

Primary Care Efficiency Measurement Using Data Envelopment Analysis: A Systematic Review

Ferruccio Pelone · Dionne Sofia Kringos · Alessandro Romaniello ·
Monica Archibugi · Chiara Salsiri · Walter Ricciardi

Received: 23 February 2014 / Accepted: 6 November 2014 / Published online: 26 November 2014
© Springer Science+Business Media New York 2014

Abstract There is a gap between the demand and supply of efficiency analyses within primary care (PC), despite the threatening financial sustainability of health care systems. This paper provides a systematic literature review on PC efficiency analysis using Data Envelopment Analysis (DEA). We reviewed 39 DEA applications in PC, to understand how methodological frameworks impact results and influence the information provided to decision makers. Studies were combined using qualitative narrative synthesis. This paper reports data for each efficiency analysis on the: 1) evaluation context; 2) model specifications; 3) application of methods to test the robustness of findings; 4) presentation of results. Even though a consistent number of analyses aim to

support policymakers and practice managers in improving the efficiency of their PC organizations, the results indicate that DEA—at least when applied to PC—is a methodology still in progress; it needs to be further advanced to meet the complexity that characterizes the production of PC outcomes. Future studies are needed to fill some gaps in this particular domain of research, such as on the standardization of methodologies and the improvement of outcome research in PC. Most importantly, further studies should include extensive uncertainty analyses and be based on good evidence-based rationales. We suggest a number of considerations to academics and researchers to foster the utility of efficiency measurement for the decision making purposes in PC.

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

Electronic supplementary material The online version of this article (doi:10.1007/s10916-014-0156-4) contains supplementary material, which is available to authorized users.

F. Pelone (✉) · A. Romaniello · M. Archibugi · C. Salsiri ·
W. Ricciardi

Department of Public Health, Population Medicine Unit,
Catholic University “Sacro Cuore”, Largo “F. Vito”, n.1, Rome, Italy
e-mail: fpelone@gmail.com

A. Romaniello
e-mail: alex.rom@gmail.com

M. Archibugi
e-mail: monicarchibugi@hotmail.it

C. Salsiri
e-mail: chiara.salsiri@gmail.com

W. Ricciardi
e-mail: wricciardi@rm.unicatt.it

D. S. Kringos
Department of Social Medicine, Academic Medical Centre (AMC)—
University of Amsterdam, Meibergdreef 9, 1105, AZ Amsterdam,
The Netherlands
e-mail: d.s.kringos@amc.uva.nl

Keywords Primary health care · Health care delivery ·
Efficiency · Data envelopment analysis · Decision making ·
Systematic review

Introduction

Health care systems aim to maintain, restore and improve the health status of the population. It is therefore crucial to assess the achievement of goals such as effectiveness, equity and responsiveness in relation to the health care resources consumed [1]. It is well recognized that primary care (PC) could play a key role in improving health outcomes, economic sustainability and the overall quality standards of health care delivery [2]. PC provides easy access first contact care, coordinates care across all levels of the health care system, and provides comprehensive health care services on a continuous basis to the majority of the population [3, 4].

The growing (complex) demand for health care, the rising expenditures and decreasing health care budgets, put priority setting at the top of the research agenda, paying due attention to the performance of PC systems and providing policymakers

with tools to better inform their decisions [5]. On account of this background, efficiency analysis and benchmarking may play an important role in enabling health policymakers [6, 7], managers and providers to improve PC structures, processes and outcomes [8], amply motivating a body of research on the overall productivity and quality of PC systems [9].

The analysis of health care efficiency is a challenging task. First, because of the unclear boundaries of some health care delivery areas (e.g. PC, public health). Furthermore, efficiency analysis must consider the high complexity of the technological transformation process of turning resources in health outcomes within these areas. Nevertheless, since the '80s, efficiency analysis in health care has been experiencing a fast growth [10].

Although there are several alternative techniques available to carry out efficiency analysis in health care (see box 1), a predominant use of Data Envelopment Analysis (DEA) with a prevalent focus on hospitals, nursing homes and physicians has been highlighted by several reviews on frontier efficiency measurement in health care [10–15].

Over the past decades, DEA has increasingly been applied within the PC context [5] showing its suitability for this setting. DEA has a number of features which make it an attractive tool for efficiency measurement of PC delivery: it can handle effectively the existence of multiple PC resources (hereinafter referred as inputs) and multiple health outcomes (hereinafter referred as outputs) in the transformation process. Furthermore, it does not require strong assumptions about the underlying technology linking the inputs to the outputs, and it measures efficiency in relative -in stead of absolute- terms.

However, while the supply-side (e.g. research) of DEA applications in PC has grown exponentially over the last decades, the demand-side (e.g. health policy) is still under-developed [10]. There is a twofold explanation for this gap: firstly, efficiency analysis needs to provide more trustworthy and reliable evidence [16]; and secondly, research needs to be more “policy-oriented”, focusing on the nature and form of (in)efficiency rather than only on its quantification [13].

It is currently unknown how to bridge the chasm between evidence and practice in the domain of PC efficiency analysis. Therefore, research is required to better address the methodological issues (e.g. differences in model specifications, selection of variables, and sensitivity analysis toward different empirical assumptions) in performing efficiency analysis in PC to improve the provision of valuable evidence to policymakers. This article aims to make a first step in this direction to increase our understanding of how DEA studies are performed in PC to inform health policy. We aim to systematically review the empirical background and findings of DEA

applications within PC to appraise their ability in generating constructive evidence to inform health policy.

This research question will support the improvement of the application of DEA in primary care. Developing a standard set of criteria in the design and execution of DEA studies in PC may be very useful, for which this review represents a starting point.

Box 1. Analytical efficiency measurement methods

Efficiency analysis	Is the study of the development of analytic tools aimed to measure the efficiency of health care organizations and systems. It develops frameworks and methods to estimate the efficient frontier in the production of health care and to derive the relative efficiency of each productive unit. Efficiency is considered in this review as the relation to which health outcome are produced in relation to the resources consumed
Analytical efficiency measurement methods	The methods developed within Efficiency analysis differ according two basic criteria ⁽²⁾ : 1) whether they are parametric or non-parametric methods (i.e. the functional form of the efficient frontier is imposed a priori or it is obtained a posteriori from the sample observations in an empirical way), and 2) whether they follow a deterministic or a stochastic approach (i.e. the distance of a production unit from the efficient frontier is entirely attributed to its inefficiency or partly due to statistical noise and random error, in addition to its inefficiency). Non-parametric and parametric techniques have complementary pros and cons. Even though it seems there is no strong empirical evidence in favour of one method over another ⁽³⁾ , nonparametric deterministic methods, such as DEA, have become more common in measuring the efficiency of health care delivery
COLS - parametric - deterministic	Corrected Ordinary Least Squares (COLS) approach assumes a given functional form of the relationship between inputs and outputs and estimates the unknown parameter of the function by ordinary least squares (OLS) regression, and the residual (the estimated error) representing inefficiency
SFA - parametric - stochastic	Within this approach the unknown parameters of the function are estimated by maximum likelihood techniques. Contrary to OLS regression, the Stochastic Frontier Analysis (SFA) model decomposes the residuals into a symmetric component representing statistical noise and an asymmetric component representing inefficiency
DEA - non-parametric - deterministic	Data Envelopment Analysis (DEA) is a non-parametric approach determining a piecewise linear efficiency frontier along the most efficient production units by means of linear programming to derive

