the level and mix of volumes could be adjusted.

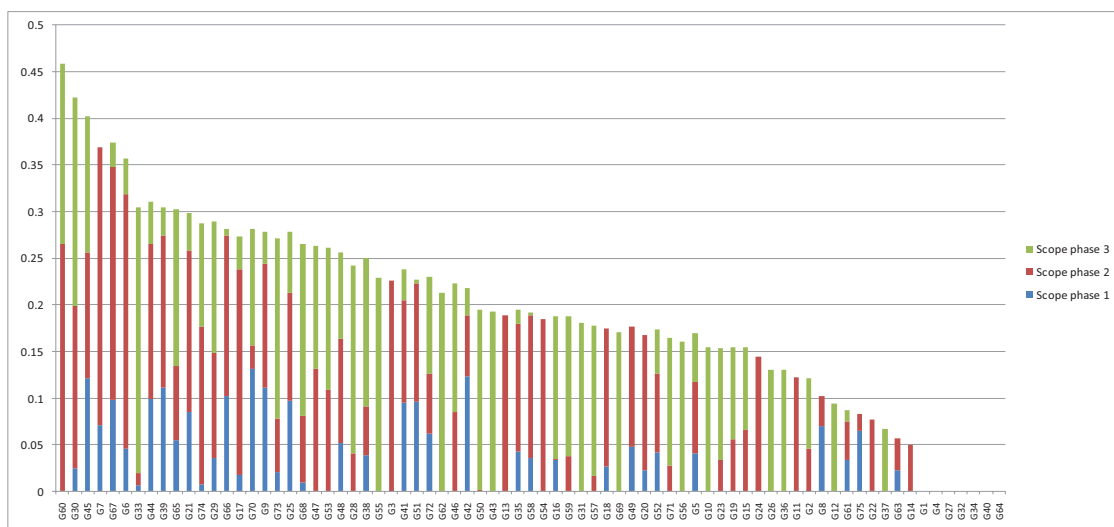


Figure 4: Scope for savings as percent of actual aggregate referrals and prescriptions expenditure - unit cost adjustments permitted

In Figure 4 we can see that Practices show different profiles for attaining cost savings. For example Practice G3 is only inefficient in allocative terms, meaning that the only way to effect the savings identified for this Practice is by changing its mix of inputs (referrals and drugs). On the other hand, Practice G55 is only inefficient in terms of unit costs, meaning that the only way to achieve the cost savings for this Practice is through a change in the unit costs it ‘pays’. Below in Figure 5 we can see two radar graphs that compare these two units with some of their main benchmarks as identified by the respective DEA model. In these graphs the input-output values for the inefficient practice (G3 and G55 respectively) have been normalized to 1.

