DÉPARTEMENT DE SCIENCE ÉCONOMIQUE
DEPARTMENT OF ECONOMICS

CAHIERS DE RECHERCHE / WORKING PAPERS

 # 0802E

Comparative Efficiency Assessment of Primary Care Models Using Data Envelopment Analysis[#]

by Olga Milliken [a,*], Rose Anne Devlin [a], Vicky Barham [a], William Hogg [b,c,d],

Simone Dahrouge [b] and Grant Russell [b,c]

March 2008

ISSN: 0225-3860

CP 450 SUCC. A
OTTAWA (ONTARIO)
CANADA K1N 6N5

P.O. BOX 450 STN. A
OTTAWA, ONTARIO
CANADA K1N 6N5

uOttawa

Faculté des sciences sociales
Faculty of Social Sciences

[*] *Corresponding author. E-mail address: olavr099@uottawa.ca*
[a] *Department of Economics, University of Ottawa, Canada*
[b] *C.T. Lamont Primary Health Care Research Centre, Élisabeth Bruyère Research Institute, Ottawa*
[c] *Department of Family Medicine, University of Ottawa*
[d] *Institute of Population Health, University of Ottawa, Ottawa*

[#] Funding for this research was provided by the Ontario Ministry of Health and Long-Term Care Primary Health Care Transition Fund. The views expressed in this report are the views of the authors and do not necessarily reflect those of the Ontario Ministry of Health and Long-Term Care.
The authors gratefully acknowledge the work of Meltem Tuna, PhD, for calculating the input scores and data cleaning. We also thank Doug Angus, MA, who was a co-investigator of the larger study in which this work was nested, provided direction for the paper.

Abstract

This paper compares the productive efficiencies of four models of primary care service delivery in Ontario, Canada, using the data envelopment analysis (DEA) method. Particular care is taken to include quality of service as part of our output measure. The influence of the delivery model on productive efficiency is disentangled from patient characteristics using regression analysis. Significant differences are found in the efficiency scores across models and within each model. In general, the fee-for-service arrangement ranks the highest and the community-health-centre model the lowest in efficiency scoring. The reliance of our input measures on costs and number of patients, clearly favours the fee-for-service model. Patient characteristics contribute little to explaining differences in the efficiency ranking across the models.

***Keywords**: Productive Efficiency; DEA; Primary Health Care*
***JEL classification**: I12, I19*

Résumé
Cet article compare l’efficience productive dans quatre modèles de prestation de soins primaires en Ontario, au Canada, en utilisant la méthodologie du DEA (Data Envelopement Analysis). Une attention particulière a été portée sur l’inclusion la qualité du service de soin santé dans la mesure de l’extrant (output). L'influence du modèle de prestation sur l'efficience productive a été séparée des caractéristiques du patient en utilisant une analyse de régression. Des différences significatives ont été trouvées dans l'efficience entre modèles et à l’intérieur de chaque modèle. En général, les arrangements avec services-payants arrivent en tête, alors que les modèles de centre de santé communautaire performent le moins en termes d’efficience. Le recours au coût et au nombre de patients comme mesure de l’intrant est nettement favorable au modèle avec services-payants. Les caractéristiques du patient contribuent peu à expliquer les différences dans le classement de l'efficience entre les modèles.

***Mots clé:** Efficience Productive; DEA; Soin Primaires*
***Classification JEL:** I12, I19*

Introduction

This paper undertakes an efficiency comparison of four distinct models of primary health care service delivery in Ontario using the methodology of Data Envelopment Analysis (DEA): fee-for-service practices including Family Health Groups (FFS/FHGs), health service organizations (HSOs), family health networks (FHN)s, and community health centres (CHCs). The analysis draws on data collected between 2005 and 2006 as part of a multidisciplinary project funded by the Ontario Ministry of Health and Long-Term Care entitled "Comparison of Models of Primary Health Care in Ontario". A comprehensive study of the performance of these models has been lacking (Muldoon, Rowan, Geneau, Hogg & Coulson, 2006).

Previous studies of efficiency in the provision of primary care have employed the number of visits and tests as an intermediate output measure (Huang & McLaughlin, 1989; Andes, Metzger, Kralewski & Gans, 2002; Linna, Nordblad & Koivu, 2003; Kirigia, Emrouznejad, Sambo, Munguti & Liambila, 2004; Rosenman & Friesner, 2004). Relatively few studies (Salinas-Jimenez & Smith, 1996; and Wagner, Shimshak & Novak, 2003) have incorporated measures of quality of care into the measurement of efficiency. In contrast, this study integrates several performance indicators measuring health service delivery and quality of care, and evaluates their cumulative influence on efficiency of primary care practices.

In Ontario, the longest standing model with the largest number of participating physicians is the fee-for-service model in which physicians are paid on a per-service basis with relatively flat fee schedule across services. Since physician income increases with the quantity of services provided under this payment model, and because physicians are better able to evaluate the patient's health care needs than are patients themselves, some claim that FFS physicians will be volume-driven (Evans, 1974; Pauly, 1980; Arrow, 1986; McGuire, 2000). The Cochrane Review (Gosden, Forland, Kristiansen, Sutton, Leese, Giuffrida et al. 2000) of the empirical literature concludes that fee-for-service, as compared to capitation, results in more primary care visit contacts, and more diagnostic and curative services, but fewer hospital referrals and fewer repeat prescriptions. Family Health Groups (FHGs) are a new model of service delivery in Ontario that is similar to pure fee-for-service. In this study traditional fee-for-service physicians and FHG physicians are lumped together and referred to as FFS. Currently there are about 7,439 pure FFS physicians in Ontario servicing some 9.2 million patients and about 2536 FHG physicians servicing 3.7 million patients (Coulson, 2005; Muldoon, Rowan, Geneau, Hogg & Coulson, 2006).