Box 1. Analytical efficiency measurement methods

	relative efficiency measures of all other production units
Stochastic DEA - non-parametric - stochastic	Stochastic DEA is a recently developed technique which combines the features of nonparametric and stochastic methods, combining the basic idea of DEA and SFA

Sources: [7]

Methods**Search strategy**

The PRISMA guidelines were followed in the search process. The search strategy focussed on original peer-reviewed papers published until August 2014, without time or language restrictions. Six databases were searched by means of a combination of specific subject headings and free text terms. Based on a preliminary search a Boolean algorithm to search MEDLINE was developed:

```
#1 = "Efficiency"[Mesh]
#2 = "Benchmarking"[Mesh]
#3 = "Data Envelopment Analysis" [Title/Abstract]
#4 = #1 OR #2 OR #3
#5 = "Primary Health Care"[Mesh]
#6 = "Physicians, Primary Care"[Mesh]
#7 = "primary health care" [Title/Abstract]
#8 = "general practice*" [Title/Abstract]
#9 = "family physician*" [Title/Abstract]
#10 = "primary care physician*" [Title/Abstract]
#11 = #5 OR #6 OR #7 OR #8 OR #9 OR #10
#12 = #4 AND #11
```

This Boolean search algorithm was run first and later transposed to Embase, EconLit, SCOPUS, ISI Web of Science and CINAHL Plus with Full Text. Additionally, a search was carried out in MEDLINE by using the 'related articles' algorithm, supplemented with hand searches of the references of related articles.

Studies were included if they carried out an efficiency assessment by means of DEA applied to PC in upper-middle and high-income countries as defined by the World Bank [17]. We defined PC as: the provision of directly accessible health care services including health promotion, disease prevention, curative care, rehabilitative care and palliative care, provided by clinicians who are accountable for addressing a large majority of personal health care needs [18].

The PRISMA guidelines were followed in the search process [19].

Data abstraction

Two authors (FP and DK) performed the overall search strategy and reviewed in parallel the titles and abstracts of all resulting articles to ascertain whether they met the eligibility criteria. The full texts of all included articles were examined in parallel by two teams of reviewers (AR-MA and FP-CS) to analyze whether they met all inclusion criteria using specific checklists designed for this study. Disagreements were resolved by peer discussion. Once consensus on article inclusion was achieved, one reviewer (DK or FP) classified each study. Each efficiency analysis was summarized in preset tables by abstracting the following details based on Jacobs et al. 2006 [7] on the applied methodological framework:

- Setting and context of analysis;
- Empirical specifications;
- Presentation and discussion of study results;
- Sensitivity analysis.

The decision not to attempt a quantitative synthesis of the study results was determined by an a priori assessment of the large number of sources of possible heterogeneity amongst studies likely to be eligible. These include the wide range of services delivered across PC, besides the methodological and modeling differences between studies. Previous reviews have used analytic techniques (i.e. meta-analysis, simulation models) to quantify the impact of modeling choices on the estimated efficiencies in reviewed studies [16, 13, 20].

Results**General description**

A total of 39 publications fulfilled the inclusion criteria [21–59] (see Fig. 1). Table 1 shows an overview of all included studies. Ten were based on data from the United States, and nine reported data on PC efficiency analysis from the United Kingdom. A substantial number of the studies were placed in other European countries, including Spain, Greece, Finland, Italy and Portugal. One reported data on 22 European countries, and six publications were placed in other EME Countries, including Brazil, Canada, Chile and New Zealand. Over two-thirds of the studies have been published in the last 15 years, showing an increased interest over time to apply DEA to analyze efficiency in PC (see Fig. 2).

Context of the analysis

Figure 3 provides an overview of the different PC settings analyzed. Largely, DEA has been used to measure the relative

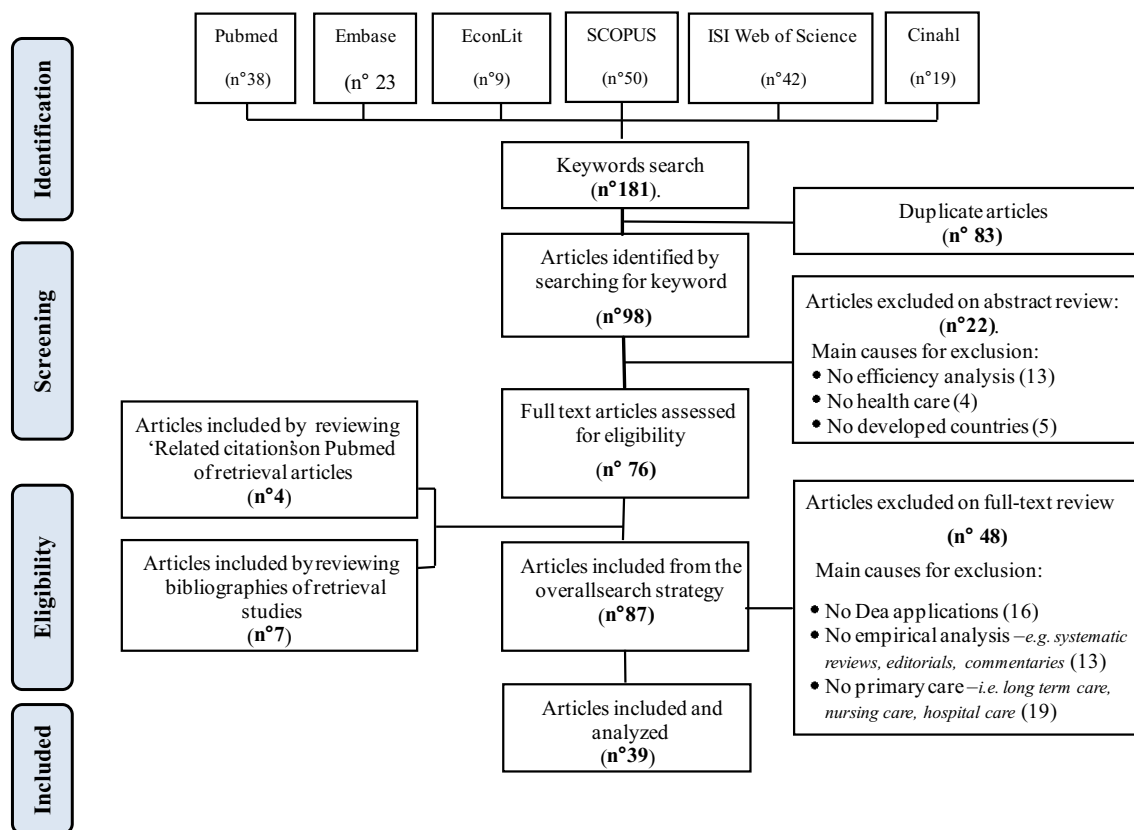


Fig. 1 Flow of included studies

efficiency of a wide range of activities provided by different PC centers or districts within national/regional PC systems [21–36]. In a large number of studies DEA has also been used to compare PC providers in delivering specific activities (e.g. dental care) and treating specific diseases (e.g. diabetes) [37–48].