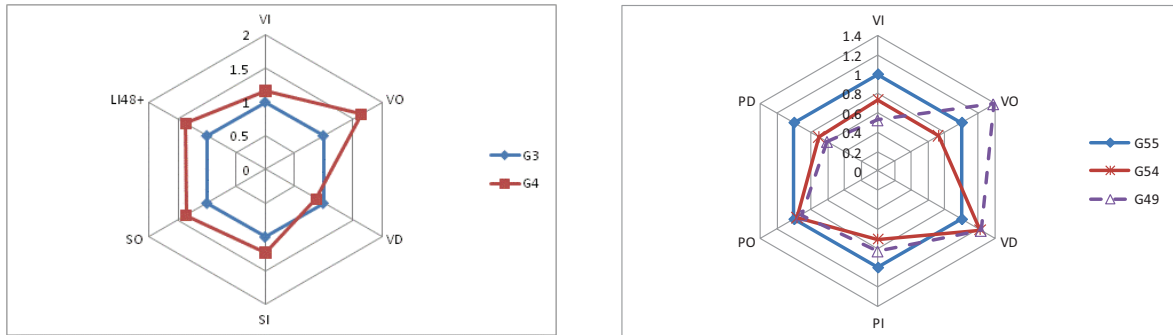


Figure 5: Practices G3 and G55 Compared with some benchmarks

Practice G3 is only inefficient in allocative terms. That is, the volumes of referrals and drugs cannot all reduce pro rata, but their mix (ratio to each other) can be altered to minimize their aggregate cost. Observed average prices for G3 are 1,835 for inpatient referrals, 104.4 for outpatient referrals, and 8.1 for drugs prescriptions, in pounds sterling. In general the minimization of costs is mainly achievable through changing inpatient referrals, as across all Practices unit costs are much higher for these referrals in comparison to outpatient referrals. In the left panel of Figure 5 we can see that practice G4 (the cost benchmark of G3) shows higher outputs (about 1.3 times more severe health conditions of patients than G3) and simultaneously shows a lower volume of drug prescriptions, a volume of outpatient referrals that is about 1.6 times of that of G3, and a volume of inpatient referrals that is only about 1.2 times of that of G3. Therefore Practice G4 shows a balance of referrals that favours the less expensive referrals in relation to the most expensive ones. Therefore model (2) identifies that the volume of outpatient referrals of G3 should increase by 15%, and that of inpatient referrals and of drug prescriptions of G3 should decrease by 3.5% and 36%, respectively. Note that the price of prescriptions does not play a major role in determining drug prescription targets, as drug prescriptions unit costs are very low, and indeed minimization of this component of cost is mainly achieved through volume changes (as will be clearer later). Therefore actions at Practice G3 to increase outpatient referrals and decrease drug prescriptions would reduce overall costs.

Practice G55 is only inefficient in input price terms (unit costs of referrals and drugs). From model (3) Practices G54 and G49 are two of its main benchmarks in terms of input prices. As we can see on the right panel of Figure 5 the two benchmark practices show similar unit costs to G55 on outpatient referrals (PO), but much lower unit costs of drugs (PD) and inpatient referrals (PI). The lower prices cannot be explained by higher volumes at the benchmarks, since both benchmark practices show lower volumes of inpatient referrals than Practice G55. Note that the price benchmark Practices are not necessarily efficient on the other dimensions,

so the differing volumes may reveal inefficiency of these price benchmarks on the other dimensions.

5.2. Overview of the changes Practices need to make to realize the savings identified

It is instructive to look at the detail of changes in referrals and prescriptions that each practice would need to make in order to achieve the cost savings identified in the previous section. We look at these changes under two alternative assumptions: that unit costs of referrals and prescriptions cannot be changed and that they can be changed.

Case 1: Changes needed when unit costs of referrals and prescriptions remain as observed

When we consider the potential for Practices to minimize aggregate expenditure by exploiting the potential to change the volume and mix of treatment pathways then Figure 6 summarises the changes needed across the 75 Practices assessed. TVI, TVO and TVD are the target volumes of inpatient, outpatient referrals and prescriptions respectively to minimise total costs under current unit prices. In Figure 6 TVI/VI, TVO/VO and TVD/VD are the target volumes expressed as fractions of the actual levels of inpatient and outpatient referrals and prescriptions, respectively.

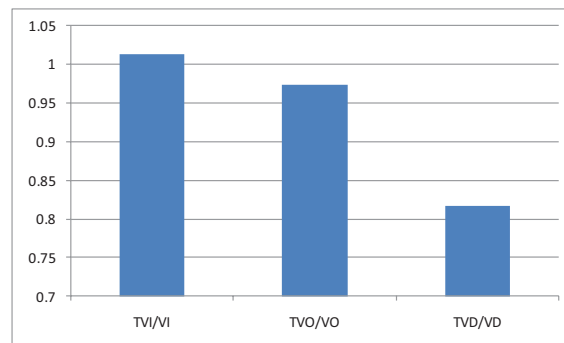


Figure 6: Referral and prescription mix changes to minimise aggregate costs under existing unit costs

As can be seen in Figure 6 the main change needed is the reduction in the volume of prescriptions down to about 81% of current volumes. Reductions in outpatient referrals needed are also considerable down to about 97% of current levels. The volume of inpatient referrals can increase slightly by about 1% if these savings in volumes of outpatients and drugs are to be attained.

Case 2: Changes needed to achieve the estimated savings when unit costs of referrals and prescriptions need not remain as observed

As we saw earlier, when we assume that through a judicious choice of health providers and treatment pathways a practice may be able to change both the mix of referrals and prescriptions and their unit costs, the total sum of savings attainable in the GMG specialty is of the order of 18% of total expenditure on referrals and prescriptions across all 75 practices.

Figure 7 summarises the mix of changes needed to realize the savings identified both in terms of volumes and in terms of unit costs. The left panel in Figure 7 identifies the scope for savings in volume terms and is therefore equivalent to the graph in Figure 6.