It has long been argued (Ellis & McGuire, 1986; Ma, 1994; Newhouse, 1996) that when patients are rostered, and physicians paid a flat fee per patient in their practice, doctors have powerful incentives to provide services in a cost-effective manner. Equally importantly, capitation-based payment schemes provide cost predictability for public health authorities. This view underlay the introduction of health service organizations in Ontario in 1975 (Gillett, Hutchison and Birch, 2001). HSO physicians register their patients and are paid a monthly capitation fee, differing with age and gender. The fee is partially clawed back if a rostered patient is provided with primary care service outside the HSO. There are 49 HSOs in Ontario with 160 physicians serving 255,000 patients. A single HSO can have multiple geographically distinct sites of operation.

A weakness with capitation-based physician remuneration is that this encourages physicians to engage in cream-skimming, that is, to roster only those patients whom they expect to be relatively healthy, or to offload costs to the non-capitated sector, e.g., by referring patients to specialists (Newhouse, 1984; Pauly, 1984; Dranove, 1987; Allen & Gertler, 1991; Ellis, 1998; Mulligan, 2002). Mixed payment mechanisms – combining some element of fee-for-service and some element of capitation-based payment – have therefore been proposed as a means of providing physicians with balanced incentives to provide quality care to patients with a variety of risk profiles, while still providing reasonable control of the overall cost of delivery of primary health care (Ellis & McGuire, 1986, 1993; Ma, 1994, Ellis 1998; Chalkey and Malcomson, 1998; Jelovac, 2001; Jack, 2005). In Ontario, Family Health Networks (FHNs) were introduced in 2001 to provide comprehensive care for their patients 24 hours a day, seven days a week. Physicians working in these networks are paid under a blended scheme, combining a capitation fee for rostered patients with specific bonuses or fee-for-service payments which encourage doctors to provide preventive services and services additional to those specified in the roster agreement. FHN physicians receive a bonus for each new patient rostered and fee-for-service payments at a rate of 10% of the provincial schedule for most services.

Whereas each of the previous three models positively link physician income to some measure of service provision, the fourth model pays physicians a straight salary. Community Health Centres (CHCs) are community-oriented, and with a primary mandate to address disadvantaged populations' needs. Introduced widely across Ontario in the 1980s, these interdisciplinary teams are one-stop health care shops for patients including physicians, nurse practitioners, nurses, physiotherapists, chiropodists, social workers and other health and community health professionals. Economic theory (Holmstrom & Milgrom, 1991; Laffont & Tirole, 1993) suggests that salaried physicians can be expected to see a low volume of patients (but will therefore provide higher-quality care). Empirical studies included in the Cochrane Review (Gosden, Forland, Kristiansen, Sutton, Leese, Giuffrida et al. 2000), as well Sorensen & Grytten (2003) (focusing on Norway), and Devlin & Sarma (2007) (focusing on Canada) show that salary payments lower the volume of services provided. However, these authors do not find that an increased number of visits/procedures is associated with better patient care. Yalnizyan &

Macdonald (2005) report on several studies supporting the view that CHCs provide higher quality care than FFS.

Methodology

DEA is a nonparametric linear programming technique (Farrell, 1957; Charnes, Cooper & Rhodes, 1978). It measures relative efficiencies, rather than absolute efficiencies, as it judges performance relative to others in the sample, and not against a theoretically constructed absolute measure of efficiency. The DEA approach permits the evaluation of practice sites' performance by a single efficiency score which takes account of the variety of different outputs produced, the quality of those outputs, and the inputs used. The definition of efficiency underlying DEA is illustrated in Figure1, for a production technology consisting of two outputs and one input.

Insert Figure 1 Here

The area south-west of GCDE in Figure1 indicates all possible combinations of output 1 and 2 per unit of the input. Empirically constructed on a particular sample of practice sites, the efficiency frontier GCDE represents points with maximal output combinations given one unit of input. Practice sites C and D are located on the efficiency frontier GCDE, hence they receive a score of 1. Practice site A does not lie on the frontier and, thus, is inefficient relative to C and D. The inefficiency of site A is measured by the ratio OA/OB, and is less than 1.

DEA allows several input and output variables, with different units of measurement, to be combined. Thus, distinct output measures, such as performance indicators (for quality of care and health service delivery), service volume and intensity, can be incorporated into the analysis. Additionally, the DEA method does not require weights to be assigned *a priori* to each input and output, instead weights are assigned by the DEA program to present every practice site in the best possible light against the others. Moreover, DEA does not require the specification of a production or a cost function. Finally, unlike parametric estimation procedures, the technique is only moderately vulnerable to the sample size. The DEA method has been extensively used for estimating efficiencies (see Hollingsworth, 2003, for review), and to the best of our knowledge, there are no theoretical grounds to prefer any one parametric or non-parametric method of efficiency measurement over another in the health care sector (Giuffrida & Gravelle, 2001).

However, DEA is not without drawbacks. Because the efficiency frontier is constructed from the sample data, it is vulnerable to data inconsistencies, outliers, or possible errors. Moreover, in the absence of time-series data for the same sample, the non-parametric nature of the method does not permit a disentangling of real efficiency from random fluctuations. In addition, the choice of inputs and outputs is mostly *ad hoc*; researchers have to rely largely on their understanding of the production processes of the industry in the study.

The output from the DEA is an array of efficiency scores which may be used to rank the practices within each model and across models; for the most part, the analysis of this paper concerns itself with this latter ranking. Importantly, these scores depend critically upon the variables used to measure inputs into production and the outputs. In the context of this paper, efficiency scores may vary because of variation across practices in the mix of patients treated. Specifically, practices which specialize in high-need patients may operate differently than otherwise. If so, then simply measuring inputs and outputs without regard to the patients being treated will yield "efficiency" scores that are not very meaningful from a policy perspective. To address this problem, we use a two-stage procedure. First, we calculate efficiency scores using the DEA method and use these scores to compare the performance across models. Second, we examine whether a variety of patient characteristics can help to explain inter-model efficiency score variation.