Particularly in the UK a number of PC DEA applications focused on general practice, comparing providers' services delivery (e.g. vaccinations) [49–56]. Three studies focussed on the efficiency of individual physicians' practice patterns in providing PC services [57–59]. Most of the efficiency analyses focused on individual PC professionals (i.e. family physicians/general practitioners, ambulatory care specialists). A number of DEA studies considered as decision making units (DMUs) PC practices, community-based PC programs or PC/outpatient programs delivered within vertical-integrated PC organizations. Five out of the 33 efficiency analyses considered as DMUs administrative entities, to compare the relative efficiency of different states [24], regions [49], municipalities [27, 28, 43] in providing PC (see Table 1).

Efficiency has been assessed in terms of different concepts including technical, scale, and cost efficiency. All of the included DEA applications were focused on technical efficiency –i.e. producing the maximum amount of output from a

given amount of input, or alternatively producing a given output with minimum quantities; in a number of DEA applications the main focus was the extent to which evaluated organizations could take advantage of returns to scale by altering its size towards optimal scale of delivery. Finally, two analyses [30, 37] were focused on comparing PC organizations in terms of cost efficiency –i.e. overall cost minimisation given observed input prices producing a given output.

Efficiency changes over time was assessed by only one study [53], by decomposing the estimated production technology of each year into indices of pure technical efficiency change, scale efficiency change and technological change of English Family Health Service Authorities (FHSAs). All other DEA studies measured the efficiency of a sample of PC DMUs on a 1 year benchmark period by using cross-sectional data.

Discussion

The body of research conducted over the past decades into the PC efficiency measurement by means of DEA allows us to answer three questions [7, 13]: What empirical assumptions

Table 1 Summary of study characteristics

Primary care setting	Country	Ref.	Level of analysis (DMU)		Empirical specifications			First author, year
			Organization type	No of units	Efficiency orientation	Technology assumption	Weight restrictions	
PC whole scope of activities	ES	[21]	PC centres	94	input	VRS	Y	Cordero Ferrera, 2014
	ES	[22]	PC centres	130	input	VRS	N	Deidda, 2014
	PT	[23]	Health care centers	22	both	CRS-VRS	N	Ferreira, 2013
	IT	[24]	EU Member States	19	output	CRS	N	Pelone, 2013
	ES	[25]	PC centres	97	input	VRS	N	Cordero-Ferrera, 2011
	CA	[26]	PC practices	137	output	CRS-VRS	N	Milliken, 2011
	CHL	[27]	Municipalities	345	output	CRS-VRS	N	Ramírez-Valdivia, 2011
	BR	[28]	Municipalities	359	output	VRS	N	Varela, 2010
	GR	[29]	PC centres	194	input	CRS-VRS	N	Kontodimopoulos, 2007
	US	[30]	a) PC practices b) Specialty care practices	a) 156 b) 346	input	VRS	N	Rosenman, 2004
	US	[31]	PC practices	115	n.r	n.r	N	Andes, 2002
	ES	[32]	PC centers	66	output	VRS	N	Pinillos, 2002
	GR	[33]	PC centres	133	input	CRS	N	Zavras, 2002
	ES	[34]	PC centres	54	input	CRS-VRS	N	Garcia, 1999
	FI	[35]	PC centres	202	input	CRS-VRS	Y	Luoma, 1996
	US	[36]	PC clinics	77	input	n.r	N	Huang, 1989
	Dental care	FI	[37]	Health centre	228	input	CRS-VRS	N
GB		[38]	Community Dental Services	n.r	input	CRS-VRS	N	Buck, 2000
Diabetes care	IT	[39]	Family physician practices	96	input	VRS	Y	Testi, 2013
	GB	[40]	General practices	14	input	CRS	N	Amado, 2009
	MX	[41]	Family units	47	output	VRS	N	Salinas-Martínez, 2009
GP prescribing patters	GB	[42]	General practices	106	input	CRS	N	Bates, 1996
Hypertension care	BR	[43]	Municipalities	66	output	VRS	N	Rabeti, 2011
Immunization activities	NZ	[44]	PC practices	24	both	CRS-VRS	N	Rouse, 2010
Otitis media	US	[45]	PCPs	160	input	CRS-VRS	Y	Ozcan, 1998
Perinatal care	GB	[46]	DHAs	83	output	CRS	Y	Thanassoulis, 1995
Sinutitis	US	[47]	PCPs	178	input	CRS	N	Pai, 2000
	US	[48]	Physicians (generalist, specialists)	176 (152, 24)	input	CRS	N	Ozcan, 2000
general practice	IT	[49]	Regions	20	input	CRS	N	Pelone, 2012
	US	[50]	PC centres	67	input	VRS	N	Rahman, 2012
	PT	[51]	DHAs	337	output	CRS	Y	Amado, 2009
	GB	[52]	FHSAs	90	input	CRS-VRS	N	Giuffrida, 2001
	GB	[53]	FHSAs	90	input	VRS	Y	Giuffrida, 1999
	GB	[54]	General practices	107	output	CRS	N	Bates, 1998
	GB	[55]	FHSAs	90	both	CRS	N	Salinas-Jiménez, 1996
	GB	[56]	General practices	52	n.r	n.r	N	Szczepura, 1993
	PCPS practice	US	[57]	PCPs	21	input	CRS	N
US		[58]	Physicians (family physicians, internists, sub-specialists)	326 (86, 169, 71)	input	CRS	Y	Chilingerian, 1997
US		[59]	PCPs	326	input	CRS	N	Chilingerian, 1996

PC Primary care, FHSAs Family health service authorities, DHAs District health authorities, PCPS Primary care physicians, GP General practice, *input* Input oriented, *output* Output oriented, Y Yes, N No, n.r: Not reported

are needed for a reliable and valid DEA model? How are empirical findings reported and summarized to inform policy-making? How is uncertainty addressed to confirm the robustness and the validity of the study conclusions?