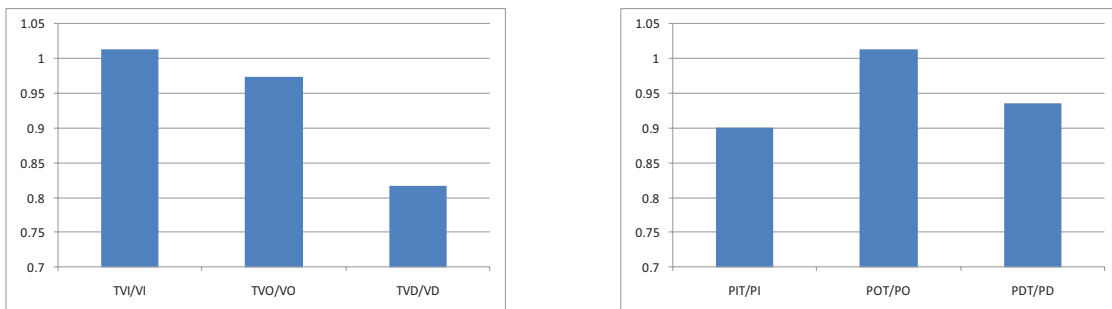


Figure 7: Referral and prescription mix and unit cost changes to minimise aggregate costs

The changes to unit costs needed are summarised in the right panel of Figure 7. Looking at that panel the ratio PIT/PI is that of the target unit cost for inpatient referrals to its actual level. Mean values of PIT and PI are used. The remaining two ratios in the right panel of Figure 7 are defined in a similar manner. The panel suggests that there is scope for substantial reductions in inpatient unit costs in the GMG specialty of the order of 10% on average. Similarly, there is scope for about 7% reduction in unit cost of drugs while outpatient costs on the other hand can rise by about 1.3% on average to compensate for the inpatient and prescription unit cost reductions.

We conclude this section by summarising in Table 7 the components of cost savings attributable to each type of volume and/or unit cost adjustment.

Table 7: Potential saving components by type of volume or unit cost adjustment

Inpatient volume adjustment	-1.05
Outpatient volume adjustment	0.49
Drugs Volume adjustments	63.65
Inpatient unit cost adjustments	17.80
Outpatient unit cost adjustments	0.22
Drugs unit cost adjustments	18.89

Clearly the bulk of cost savings can be achieved through reducing the volume of drug prescriptions. Two further substantial components of savings of 17%-19% of total potential savings are possible through reducing the unit cost of inpatient referrals and drugs. The remaining components of cost savings can be considered residual.

6. Conclusion.

In this paper we have investigated the scope for health care cost savings at 75 GP Practices in relation to patients presenting with conditions that lead to treatment in the GMG specialty. The analysis followed that in Thanassoulis et al. (2012) where overall expenditure on health care across all medical specialties was investigated in aggregate.

The analysis at the specialty level reported here was undertaken, in order to control more accurately for the medical conditions of those receiving medical care and thereby incurring the costs being modelled. An additional benefit of an analysis at Specialty level is that we gain better insight into the cost-efficient patient management processes as we identify such processes at specialty level thus avoiding overlooking potential efficiencies through averaging across efficient and inefficient component costs of a Practice.

Two of the key contributions of this paper, apart from the identification of cost savings possibilities, are both the process of identifying input-output variables for assessments at specialty level and the variables themselves. The process of identifying input-output variables is important because medical data at patient level are strictly confidential and so the analysis needs to identify medical conditions to control for, yet without straightforward access to the diagnoses of those conditions. We exploited the fact that for those patients that had had in the time period under considerations (2007/8) at least one inpatient appointment we had at anonymised patient level the diagnosis. However, this information was not available for those that had an outpatient appointment. Using data from the inpatient record and combining it with more general anonymised patient level data on age, gender and deprivation level of geographical location we identified proxy variables that explained on the order of 90% of the variation across Practices of volumes and costs of inpatient and outpatient referrals and over 80% of the variation in volume and costs of drugs used. These variables enabled health care costs to be set against a good representation of the disease weight driving those costs.

Our analysis suggests that for the 75 Practices analysed the largest percentage of savings at the GMG specialty comes from changes in the mix of referrals and drug prescriptions. In particular, the largest components of potential savings are available through control of the use of drugs. The single largest component, accounting for some 63.65% of total potential savings, relates to a potential reduction in the volume of drug prescriptions. The second largest component, accounting for just under 19% of total potential savings is due to a reduction in the unit cost of prescriptions. Finally considerable savings of the order of 17% are available through better monitoring of inpatient unit costs. There is, on the face of it, the counter-intuitive finding the foregoing reductions appear to be contingent on a marginal rise of inpatient referrals. This suggests that in some cases spells of inpatient episodes can reduce perhaps long term drug use. This finding can be further investigated by analyzing longitudinal data at patient level. Savings from outpatient referrals in terms of volume or unit costs are marginal, suggesting that the biggest sources of inefficiency are related to inpatient unit costs and above all volume of drug prescriptions.