Data and the choice of input and output variables

The data used in this study were collected over the period of 2005-2006. 137 practice sites in Ontario were chosen randomly, stratified by model type. Information on the recruitment and representativeness of the practices is presented elsewhere (Hogg et al., 2008). This paper is based on an analysis of 109 practice sites from the four models under study: 19 CHCs, 27 FHN sites, 32 FFS sites and 31 HSO sites. 28 practice sites were excluded from the total sample size sites due to the lack of expenditure and/or patient data or inconsistencies in the reported variables.

As the DEA approach requires that primary care practice sites be as homogeneous as possible, we consider only the clinical primary care component of the services provided. For the purpose of this study, clinical primary care is defined as one-on-one encounters with a physician, a nurse practitioner, a registered practical nurse, a nurse or a nursing assistant for the purpose of clinical medical care. While clinical primary care services are virtually the only services provided in HSOs, FHNs and FFS, CHCs provide additional services. The most challenging task was to disentangle the cost of the clinical services component from the additional services provided by CHCs, such as group activities, outreach community services, counselling and education. Each CHC was requested to provide the personnel data and operations expenditure data attributed directly to clinical primary care services. Capital costs were distributed to the clinical primary care component proportionally to the office space occupied by the clinical primary care service unit; while overhead costs, including administration expenditures, were distributed proportionately to the clinical primary care budget.

The primary care practice sites in this study consist mostly of family physicians with few specialists present. Both types of physicians are addressed as physicians in this paper. The term ‘provider’ refers to both physicians and nurse practitioners (NPs). To the best of our knowledge no consensus is found in the health care literature with regard to the substitutability of doctors and nurse practitioners (Laurant M. et al., 2005). The DEA results presented here are calculated under the assumption that one NP equals one-half of a physician; although the rankings of the different practice sites change slightly under alternative specifications of this substitutability ratio, the collective ranking of the models does not.

The number of patients in the practice site relies on numbers that are reported by the practice site, i.e., the number of patients seen at least once within the past year. Practice site annual costs are comprised of physician incomes, the salaries of medical and administrative personnel, operating costs, and maintenance and capital costs (including rent and depreciation of capital assets). Output measures include the average number of visits per patient in the practice site, and performance indicators measuring technical quality of care and health service delivery. These indicators are based on the best practices reported in the health care literature (Shi, Starfield & Xu, 2001) and on established guidelines; they have been calculated using the data extracted from patient’s charts and patient surveys. The calculation procedure is available upon request. Table 1 provides the definition of the variables used in the analyses.

INSERT TABLE 1 HERE

Table 2 provides some descriptive statistics of the output and the input variables. Notice that the figures in table 2 are ***not*** adjusted for patient characteristics which can potentially influence outcomes in primary care service delivery; these are adjusted at the second stage of the analysis.

INSERT TABLE 2 HERE

It is interesting to observe from table 2 that CHCs have the highest means for four out of the seven output variables when compared to the other three models. The Wilcoxon-Mann-Whitney test confirms that the CHCs are statistically different from the other models with respect to health promotion, chronic disease management, comprehensiveness and visits per patient. This suggests that the CHC model performs well in terms of quality of care and performance measures, and that CHC providers see their patients as often as do providers in FFS sites.

The HSO model is statistically different from every other model with respect to both continuity and access to primary care services scores and has the highest average for these variables. FHNs and CHCs lead in delivery of preventive services. The fact that FHNs are performing well with respect to this indicator should not be surprising given that FHN providers receive financial bonuses, which vary with the achieved level of preventive services.

Practice site costs per provider, practice site costs per patient, and provider-patient ratio are input variables used in the DEA (table 2). In general, the mean values for these variables are the lowest in the FFS model and the highest for the CHC model. These differences are statistically significant.

DEA Scenarios and Results

The practice output comprises of seven variables described in table 2. Three different input scenarios are estimated; the first scenario uses the cost per provider as the input measure; in scenario 2 input is the cost per patient and the final scenario uses the provider-patient ratio as input. We examine the sensitivity of the model's ranking to the exclusion or inclusion of physician income in the cost data.

Table 3 synthesizes the efficiency scores by presenting the number of practice sites found at decile intervals from 0 to 1, where 1 is the most efficient practice site. There is clearly significant variation across models and within a model; some differences are discernable for the input scenarios.

INSERT TABLE 3 HERE

Input Scenario 1: Total Cost/Provider

In scenario 1, with the total cost per provider as the input measure, the mean of the resulting efficiency scores is 0.6 and the standard deviation is 0.17. The distribution of these efficiency scores by quartiles is shown in figure 2. The first quartile shows the highest efficiency scores and the fourth quartile the lowest; each bar indicates the percentage of the total number of practice sites for each model in a respective quartile. Observe that the FFS model has the highest representation in the first and the second quartile compared to the other models. The CHC model has over 40% of its practice sites located in the lowest quartile, while another 37% of its sites are in the third quartile. The efficiency scores of both the HSO and FHN sites are more evenly distributed across all quartiles in comparison to FFS and CHCs. However, the FHN model appears to perform better in terms of efficiency scoring than the HSO model: FHN sites have higher representation in the 1^{st} and the 2^{nd} quartile, whereas HSO sites – in the 3^{rd} and the 4^{th} quartile. The explanation for the models' ranking is that the correlation between the total cost per provider variable and the resulting efficiency score variable is very high - 87%. This means that the input variable is driving the results: CHCs have, on average, the highest total cost per provider and FFS sites have the lowest total cost per provider.

We also investigate whether including physician incomes in the total cost influences the efficiency score ranking. As only about one half of physicians in each practice site were approached to self-report their annual before-taxes, there was some concern regarding the accuracy of the physician income data for the FFS, FHN and the HSO models. Given the potential biases, we excluded physicians' income from the cost data. This worsens the ranking of the CHC practices, suggesting that the low CHC rankings are driven by costs other than physician salaries.