What empirical assumptions are needed for a reliable and valid DEA model? A valid and reliable DEA model should specify four assumptions [7]: (1) the technology assumption of the PC delivery process (i.e., whether to assume constant or

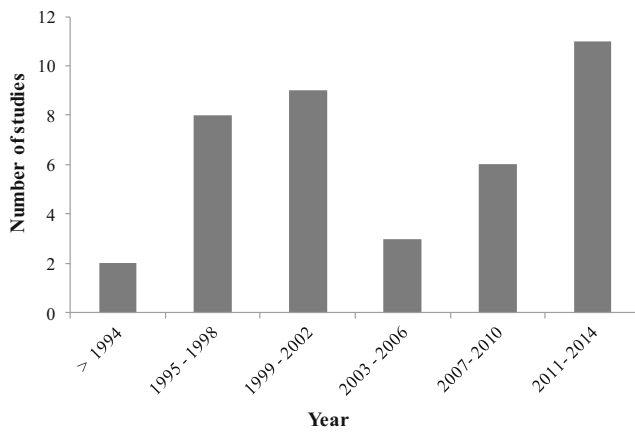


Fig. 2 Number of DEA applications in primary care

variable returns to scale), (2) the efficiency analysis orientation (i.e., whether to assume an input or an output orientation), (3) the input–output variables, and (4) whether to incorporate weight restrictions.

Defining the technology assumption The choice of a Constant Returns to Scale (CRS) or a Variable Returns to Scale (VRS) should depend on the context, the level and perspective of analysis, in addition to the technology that links the inputs to the outputs in the transformation process [7]. For example, if the DEA exercise is conducted from a policy maker's point of view (i.e. the analyst aims to measure efficiency regardless of any managerial factors linked to the PC delivery process), then a CRS approach is appropriate. If the PC efficiency analysis uses a managerial point of view (i.e. focussing on the extent to which elements such as scale of operations or providers' practice style affect productivity in PC delivery) a VRS approach is advisable.

A VRS approach was opted in 12 efficiency analyses [21, 22, 25, 24, 28, 30, 32, 39, 43, 41, 50, 53] –see Table 1. For example, Rabetti and Freitas [43] adopted a VRS DEA model under the assumptions that any variation in the size of DMUs

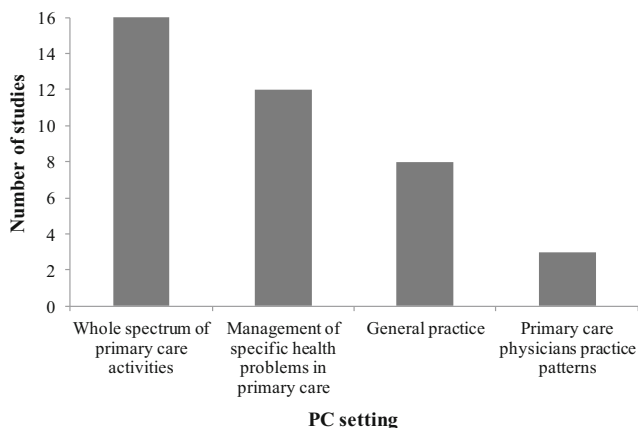


Fig. 3 Scope of DEA applications in primary care

(municipalities) under scrutiny might interfere in the production scale of service and results (provision of hypertension care services). Several studies provided no clear empirical assumptions underlying their choice of applying a VRS model in their DEA applications [28, 41].

Fourteen studies applied a CRS technology assumption, justifying this choice by three types of assumptions:

1. The analysis revealed little evidence on scale economies in the production process –i.e. there was a linear relationship between inputs and outputs [40, 45, 42, 47, 51, 54, 57];
2. It is a sensitive model for seeking out technical and scale efficiency –i.e. it can be used to determine whether returns to scales are locally increasing or decreasing [58, 59];
3. Using ratios adjusted for the DMUs size as input–output data rather than absolute numbers, which removes any information about the size of DMUs [34, 49, 55].

Eleven included DEA applications focused on measuring scale (in)efficiency, using both a CRS and a VRS DEA model on the same dataset [23, 26, 27, 29, 34, 35, 44, 37, 38, 46, 52].

Defining the efficiency analysis orientation The choice of input or output orientation should be based on what the practice managers are able to control better [7] –the resources or the outcomes of the production process. Three out of thirty-nine included analyses ran both an input-oriented and an output-oriented DEA model, to know simultaneously the proportional reduction in input usage possible for a given level of output and the proportional expansion in output levels possible keeping the input fixed [23, 44, 55].

Rouse et al. [44] found that scale efficiency scores were higher in case of input-orientation compared to an output-orientation, for each DMU in the sample. The results of Salinaz-Jiménez et al. [49] DEA study go in the opposite direction of the first analysis, as DMUs under investigation were shown to be more efficient according to an output orientation than an input orientation. The authors explained this finding as depending on the number of input–output variables used in their DEA model –i.e. a DMU obtained higher efficiency scores under an output-oriented model than the alternative, because DEA was more demanding since the potential improvement was divided over 7 output variables rather than only one input.

Largely, included studies examined the relative efficiency of PC DMUs by using an input orientation –i.e. keeping output fixed and explore the proportional reduction in input usage, while nine analyses were output-oriented DEA studies –i.e. keeping input fixed and explore the proportional expansion in output that is possible (see Table 1).

All of the reviewed output-oriented DEA applications were run to quantify the relative efficiency among a sample of

DMUs considering outputs as variables under-control of the decision makers.

Some analysis [31, 56] did not provide any information on the orientation opted for the efficiency analysis or on the technology assumption about the production process.

Selecting and measuring the input–output variables DEA results are sensitive to the choice of variable specifications, determining most of the heterogeneity [7, 16]. Most studies included in this review selected their variables based on data availability, rather than collecting new data to construct the best possible measure. Thus, many studies considered PC activities as a measure of PC outputs, given the difficulty in measuring PC health outcomes. Our results indicate that currently the golden standard is to rely on PC quality indicators when there is evidence that such measures are a proxy close to health improvements.

DEA can handle several inputs and outputs in the transformation process. However, because it is a data-driven methodology, results depend on the relevance of the included variables [16]. Additional file 1 shows differences and similarities with regard to input–output variable selections of efficiency models.

Considering the outputs, only two studies defined them in terms of health outcomes produced [47, 57]. Wagner et al. [57], whom compared the relative efficiency of PC physicians, advanced an approach of adding variables to the model in a ‘stepwise’ manner to better understand the DEA results.

A higher number of studies (20 – see additional file 1) compared the relative efficiency of the PC DMUs on the basis of PC activities (e.g. visits performed, examinations provided). Several studies defined outputs in terms of PC quality indicators [21, 26, 30, 40–42, 57]. This may overcome the drawbacks of basing efficiency models on PC activities, such as the inability to take account of the quality of care delivered, therefore possibly rewarding PC DMUs that produce more activities than other DMUs just because they are operating at a lower quality level. A set of variables for which the outputs included both activity levels and quality measures was adopted in five DEA applications [21, 43, 52, 53, 59].