References

- Amado, C.A.F. and Santos, S.P. (2009) Challenges for performance assessment and improvement in primary health care: The case of the Portuguese health centers, *Health Policy* 91, 43-56.
- Amado C.A.F. and Dyson R.G. (2009). Exploring the use of DEA for formative evaluation in primary diabetes care: An application to compare English practices. *Journal of the Operational Research Society* 60, 1469-1482.
- Amado, C.A.F, and Dyson, R.G. (2008) On comparing the performance of primary care providers, *European Journal of Operational Research* 185, 915-932.
- Barnett, M.L. , Keating, N. L., Christakis, N. A., O'Malley, A. J., and Landon, B. E. (2011) Reasons for Choice of Referral Physician Among Primary Care and Specialist Physicians, *Journal of General Internal Medicine* 27(5):506–12
- Chilingerian, J. A. and Sherman, H.D. (2004) Health care applications in Cooper, Seiford and Zhu (Eds), *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, 481-537.
- Dervaux, B. , Leleu, H. , Minvielle, E., Valdmanis, V., Aegerter, P., and Guidet, B. (2009), Performance of French intensive care units: A directional distance function approach at the patient level, *Int. J. Production Economics* 120, 585–594.
- Dismuke, C. E. and Sena, V. (2001) Is there a Trade-Off between Quality and Productivity? The Case of Diagnostic Technologies in Portugal, *Annals of Operations Research* 107, 101–116,
- Donabedian A. (1980). *The Definition of Quality and Approaches to its Assessment*. Ann Arbor, Michigan: Health Administration Press.
- Hollingsworth, B. (2008), The measurement of efficiency and productivity of health care delivery, *Health Economics* 17, 1107-1128.
- Kristensen, T., Laudicella, M. Ejersted, C. and Street, A. (2010). *Health Economics Cost variation in*

diabetes care delivered in English hospitals, *Diabetic Medicine*, 27, 949–957

- Linna, M., Nordblad, A. and Koivu, M. (2003) Technical and cost efficiency of oral health care provision in Finnish health centres, *Social Science and Medicine*, 56, 343-353.
- Luoma, K., Jarvio, M-L., Suoniemi, I., and Hjerpe, R. (1996) Financial incentives and productive efficiency in Finnish health centres, *Health Economics* 5, 435-445.
- Laudicella, M. Olsen, K. R. and Street, A. (2010). Examining cost variation across hospital departments: a two-stage multi-level approach using patient-level data *Social Science & Medicine* 71, 1872-1881
- Ozcan, Y.A. (1998) Physician benchmarking: measuring variation in practice behavior in treatment of otitis media, *Health Care Management Science* 1, 5-17.
- Puig-Junoy, J., 1998. Technical efficiency in the clinical management of critically ill patients. *Health Economics* 7, 263–277.
- Puig-Junoy, J. and Ortún, V. (2004) Cost efficiency in primary care contracting: A Stochastic frontier cost function approach, *Health Economics* 13, 1149-1165.
- Sørensen, T.H., Olsen, K.R., and Gyrd-Hansen, D. (2009) Differences in general practice initiated expenditures across Danish local health authorities – a multilevel analysis, *Health Policy* 92, 35-42.
- Staat, M. (2003), The efficiency of treatment strategies of general practitioners, *European Journal of Health Economics* 4, 232-238.
- Staat, M. (2011), Estimating the efficiency of general practitioners controlling for case mix and outlier effects, *Empirical Economics* 40, 321-342.
- Szczepura, A., Davies, C., Fletcher, J. and Boussofiane, A.(1993) Efficiency and effectiveness in general practice, *Journal of Management in Medicine* 7/5, 36-47.
- Thanassoulis, E., Boussofiane, A., and Dyson, R. G. (1995) Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. *European Journal of Operational Research* 80, 588.
- Thanassoulis, E., Portela, M. C.A.S. and Despic, O. (2008) DEA - The mathematical programming approach to efficiency analysis. In Fried, H.O. Lovell, C.A.K., and Schmidt, S.S. editors, *The Measurement of Productive Efficiency and Productivity Growth*, Oxford University Press.
- Thanassoulis, E., Portela, M.C.A.S. and Graveney, M. (2012) Estimating the scope for savings in referrals and drug prescription costs in the General Practice units of a UK Primary Care Trust, *European Journal of Operational Research* 221, 432-444.
- Tone, K, and Tsutsui M.(2007) Decomposition of cost efficiency and its application to Japanese-US utility comparisons, *Socio-Economic Planning Sciences* 41, 91-106.
- Zavras, A.I., Tsakos, G., Economou, C., and Kyriopoulos, J. (2002) Using DEA to Evaluate Efficiency and formulate Policy within a Greek national primary health care network, *Journal of Medical Systems* 26/4, 285-292.