Insert Figure 2 here

Another potentially useful way of presenting the results from the DEA analysis is to examine the characteristics of the 10% highest-ranked and 10% lowest-ranked practice sites. The FFS model is represented most in the *top* decile, while no CHCs are found in this group. CHC and HSO sites are represented most in the *bottom* decile. Comparing the top and bottom deciles, high efficiency practice sites have more visits per patient per year, undertake more preventive measures, are better in chronic disease management, promote health more actively, and provide more comprehensive care. Efficient practice sites have fewer nurses and fewer nurse practitioners than the least efficient group, and are more likely to be in an urban area than in a rural one. Top performers also employ fewer administrative personnel. Interestingly, however, is that there are almost no differences with respect to continuity of care, access, and the number of physicians in the practice.

Input Scenario 2: Total Cost per Patient

The total cost per patient is used as the input measure in Scenario 2. The resulting DEA efficiency scores have a mean of 0.44 and standard deviation of 0.23. The top performer in this scenario is the FFS model with 40% of its sites in the first quartile and over 70% in the first two quartiles, followed by the HSO model with 58% of all sites in the first two quartiles. The ranking of CHC practice sites is worsened: almost 80% of sites located in the lowest efficiency score quartile are CHCs. This is explained by the fact that the total cost per patient in CHCs is significantly higher, on average, than in other models. Furthermore, the correlation between total costs and total patient roster sizes is the lowest for the CHC model (21.8% compared to 73.8% for the FFS model), thus it is not the number of patients that is driving up the costs of CHCs.

Input Scenario 3: Provider/Patient

Using the provider-per-patient ratio as the input measure, the average efficiency score for the models is 0.41 with a standard deviation of 0.21. The best performers are HSO and FFS models, which are well represented in the first and the second quartile. The worst performer is again the CHC model, which is highly represented in the last quartile and is least represented in the first two quartiles.

Comparing Input Scenarios

The FFS model is the most efficient performer, dominating the 1st and the 2nd quartiles in scenarios 1 and 2 and sharing this top position with the HSO model in Scenario 3. The HSO model is the next best performer in Scenarios 2 and 3, but not in scenario 1 where the FHN model shines. The worst performer is the CHC model in all three scenarios.

The efficiency ranking across models is determined largely by the input variable ranking. The low correlation between the roster size variable and the total expenditure variable may be associated with problems in the self-reported patient roster sizes. There is high variation in patient roster sizes for both top and bottom performers.

Stage Two: Controlling for Organizational and Patient Characteristics

Variation in the efficiency scores can be caused by characteristics of the practice sites themselves as well as the characteristics of the patients served by these practice sites. To this end, we use the efficiency scores as the dependent variables in a regression analysis which takes account of a variety of patient and other characteristics. The choice of regression technique varies according to the problem at hand. Linna, Nordblad, & Koivu (2003) for instance, employed a Tobit procedure to examine the influences of a number of factors on efficiency scores in oral health care provision. Tobit is an ideal procedure whenever the data are censored at one or both ends of the distribution. In the problem considered in this paper, very few observations were found at the high end – only five observations out of 109 in scenario 1 – and no observations were found at the low end of the range. As a result, we employ a simple ordinary least squares (OLS) technique in the analysis that follows (using a Tobit procedure, however, makes very little difference to the results).

INSERT TABLE 4 HERE

INSERT TABLE 5 HERE

The explanatory variables used in the analysis are defined in table 4; descriptive statistics are presented in table 5. The regression results are presented in table 6 for each of the three input scenarios. The OLS model for scenario 1 has the highest explanatory power (R^2 adj. = 0.274) and it is also the scenario in which we have the most confidence as it uses data on the number of providers per practice rather than on the number of patients. The impact of several patient profile variables on efficiency changes dramatically from scenario 1 to scenarios 2 and 3 – likely because of the poor quality of the data on the number of patients in each practice upon which the last two scenarios are based. The discussion below focuses mainly on the first scenario.

INSERT TABLE 6 HERE

Table 6 reveals that CHCs are less efficient than FFS practices in scenarios 1 and 2. However, whenever provider per patient is used as the input, the difference between CHCs and FFSs is negative but statistically weak (20% level of significance). As before, whenever total costs are used as part of the input measure, CHCs fare very badly. HSOs are less efficient than FFS when costs are used, especially in the first scenario, whereas FHNs are less efficient than FFS in the first scenario, but not otherwise.

We calculate the percentage difference in efficiency scores across the different models, holding constant all other influences. In scenario 1 the estimated coefficient on the CHC dummy variable is -0.194. This means that, in comparison with the FFS model, once the influences described in Table 7 are taken into account, the average CHC efficiency score is 28% lower (i.e., 0.194 divided by the average FFS efficiency score of 0.691). The average HSO score is 17% lower than the average FFS practice, while the score of the FHN is 9% lower.

Certain patient characteristics help to explain the efficiency scores. In scenario 1 we find that, *ceteris paribus*, the proportion of patients over age 65 contributes negatively to the efficiency score of the practice site, an influence which is reversed in scenarios 2 and 3. The estimated coefficient on MALE is positive (at the 12% level of significance) in Scenario 1, and negative and statistically significant for the other two scenarios. Notice that HSOs improve their efficiency ranking in scenarios 2 and 3 compared to scenario 1, and they are also characterized by the largest share of patients over 65 years old and of male patients compared to the other models. In contrast, the CHC model has the lowest share in these patient characteristics.

For the most part, having more immigrants among a patient population does not seem to affect efficient scores, corroborating the findings of Sarma, Devlin & Hogg (2007), and consistent with the "healthy-immigrant" effect of McDonald & Kennedy (2004) and Deri (2005).

Surprisingly, the socio-economic status of patients is not found to influence the efficiency score. Having a higher percentage of patients who are unemployed does not affect very much the efficiency of practices, nor does the income and education of patients. Multicollinearity across the model dummy variables and various patient characteristics may be thwarting our attempt to tease out the effect of patient socio-economic variables in the regression analysis. Unfortunately, the exclusion of the variables which are correlated with the model dummies does not improve the significance level of other patient characteristics in explaining the variation in efficiency scores.