Regarding input, there are three main input categories: labour, capital, and consumable resources (e.g. consumed drugs). These variables can be measured in physical unit or in monetary terms, as an overall aggregate measure or a set of disaggregated measures.

A single measure of aggregated input in terms of costs –e.g. expenditure, operating costs– has been opted by six studies [23, 27, 28, 35, 54, 55]; while the most common approach was to consider a disaggregated combination of inputs –i.e. mix of labour or capital resources; five DEA models considered both single and disaggregated measures of inputs (see additional file 1). Inputs were measured in physical units in either total working hours or working hours by skill level [23, 36, 38, 40,

41, 47, 56], or in absolute number of individuals by skill level [21, 22, 25, 29, 32, 33, 51, 53]. Only four studies divided their input mix into two categories (i.e. capital and labour inputs), adding to the mix of professionals some proxy measures of the physical capital used in the production process (e.g. square footage of facilities) [30, 31, 34, 37, 43].

Incorporating weight restrictions The introduction of weight restrictions always involves value judgements. It can reflect the decision makers’ preferences in the process of assessing efficiency [60]. However, most of the PC efficiency studies did not incorporate any restrictions to the input–output weights in specifying their efficiency models, partly to avoid compromising the objectivity of DEA and partly because of the lack of decision makers’ involvement in the modelling process. Little has been done to reflect decision makers’ preferences in the process of assessing efficiency [e.g. 58] and to reflect the relative importance of certain input–output combinations [e.g. 47].

A set of unrestricted input–output weights are endogenously determined in the DEA model without the need for subjective judgments, assigning to each DMU its best attainable efficiency score [60]. Nevertheless, eight studies included value judgments in the DEA models, for four types of reasons:

1. To incorporate prior views on the value of individual inputs and outputs [35, 53];
2. To relate the value of certain inputs and/or outputs [21, 47, 51];
3. To delineate how certain physicians’ preferences (e.g. more drugs than visits or hospitalization) influence the healthcare delivery process [39]
4. To incorporate prior views on (in)efficient DMUs [46, 58].

The DEA applications included in the third category incorporated value judgments by means of two stage DEA models –firstly, an unbounded DEA model was run and unrestricted weights were calculated. Secondly, the studies imposed upper and lower bounds for each factor weight on the standard DEA model– to incorporate prior strategic intentions of the management in the construction of a best practices production frontier for physicians.

How are empirical findings reported and summarized to inform policy-making? Our results point to the following recommendations that are important to consider when summarizing the empirical findings. Firstly, the focus in the reporting of the empirical findings should be on transparency. Therefore, it is useful putting such information in summary tables or to describe it in a narrative way. Secondly, it is important to consider all implementation issues such as the feasibility, the suitability, and the relevance of adopting

efficiency changes in the specific study context. Finally, in discussing results it is necessary to compare the obtained results with those obtained in similar setting, to help policy makers in using of such evidence.

We found a large heterogeneity among the reviewed DEA applications, as regard to the presentation and discussion of their empirical findings (see Table 2). Few papers [31, 33, 43, 44, 37, 49, 52] compared their empirical findings with related studies. Only four authors [44, 52, 57, 58] discussed the applicability of their findings to other PC settings (generalizability of results). A number of studies mentioned other important factors to be considered in the efficiency decision under consideration, such as equity [40, 51], access [45, 48], effectiveness [40] and financial incentives [26].

We classified the DEA studies into two categories according to the type of results: methodology-oriented (i.e. the findings were discussed focusing on the methodological added value of the DEA exercise) and practice-oriented applications (i.e. the findings were analyzed focusing on the empirical added value of the DEA exercise for the evaluation context).

The five examples of the ‘methodology-oriented’ DEAs in PC provided varying results. Chilingirian and Sherman [58] demonstrated how a ‘cone ratio DEA model’ can incorporate strategic thinking and executive accountability when establishing clinical benchmarks. Several DEA applications provided insight into the variable selection, highlighting the central role of quality indicators [55] and health outcomes [47] in applying DEA to PC. In 2003, Wagner [57] proposed a method for the selection of model inputs and outputs, using a “stepwise” approach, and a sensitivity analysis using super-efficiency scores, demonstrating advantages for physician profiling. Finally, Giuffrida and Gravelle by comparing DEA and SFA estimates of efficiency in PC, found that one method could not be preferred to another [52].

Most of the studies (32 out of 39) presented their findings providing important implications for practice. Input-oriented DEA models (either CRS or VRS) provided policymakers with information on the potential gains –in delivering PC- in terms of resources use. The output-oriented VRS DEA applications showed that efficiency scores for a given PC practice tended to improve as the output formulation became more complex [37, 42]. Both CRS and VRS output-oriented DEA papers, reported and discussed their empirical findings, in an effort of offering decision makers information about how to improve the relative efficiency of their PC delivery.

Several analyses provided insights toward a better understanding about the effects of scale size on the productivity and efficiency achievement for PC units under consideration (see Table 2).

Only few DEA applications discussed how to put their findings into practice. They discussed issues of implementation [31, 40, 48, 51] or the feasibility of adopting efficiency

changes given the existing operational constraints [29, 37]. Kontodimopoulos et al. (2007) identified that the theoretically possible efficiency improvements resulted from the analysis were plausible and desirable, but from a societal perspective, the reduction of resources needed to obtain efficiency improvements is undesirable for populations with limited health care options [29].

How is uncertainty addressed to confirm the robustness and the validity of the study conclusions? DEA does not require any empirical assumption on the shape of the functional form linking the inputs to outputs. This reduces the exposure to the theoretical assumptions of its empirical specification, compared to studies adopting parametric approaches. DEA-based studies have a higher degree of uncertainty than statistically-based studies using models of uncertainty on the robustness of results [61, 62]. DEA does not offer any diagnostics to assess the appropriateness of the model with respect to a number of methodological issues, such as: outlier observations, measurement accuracy and random error, the choice of the type and number of variables with respect to the total DMUs under examination, and the application of sensitivity analysis –to ascertain the robustness of the results, and construct data ranges within which results remain unchanged [63].

There are some examples of PC DEA models dealing with these detriments. While no study attempted to address biases of results due to statistical error, few analysts tried to improve the robustness of their results by discriminating the effect of atypical input–output combinations by means of a ‘super-efficiency analysis’ [28, 57]. This method involves rerunning the DEA model, removing in turn each efficient DMU, and calculating a measure of the resulting change. As a result it is possible to examine the degree to which the DMUs influence the efficiency solutions [64].

Some studies undertook procedures aimed to select the most relevant input–output variables to enhance the soundness of the results of the DEA model. Other studies performed statistical tests (i.e. Spearman correlation) to ascertain that input variables were positively correlated with output variables to assure the existence of a trade-off between inputs and outputs in obtaining efficiency [28, 55] –i.e. an increase in inputs should reduce efficiency while a decrease in outputs should increase efficiency. Another way to improve the feasibility of the DEA model with regard to variables selection was suggested by Giuffrida and Gravelle [52], who tested the robustness of the results obtained by running two DEA models with various output variables (see additional file 1).

Finally only few studies, in the effort to deal with random noise or measurement error in their data, carried out a sensitivity analysis (i.e. bootstrapping procedure) of the DEA model, to construct confidence intervals of the efficiency scores [21–24, 27, 41, 49, 53].

Table 2 Summary of study findings

Ref.	Efficiency scores	Summary of study results	First author, year
[21]	Mean TE ^A : model 1 0.828; model 2–0.751; model 3 0.860 0.833; model 4 0.873	Including information about both quality of care and environmental variables in the analysis results in efficiency measures that can be interpreted as an accurate reflection of performance.	Cordero Ferrera, 2014
[22]	Mean TE ^B : model 1–0.733; model 2: 0.613	Health information technology may lead to higher-quality care by introducing technology in the production function.	Deidda, 2014
[23]	Mean CRS TE ^C 0.918 1.099; Mean VRS TE ^C 0.928 1.084;	A significant level of inefficiency was observed, although there was a general improvement in efficiency between 2009 and 2010. The dissemination of best practices is likely to improve efficiency of PC while maintaining or even improving quality and equity.	Ferreira, 2013
[24]	Mean TE: 0.97 (structure)–0.80 (Outcome)	The structure–process DEA model shows that 10 countries are relatively efficient in delivering their processes, using the best mix of structure dimensions, with efficiency scores equal to the unity or 100 % relative to the other PC systems	Pelone, 2013
[25]	Mean TE ^B : model 1–0.633; model 2: 0.731	Including environmental variables in the DEA model leads to an efficiency indicator that reflects the performance of PC units more properly.	Cordero-Ferrera, 2011
[26]	Mean TE: 0.88 (sd 0.14)	How practices are organized and how physicians are remunerated affect the costs and the efficiency associated with providing patient care. CHCs –community-oriented with a primary mandate to address the needs of disadvantaged populations– fare the worst when it comes to relative efficiency scores.	Milliken, 2011
[27]	Mean CRS TE ^D : 0.441 (sd 0.208); 0.3165 (sd 0.229); Mean VRS TE ^D : 0.708 (sd 0.226); 0.562 (sd 0.262)	Urban municipalities are more efficient than rural ones in delivering primary care services.	Ramírez-Valdivia, 2011
[28]	Mean TE: 0.455 (sd 0.246); Mean SE: 0.566 (sd 0.279)	Twenty-three out of the 359 (6.41 %) municipalities were considered efficient in the use of public funds in public PC. Partly the inefficiency was associated with range of action of municipalities (i.e. scale inefficiency).	Varela, 2010
[29]	Mean CRS TE: 0.673; Mean VRS TE: 0.771; Mean CRS SE: 0.873	Smaller PC centres appeared to suffer mostly from scale inefficiencies and larger ones from technical inefficiencies. Furthermore, IKA facilities evidently outperformed NHS ones.	Kontodimopoulos, 2007
[30]	Mean TE: 0.878; Mean AE: 0.913; Mean CE: 0.75; Mean CE: 0.876;	Single specialty practices were more efficient than multispecialty practices, irrespective of the focus of single specialty practices on PC or specialty care. The majority of the inefficiency was due to technical inefficiency. Larger practices are able to capture efficiencies of scope, but incur inefficiencies of scale.	Rosenman, 2004
[31]	TE range: 0.117–0.987 (7 best performers)	The key to increase efficiency may be the use and manage of a smaller number of inputs rather than focusing on expanding the size and complexity of practices and their total charges.	Andes, 2002
[32]	Mean TE: 1.203 (range 1.836–1.046)	The results do not indicate significant differences in the (in)efficiency of PC centres in the regions evaluated (i.e. La Rioja, Navarre and Alava)	Pinillos, 2002
[33]	Mean TE range: 0.66–0.812	Medium-sized PC centres were more efficient than larger ones, while small units showed lower efficiency ratings. A high level of efficiency of the health centres with laboratory activity (87 %) indicates an appropriate infrastructure model for the organization of PC services.	Zavras, 2002
[34]	Mean TE: 0.842 (sd 0.132); PTE: 0.914 (sd 0.114); SE 0.922 (sd 0.092)	There is a significant amount of sensitivity to model specification, pointing towards the importance of model validation.	Garcia, 1999
[35]	Mean CRS TE: 0.881 (0.885); Mean VRS TE: n.r	Differences in productive efficiency among Finnish health centres implied that the input saving potential of health centres would be at least 13 %.	Luoma, 1996
[36]	n.r	Thirteen out of 77 rural primary care programs were found to be inefficient.	Huang, 1989
[37]	Model 1 ^F : Mean TE: 0.83; Mean CE: 0.80 model 2 ^F : Mean TE: 0.75; Mean CE: 0.65	High variation in productive efficiency between health centres in different municipalities. The level of cost inefficiency was between 20 and 30 %. Changing the unit of output measurement from visit-to patient-based measures markedly affected the ranking of dental health centres.	Linna, 2003
[38]	Mean CRS: 0.635 (95 % CI 0.498–0.737); Mean VRS: 0.673 (95 % CI 0.326–0.858)	The relative efficiency of the CDS varies widely in England – on average the CDS is operating at 75 % of efficient levels compared to best-practice services.	Buck, 2000
[39]	Mean VRS TE: 0.86 (0.15)	35 of 96 practices were efficient based on the standard DEA model. The number of efficient practices decreased based on three restricted models that explored various behavioural preferences of physicians.	Testi, 2013
[40]	Mean TE: 0.87	Three out of the 14 practices were found technically efficient.	Amado, 2009
[41]	Mean TE: model 1 ^G 0.693 (95 % CI 0.61–0.67); model 2 ^G 0.742 (95 % CI 0.67–0.82)	Seven family units were technically efficient in providing services and nine in achieving health goals in treating patient with diabetes.	Salinas-Martínez, 2009
[42]	n.r	Seven out of the 107 general practices in the sample were technically efficient. Efficiency scores for a given practice tended to improve as the output formulation became more complex.	Bates, 1996

Table 2 (continued)