The results reported in table 6 show that the fact that the practice serves patients with particular needs does affect efficiency. The larger the proportion of patients who perceive their health as being good, the more efficient the practice site becomes (especially in scenario 1). Similarly, the larger the proportion of patients with conditions lasting more than one year, the lower the efficiency score (again, in scenario 1). Patients with several chronic conditions appear to have a positive impact on efficiency in scenario 1; however this variable is highly negatively correlated with the good health variable, so its estimated coefficient may be unstable. Eliminating both the long-term conditions and chronic conditions variables from the analysis increases the adjusted R-squared for the regression; the impact of "GOOD HEALTH" on efficiency persists.

Geographic location and other environmental characteristics captured by the rurality index have a positive influence on efficiency scores in scenario 1. Practices in rural areas seem to be more efficient, *ceteris paribus*. Once again, the results for the other two scenarios are different, likely because HSOs - which rank higher in scenario 2 and 3 - are located mostly in one urban area. The age of the practice site does not seem to have a statistically significant impact on efficiency. The fact that the number of HSO practices was frozen for many years, and that the FHN model is relatively new, may mean that there is not enough variation in the practice year variable to exert a significant effect on efficiency.

Conclusion, limitations, and policy implications

This paper compares and contrasts four different models of primary care delivery in Ontario with a view to identifying, if possible, which model is the most efficient and under what circumstances. It is clear that how output and inputs are measured matters. One of the innovative elements of the data set used in the paper is its focus on both qualitative and quantitative indicators of output. And one of the clear advantages of the DEA procedure is its ability to incorporate a variety of inputs and outputs and then weigh them in order to present each practice in the best possible light.

If we look only at the raw data on the qualitative indicators of output, CHC practices perform reasonably well. On average, they achieve scores higher than the other three models for three of the eight performance indicators: health promotion, chronic disease management, and comprehensiveness and fare relatively well on prevention and access. The HSO model has the highest average for the continuity of care and access to primary care services variables. FHNs are the best in terms of preventive services.

However, once costs are added to the mix in the DEA analysis, CHCs are the least efficient practice sites virtually across the board, whereas the FFS model performs the best. The efficiency scores of both the HSO and FHN sites are more evenly distributed across all quartiles in comparison to FFS and CHCs. The data show that these efficiency score rankings are driven by the costs of running the practice.

A number of reasons explain the poor efficiency scores of the CHCs. The link between performance indicators and costs may be non-linear, and it therefore may be relatively inexpensive to achieve a low level of performance, but very costly to push these indicators beyond any given threshold. Practices with higher-than-average performance on quality of care and service delivery indicators may require many more resources than practices providing a lower quality of care. Another reason is related to the limitations of the study which focused on clinical primary care; thus the collected expenditure/cost data do not take account of the costs associated with shifting health care from primary care providers onto hospitals, emergency rooms, specialists and outsourced diagnostic services, as well as the costs of prescribed drugs. The costs absorbed by patients, insurance payers, provincial governments and society as a whole are likely to be significant. If the better quality care provided by CHCs were to reduce these expenses it could diminish or eliminate the cost differences found between the CHC model and the others. Finally, it may well be that CHCs are funded are too small relative to the large fixed costs necessary to operate a multidisciplinary health centre, and hence their average costs are very high.

It is also interesting to ask why fee-for-service practices fare so well in the DEA analysis. Part of the story, undoubtedly, is that the broader costs of the FFS approach are borne outside of the practices themselves – costs such as the reputed over-use of specialists. Nevertheless, FFS physicians clearly face incentives to see as many patients as possible given that their remuneration depends upon the number of visits conducted per period of time, while, at the same time, they would want to minimize the costs of running their individual practices because physicians themselves are the residual claimants to the proceeds of the practice.

Overall, the efficiency scores of the FHN and HSO models lie somewhere in-between FFS and CHC practices. In terms of quality of care, FHNs and HSOs are on average at least as good as the FFSs or better, particular in the area of prevention – not surprisingly since physicians in these models receive financial bonuses for reaching certain performance indicators measuring quality of care. In terms of the costs of remunerating physicians, however, FHNs and HSOs are overall more expensive than FFS practices. This may be explained by the fact that physicians who are paid prospectively on a capitation basis have to be compensated for bearing the financial risks associated with running their practices, while CHC physicians who are also paid prospectively (on a salary base) do not bear such risk, as they do not own the centres.

The findings of this paper show clearly that practice type matters. How practices are organized and how physicians are remunerated affect the costs associated with providing patient care. However, we cannot say unequivocally that one type of primary care model dominates. In particular, further research, which better tracks the relationship between the primary care model and the use of other health system resources, is necessary in order to better understand which approach makes the best use of public resources.

REFERENCES

Andes, S., Metzger,L.M., Kralewski,J., & Gans,D. (2002). Measuring efficiency of physician practices using data envelopment analysis. *Managed Care (Langhorne, Pa.),* 11(11), 48-54.

Allen, R. & Gertler, P. (1991). Regulation and the provision of quality to heterogeneous consumers: the case of prospective pricing of medical services. *Journal of Regulatory Economics,* 3, 361-375.

Arrow, K.J. (1986). Agency and the market. In K.J. Arrow amd M.D. Intriligator (Eds.), Handbook of Mathematical Economics, V.3. (pp.1183-1195). North Holland: Elsevier Science Publishers.

Chalkley, M. & Malcomson, J. M. (1998). Contracting for health services when patient demand does not reflect quality. *Journal of Health Economics*, 17, 1-19.

Charnes, A., Cooper, W. W. & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2, 429-444.

Coulson,D. (2005). Primary health care models in Ontario. Toronto, Ontario, Coulson & Associates.

Deri, C. (2005). What lessons can we learn from the body mass of immigrants? Canadian Economics Association annual meeting.

Devlin,R., & Sarma,S. (2007). Do physician remuneration schemes matter? The case of Canadian Family Physicians. Canadian Health Economics Study Group.

Dranove, D. (1987). Rate-setting by diagnosis related groups and hospital specialization. *RAND Journal of Economics*, 18, 417-427.

Ellis, R. P. (1998). Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics*, 17, 537-555.