Ref.	Efficiency scores	Summary of study results	First author, year
[43]	Mean TE: 1.65 (range 1.06–2.90)	Municipalities were more efficient in their services production (37.8 %) than in their results production (16.6 %).	Rabetti, 2011
[44]	Mean CRS: 0.60 (sd 0.24); Mean output VRS: 0.69 (sd 0.24); Mean input VRS: 0.69 (sd 0.21); Mean input SE: 0.87 (sd 0.14); Mean input SE: 0.80 (sd 0.17)	Practice size, socioeconomic environment, service delivery process as well as regular administrative tasks are major cost and performance drivers for general practices in immunization activities.	Rouse, 2010
[45]	Mean TE: 0.796 (sd 0.198)	Of the 160 PCPs, 46 (28.8 %) were found to be relatively efficient. The remaining 114 (71.2 %) were classified as inefficient if compared to efficient PCPs.	Ozcan, 1998
[46]	n.r.	A set of inputs- outputs variables were advanced for perinatal care. The output set incorporates both activity levels and quality measures.	Thanassoulis, 1995
[47]	Inefficient physicians: 0.71	Inefficient physicians consumed significantly more resources and were 48 % more costly than efficient physicians.	Pai, 2000
[48]	Inefficient generalists 0.71; inefficient specialists 0.73	Substantial regional variation was mainly attributed to the differences in use of prescriptions and laboratory procedures. Relevant differences are not revealed in TE between specialist and generalists in treating patients with sinusitis, but specialists were found to be more costly than generalists.	Ozcan, 2000
[49]	Mean TE: 0.95	Six Northern Regions were identified as efficient using the best combinations of general practitioners to deliver a given level of GP outcomes.	Pelone, 2012
[50]	Mean TE 0.78	The results also showed that clinics that employed fewer physicians compared to all “licensed” practitioners were likely to be more efficient. In addition, providing transportation services to patients also enhanced clinics’ efficiency	Rahman, 2012
[51]	Mean TE: 0.844 (sd 0.69)	The health centre with the lowest efficiency score was the one with the lowest list size per doctor, as the DHA presenting the highest list size per doctor was likely to be technically efficient.	Amado, 2009
[52]	Mean CRS TE: 0.984 (sd 2.07); Mean VRS TE: 0.992 (sd 1.55)	In comparing DEA and SFA estimates of efficiency in PC, there are no general grounds for preferring one method over another. Nevertheless, the temporal robustness of the results of the SFA over the DEA suggest SFA to be more consistent.	Giuffrida, 2001
[53]	Mean PTE: 0.996 (range 0.952–1) ^H	The productivity increase over the period considered is attributed to pure technical efficiency improvement and positive change in scale efficiency, while the technology does not show significant change.	Giuffrida, 1999
[54]	TE range: 0.32–1	The lower efficiency the greater proportion of patients who were exempted from payment of prescription charges. A low correlation was found between efficiency and the number of temporary residents.	Bates, 1998
[55]	Output TE range: 0.732–1; Input TE range: 0.783–1.	Central role of quality indicators in PC is confirmed, by examining the extent to which data envelopment analysis (DEA) gives useful insights into FHSAs performance in terms of quality.	Salinas-Jiménez, 1996
[56]	TE range: 0.35–1	Efficiency ratings did not vary according to practice size for surgeries with a single location. For surgeries with multiple locations, larger surgeries were found to be on average more efficient.	Szczepura, 1993
[57]	TE range: 0.645–0.971 (11 best performers)	A method for the selection of model inputs and outputs, using a “stepwise approach”, and a sensitivity analysis using super-efficiency scores was advanced and validated. DEA was likely to carry advantages for physician profiling and usefully augments the current ratio based reports	Wagner, 2003
[58]	Mean TE: a) 0.92 0.89 0.90a; b) 0.86 0.82 0.82 a	A cone ratio DEA model would incorporate strategic thinking and executive accountability when establishing clinical benchmarks.	Chilingerian, 1997
[59]	TE range: 0.21–1	A multi-stage DEA technique was used to locate specific types of inefficient physicians. Among 326 subjects, specialists were found to have higher average costs than generalists and to be more efficient than high cost generalists.	Chilingerian, 1996

GP General practice, CDS Community dental service, CHCs Community health centres, FHSAs Family health service authorities, MHS National health system, IKA Social Security Foundation, SFA Stochastic frontier analysis, DEA Data envelopment analysis, PTE Pure technical efficiency, TE Technical efficiency, AE Allocative efficiency, CE Cost efficiency, SE Scale efficiency, VRS Variable returns to scale, CRS Constant returns to scale, sd Standard deviation, CI Confidence interval

^A The three DEA models differ each other for output variables incorporated in the analysis: 1: activities indicators; model 2: quality indicators; model 3: activity combined with quality indicators unrestricted and restricted; model 4: including environmental variables (4-stage DEA); ^B The two DEA models differ each other for the input-output variables incorporated in the analysis: model 1: basic DEA; model 2: including environmental variables; ^C Technical efficiency values refer to the fourth model reported in the study “Global with total costs”. They refer to the input orientation and | to the output orientation; ^D Technical efficiency values refer to the fourth model reported in the study “Global with total costs”. They refer to the urban and | to the rural municipalities; ^E The two DEA models differ each other for output variables were incorporated in the analysis—model 1: patient output DEA; model 2: visit output DEA; ^G The two DEA models differ each other for output variables were incorporated in the analysis—model 1: health care activities; model 2: quality indicators; ^H Technical efficiency values refer to 1990/91, the baseline year of the DEA analysis. The findings are not reported respect to the Malquist (productivity) analysis; ^I: a) cone ratio (bounded) model; b) standard (unbounded) model

Table 3 Recommended criteria for performing efficiency analyses in primary care using DEA