Ellis, R.P. & McGuire, T., G. (1986). Provider Behaviour under prospective reimbursement: cost sharing and supply. *Journal of Health Economics*, 6, 129-151.

Ellis, R.P. & McGuire, T., G. (1993). Supply-side and demand-side cost sharing in health care. *Journal of Economic Perspectives,* 7(4), 135-151.

Evans, R. G. (1974). Supplier-induced demand: Empirical evidence and implications, in: M. Perlman, ed., *The economics of health and medical care* (London).

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A.* 120(3), 252-281.

Gillet, J., Hutchison, B. & Birch, S. (2001). Capitation and primary care in Canada: Financial incentives and the evolution of health service organizations. *International Journal of Health Services*, 31(3), 583-603.

Giuffrida, A., & Gravelle, H. (2001). Measuring performance in primary care: Econometric analysis and DEA. *Applied Economics*, 33, 163-175.

Gosden,T., Forland,F., Kristiansen,I.S., Sutton,M., Leese,B., Giuffrida,A., Sergison,M., & Pedersen,L. (2000). Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. The *Cochrane Database Of Systematic Reviews 2000,* (3), CD002215.

Hogg et al., 2008

Hollingsworth,B. (2003). Non-parametric and parametric applications measuring efficiency in health care. *Health Care Management Science 6*(4), 203-218.

Holmstrom, B. & Milgrom, P. (1991). Multitask principal-agent analyses: incentive contracts, asset ownership and job-design. *Journal of Law, Economics and Organization*, 7, 24-52.

Huang, Y. G., & McLaughlin, C. P. (1989). Relative efficiency in rural primary health care: an application of data envelopment analysis. *Health Services Research 24*(2), 143-158.

Jack, W. (2005). Purchasing health care services from providers with unknown altruism. *Journal of Health Economics*, 24, 73-93.

Jelovac, I. (2001). Physicians' payment contracts, treatment decisions and diagnosis accuracy. *Health Economics*, 10, 9-25.

Kirigia, J. M., Emrouznejad, A., Sambo, L. G., Munguti, N., & Liambila, W. (2004). Using data envelopment analysis to measure the technical efficiency of public health centers in Kenya. *Journal Of Medical Systems*, 28(2), 155-166.

Laffont, J.-J. & Tirole, J. (1993). A theory of incentives in procurement and regulation. Cambridge, MA: MIT Press.

Laurant, M., Reeves, D., Hermens, R., Braspenning, J., Grol, R., & Sibbald, B. (2005). Substitution of doctors by nurses in primary care. *The Cochrane Database Of Systematic Reviews 2005* (2), CD001271.

Linna,M., Nordblad,A., & Koivu,M. (2003). Technical and cost efficiency of oral health care provision in Finnish health centres. *Social Science & Medicine,* 56(2), 343-353.

Ma, C. A. (1994). Health care payment systems: cost and quality incentives. *Journal of Economics & Management Strategy,* 3(1): 93-112.

McDonald, J. T., & Kennedy, S. (2004). Insights into the 'healthy immigrant effect': health status and health service use of immigrants to Canada. *Social Science & Medicine,* 59(8), 1613-1627.

McGuire, T. (2000). Physician Agency. In A. J. Culyer, & J. P. Newhouse (Eds.), Handbook of health economics. Amsterdam: Elsevier.

Muldoon, L., Rowan, M., Geneau, R., Hogg, W., & Coulson, D. Models of primary care service delivery in Ontario: Why such diversity? *Healthcare Management Forum*. 18-23. 2006.

Mulligan, P. K. (2002). Capitation: the wrong direction for primary care reform. *Canadian Family Physician*, February.

Newhouse, J. P., (1984). Cream skimming, asymmetric information, and a competitive insurance market. *Journal of Health Economics*, 3, 97-100.

Newhouse, J. P. (1996). Reimbursing health plans and health providers: efficiency in production versus selection. *Journal of Economic Literature*, XXXIV, September, 1236-1263.

Pauly, M.V. (1980). Doctors and their workshops: economic models of physician behaviour. Chicago: University of Chicago Press.

Pauly, M. (1984). Is cream-skimming a problem for the competitive medical market? *Journal of Health Economics*, 3 87-95.

Rosenman, R., & Friesner, D. (2004). Scope and scale inefficiencies in physician practices. *Health Economics*, 13(11), 1091-1116.

Salinas-Jimenez, J., & Smith, P. (1996). Data envelopment analysis applied to quality in primary health care. *Annals of Operational Research,* 67, 141-161.

Sarma, S., Devlin, R., & Hogg, W. (2006). Physician production of primary care in Ontario: Evidence from the 2004 National Physician Survey, Canada. Canadian Health Economics Study Group.

Shi, L., Starfield, B., & Xu, J. (2001). Validating the Adult Primary Care Assessment Tool. *Journal of Family Practice*, 50(2), E1.

Sorensen, R. J., & Grytten, J. (2003). Service production and contract choice in primary physician services. *Health Policy*, 66(1), 73-93.

Wagner, J., Shimshak, D., & Novak, M. (2003). Advances in physician profiling: the use of DEA. *Socio-Economic Planning Sciences,* 37, 141-163.

Yalnizyan, A. & Macdonald, D. (2005). CHC cost effectiveness: a review of the literature. Association of Ontario Health Centres, June.

Figure 1: Efficient Frontier

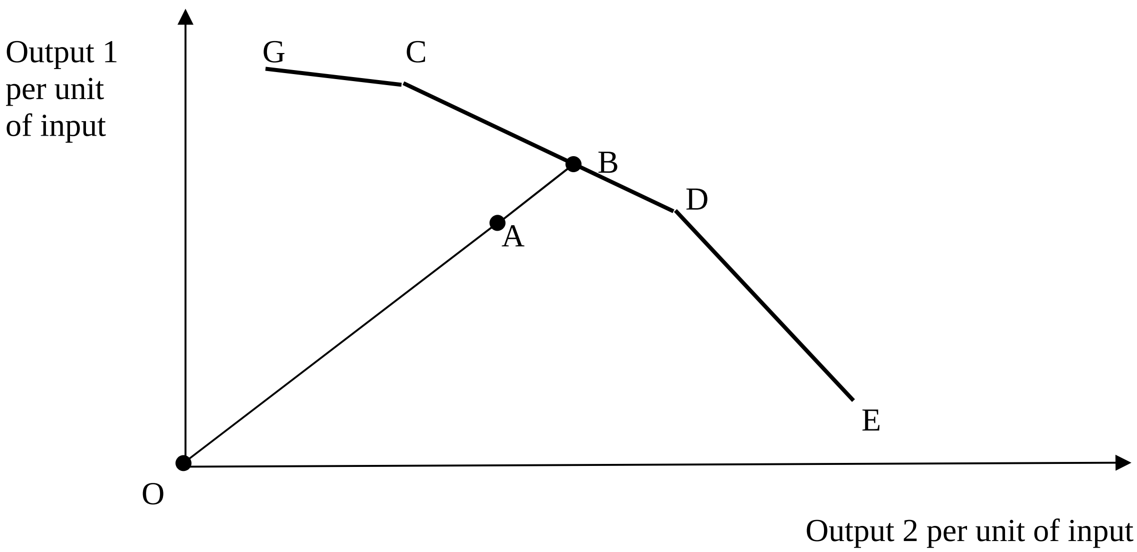

Figure 2

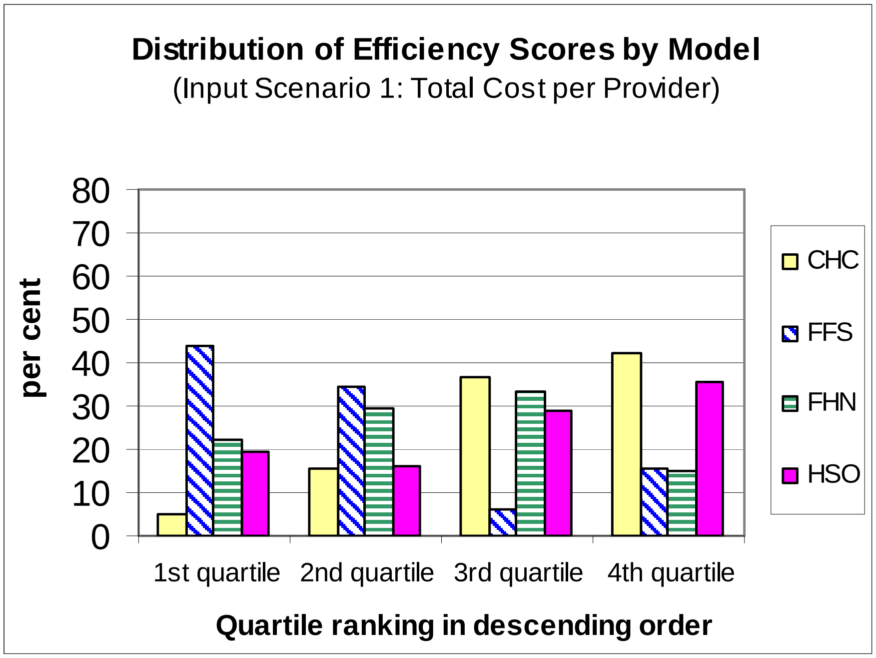
Distribution of Efficiency Scores by Model
(Input Scenario 1: Total Cost per Provider)
per cent
80
70
60
50
40
30
20
10
0
1st quartile
2nd quartile
3rd quartile
4th quartile
Quartile ranking in descending order
CHC
FFS
FHN
HSO

Table 1: Input and output measures

Variable name	**Definition**
I. Output Variables	
1) Intensity of Service Utilization	
Visits/patient	Average number of visits per patient
2) Technical Quality of Care Indicators	
Prevent	Prevention score [0,1]
Chronic_Mngt	Chronic disease management score [0,1]
Health Promo	Health promotion score [0,1]
3) Service Delivery Indicators	
Access	Access to primary care services score [0,1]
Contin	Continuity of care score [0,1]
Compreh	Comprehensiveness of care score [0,1]
II. Input Variables	
Provider/patient	Provider’s FTEs per 1000 patients
Cost/provider	Full practice site’s expenditure per provider’s FTE
Cost/patient	Full practice site’s expenditure per patient

Table 2: Descriptive Statistics [a]

Model	CHC		FFS		FHN		HSO	
Variable	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Output Variables								
Visits/patient[b]	5.2	*0.94*	5	*1.55*	4.06	*1.23*	3.68	*0.88*
Prevent	59	*16.6*	53.8	*14*	60.5	*14.3*	48.5	*12.6*
Chronic_Mngt	71.7	*16.3*	57.9	*20*	59.3	*19.3*	62.2	*16.2*
Health Promo	49.6	*10.8*	40.6	*10.4*	42.5	*9.7*	40.1	*12.7*
Access	3.0	*0.19*	2.94	*0.217*	3.0	*0.28*	3.3	*0.16*
Contin	3.4	*0.21*	3.5	*0.20*	3.5	*0.16*	3.6	*0.16*
Compreh	76.5	*13.1*	61.9	*16.8*	61.3	*17.4*	60.6	*15.1*
Input Variables								
Provider/patient	1.23	*0.58*	0.72	*0.37*	0.77	*0.3*	0.67	*0.34*
Cost/provider	390789	*67893*	250962	*61496*	292584	*76073*	313110	*74680*
Cost/patient	480.6	*242.5*	172.8	*76.5*	214.5	*73.8*	199.4	*83.2*

[a] The statistics are calculated on the practice site's level and then averaged across models

[b] In efficiency score calculations output measure *average visits per patient in one year* may create an unfavourable bias against sites in which physicians take longer time with a patient during a single visit rather than booking several visits. However, the comparison of efficiency scores with the "visit" variable and without it indicates no significant difference in model ranking.

Table 3: DEA efficiency scores by input scenario and model

RANGE	SCENARIO 1					SCENARIO 2					SCENARIO 3				
	All N=109	CHC n=19	FFS n=32	FHN n=27	HSO n=31	All N=109	CHC n=19	FFS n=32	FHN n=27	HSO n=31	All N=109	CHC n=19	FFS n=32	FHN n=27	HSO n=31
1.0	5	0	3	2	0	6	0	4	1	1	2	0	1	0	1
0.9-0.999	1	0	1	0	0	3	0	0	2	1	1	0	0	0	1
0.80-0.899	9	0	5	1	3	1	0	0	0	1	5	1	2	1	1
0.70-0.799	14	1	5	5	3	4	1	1	1	1	4	1	1	2	0
0.6-0.699	23	3	10	6	4	6	1	2	1	2	8	0	3	1	4
0.5-0.599	24	4	2	8	10	12	0	8	1	3	9	0	3	3	3
0.4-0.499	22	8	4	3	7	20	1	6	4	9	18	1	6	4	7
0.3-0.399	9	2	1	1	5	21	0	5	10	6	21	5	4	4	8
0.2-0.299	2	1	0	1	0	24	8	4	6	6	27	5	9	9	4
0.1-0.199	0	0	0	0	0	10	7	1	1	1	13	6	1	3	3
mean	0.604	0.504	0.691	0.628	0.559	0.438	0.255	0.520	0.439	0.467	0.410	0.310	0.439	0.387	0.460
max	1	0.779	1	1	0.857	1	0.709	1	1	1	1	0.831	1	0.848	1
min	0.289	0.293	0.393	0.289	0.359	0.081	0.081	0.196	0.198	0.125	0.115	0.118	0.118	0.167	0.115
sd	0.169	0.122	0.172	0.172	0.144	0.226	0.165	0.227	0.233	0.198	0.211	0.192	0.224	0.186	0.215

Table 4: Explanatory variables for regression analysis

Variable name	Definition
	I. Organizational Structure
CHC	CHC dummy (=1 if the practice site is a CHC site, 0 otherwise)
FHN	FHN dummy (=1 if the practice site is a FHN site, 0 otherwise)
HSO	HSO dummy (=1 if the practice site is a HSO site, 0 otherwise)
Practyr	The number of years in operation in this model
	II. Patient Profile (a proportion of patients with a certain characteristic to the total number of patients who filled out exit questionnaires)
	1. Age/gender profile
Age65	Proportion of patients age 65 or over
Male	Proportion of male patients
	2. Socio-economic status
Immig2	Proportion of patients-immigrants who have been to Canada for 2 or fewer years[a]
Unempl	Proportion of unemployed, excluding housewives and househusbands and those who study
LowIncome	Proportion of patients living in a household with the household income lower than $20,000 per an equivalent household member
LowEdu	Proportion of patients with education less than high school
	3. Health status of patients
GoodHealth	Proportion of patients who perceived their health being excellent or very good[b]
Cond	Proportion of patients with physical, mental or emotional condition that have lasted or are likely to last longer than one year
Chronic	The number of chronic conditions per patient (Average per practice)
	III. Environmental variable
RI	Rurality index. Higher index is associated with relatively remote areas with lower level of services available.

[a]A threshold of 5 years yields equal significance in the analysis.
[b]A variable representing the proportion of patients who perceived their health being fair or poor has not been identified as a significant predictor.

Table 5: Summary statistics for explanatory variables in regression analysis

Model	Age65	Male	Immig2	Unempl	Low Income	LowEdu	Good Health	Cond	Chronic	Practyr	RI
CHC											
mean	14.5	26.1	18.7	12.8	0.33	23	41.6	47	1.9	17.5	13.5
sd	11.9	13.2	22.4	8.2	0.15	10.5	8.5	14.8	0.5	7.3	17.1
FFS											
mean	20.2	34.6	10.4	5	0.15	14	46	42.9	1.75	16.3	13.1
sd	9.7	17.7	12.1	4.7	0.09	10	9.9	10.8	0.47	10.2	18.5
FHN											
mean	22.8	34.3	3.3	3.6	0.15	14.1	47.6	40.7	1.8	2.34	18.3
sd	9.9	13	3.7	3.1	0.09	9.5	12.6	6.8	0.45	1.17	20.8
HSO											
mean	24.8	39.8	2.91	3.43	0.15	16.4	50.7	37.7	1.68	15.3	8
sd	11.1	9.69	4.21	3.33	0.08	8.51	10.1	10	0.35	6.7	9.23
Total											
mean	21.2	34.6	7.89	5.54	0.18	16.3	47	41.6	1.78	12.7	12.9
sd	11	14.3	13	5.87	0.12	9.98	10.8	10.9	0.45	9.39	16.9

Table 6: The influence of patient and practice characteristics on efficiency

Explanatory variable	Scenario 1		Scenario 2		Scenario 3	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Constant	0.518	0.000	0.649	0.001	0.547	0.004
CHC	-0.194	0.000	-0.244	0.002	-0.091	0.197
HSO	-0.120	0.005	-0.090	0.131	-0.017	0.753
FHN	-0.063	0.197	-0.074	0.291	-0.058	0.366
Age65	-0.005	0.004	0.006	0.026	0.007	0.002
Male	0.002	0.118	-0.003	0.062	-0.003	0.038
Immig2	0.000	0.800	-0.000	0.890	0.001	0.661
Unempl	0.003	0.388	-0.004	0.425	-0.009	0.086
LowIncome	-0.011	0.954	0.152	0.584	0.118	0.647
LowEdu	-0.003	0.158	0.003	0.359	0.004	0.151
GoodHealth	0.004	0.033	0.001	0.707	-0.000	0.923
Cond	-0.003	0.137	0.001	0.772	0.001	0.548
Chronic	0.083	0.119	-0.143	0.059	-0.113	0.105
RI	0.002	0.013	-0.002	0.253	-0.003	0.023
Practyr	0.000	0.884	0.001	0.689	-0.000	0.943
Adj. R-squared	0.2740		0.181		0.193	