Analysis component	Recommended criteria	Reference
Defining the scope of the efficiency analysis	<p>Analyst should define a well-formulated research question in terms of:</p> <ul style="list-style-type: none"> - Setting/context of analysis –i.e.: whole spectrum of PC activities, specific PC activities (e.g. general practice), management of specific health problems; [53] - Level of organization under consideration –i.e.: PC professionals (e.g. general practitioners, ambulatory care specialists), PC practices or PC/outpatient programs delivered within vertical-integrated organizations, administrative entities (e.g. PC systems at a local, regional or national level); [7] - Analytical time frame–i.e. cross-sectional (in-time), or longitudinal (on-time) efficiency analysis - Perspective and target audience of the efficiency analysis –i.e. research, organization/managerial, regulator or society perspective <p>Define the technology assumption</p>	Table 1
DEA	<p>Analyst should opt a CRS or VRS accordingly to the technology that links the inputs to the outputs in the transformation process you are evaluating.</p> <p>Constant returns to scale may be assumed if there is a one-to-one, linear relationship between inputs and outputs. If an increase in a unit's inputs does not produce a proportional change in its outputs then the unit exhibits variable returns to scale. Finally to seek out scale efficiency (SE), both CRS and VRS DEA models are run on the same data, and any change in measured efficiency can be attributed to the presence of SE.</p> <p>In addition, consider the perspective and context of your analysis in deciding whether running a CRS or a VRS efficiency model.</p> <p>Define the efficiency analysis orientation</p>	Table 1 [7]
Model specifications	<p>Analyst should use an input or an output orientation based on what the managers of the organizations under analysis are able to control better (the resources or the outcomes of the production process). Largely, included studies examined the relative efficiency of PC DMUs by using an input orientation –by considering how much input quantities could be reduced while still maintaining the output target.</p> <p>Select and measure relevant input–output variables</p> <p>As for the output variables, the golden standard is to select relevant PC quality indicators when there is evidence that such measures are a proxy close to health improvements –rather than health care activities.</p> <p>The input side of efficiency analysis is usually considered less problematic than the output side –since the inputs can often be measured more accurately than outputs. Accordingly the scope of the DEA exercise, the Analyst has to consider two main issues for selecting an adequate set of input variables :</p> <ul style="list-style-type: none"> - Level of inputs disaggregation, e.g. total cost VS labour +capital costs - Measurement of selected input: monetary value VS cost terms <p>Incorporate (if appropriate) weight restrictions in valuing input–output variables</p> <p>Analyst should apply weight restrictions:</p> <ul style="list-style-type: none"> - To incorporate prior views on the value of individual inputs and outputs; - To relate the value of certain inputs and/or outputs; - To delineate how certain physicians' preferences (e.g. more drugs than visits or hospitalization) influence the healthcare delivery process [39] - To incorporate prior views on (in)efficient DMUs. <p>Most of the included PC DEA studies did not incorporate any restrictions to the input–output weights in specifying their efficiency models, partly to avoid compromising the objectivity of DEA and partly because of the lack of decision makers' involvement in the modelling process.</p>	Table 1 [5, 7, 13] Additional file 1 [44] [35, 53] [21, 47, 51] [39] [46, 58]
Presenting and interpreting findings	<p>the focus in the reporting of the empirical findings should be on transparency. Indeed, it is recommended putting such information in summary tables or to describe it in a narrative way.</p> <p>it is important to consider all implementation issues such as the feasibility, the suitability, and the relevance of adopting efficiency changes in the specific study context.</p> <p>in discussing results it is necessary to compare the obtained findings with those obtained in similar settings, to help policy makers in using of such evidence</p>	Table 1 [5, 13]
Addressing the uncertainty in modeling DEA	<p>Analyst should test a variety of model specifications using sensitive analysis to testify the robustness of the results, and construct data ranges within which results remain unchanged:</p> <ul style="list-style-type: none"> - discriminating the effect of atypical input–output combinations by means of a 'super-efficiency analysis' - performing statistical tests (i.e. Spearman correlation) to assure the existence of a trade-off between inputs and outputs in obtaining efficiency. - testing the robustness of the results obtained by running a number of DEA models with varying input- output variables. - carrying out a sensitivity analysis (i.e. bootstrapping procedure) of the DEA model, to construct confidence intervals of the efficiency scores 	[7, 57] [28, 57] [28, 55] [52] [41, 48]

Conclusions

We found 39 studies that have used DEA to compare the efficiency of primary health care providers, a few of which provide end users with insight on the implementation process or on the feasibility of adopting efficiency changes in a determined policy context, and almost all studies are cross sectional in nature rather than longitudinal. Given the current state of the art, the overall performance of DEA studies within PC to inform health policy is far from optimal and could be further improved given the potential is shown to provide evidence-based health policy information.

The included DEA studies in PC revealed a consistent difference in terms of empirical findings. This diversity could be explained partly by different methodological constructs (i.e. input–output variables selection and other model specifications such as the technology assumption and the weight restrictions), and partly by a wide range of evaluation contexts (i.e. type of PC organizations, level of analysis and time frame) in composing the study designs.

A number of studies aimed to support policymakers and practice managers in improving the efficiency of their PC organizations. However, still a part of the studies provided findings aimed to improve the theoretical understanding in the empirical use of DEA and were mainly targeted at researchers and health economists. This indicates that DEA—at least when applied to PC—is a methodology still in progress and needs to be further advanced and standardized to meet the complexity that characterizes the production of PC outcomes.

DEA studies in PC put very little effort in enhancing the accuracy of study findings by assessing the sensitivity of the results to different model specifications. Few studies have adopted approaches to improve the model specifications—i.e. selection of input and outputs; weight restrictions to incorporate exogenous value judgments in modelling DEA; super-efficiency model to deal with outliers observations; bootstrapping technique to make the model more robust—aimed to provide decision makers with more reliable information.

This systematic review contributes to the ability of stakeholders involved in PC performance improvement in interpreting understanding the strengths and limitations of DEA. Another important strength is the extensive scope and all-inclusive approach of this study, bridging the gap between health economists, health services researchers and PC decision makers. Although the DEA literature on efficiency analysis in healthcare has focused mainly on comparing the efficiency of hospitals and nursing homes rather than PC providers, several of our findings were also confirmed by such studies [5, 10–14].

However, this literature review has two limitations. First, it was not possible to search the “grey literature” for additional DEA studies in PC, as well the review process is also open to

publication bias, as we have only included the published articles. We tried to overcome this limitation by refining our search with the ‘related articles’ search on Pubmed, supplemented with hand searches of the references of related articles. The second limitation is the quality of the original studies. Due to an absence of quality tools, it is currently not possible to evaluate whether the included efficiency studies assessed the scientific soundness of their measures.

Recommendations

Based on our systematic review and a number of existing methodological guidelines [5, 7, 13], we developed several points of importance for performing DEA studies in the primary health care context (see Table 3). In addition to including these parameters, DEA applications should be transparent in the applied methodology and results, and be reproducible. It is recommended that researchers undertake joint efforts to improve the consistency and appropriateness of efficiency methodologies and to optimize the utility for the consumer. Further research is needed to fill the gaps in some measurement areas, such as on the use of health outcomes as outputs or capital resources as inputs, on the causes of productivity change and its decomposition, and to identify the determinants of improved performance of PC services delivery. Furthermore, improving the application of sensitivity tests is imperative to analyze how uncertainty in the efficiency results can be attributed to different variations in DEA model specifications.

For policymakers and clinical managers it is recommended to integrate efficiency evidence with other relevant factors of the overall performance of their organizations, such as equity, responsiveness and effectiveness. All stakeholders involved in PC decision making could benefit from incorporating regular efficiency measurements using DEA into their organizations, and from playing an active role in the DEA modeling phase to facilitate the relevance and the implementation of the results into practice.

Conflicts of interest None declared. The authors have no current external funding sources for this study. FP completed the work as part of a PhD’s degree.

References

1. WHO (2000) The World Health Report 2000. Health Systems: Improving Performance. Geneva, Switzerland: World Health Organization.
2. WHO (2008) The World Health Report 2008. Primary health care - Now more than ever. Geneva, Switzerland: World Health Organization.

3. Kringos DS, Boerma WGW, Hutchinson A, van der Zee J, Groenewegen PP (2010) The breadth of primary care: a systematic literature review of its core dimensions. *BMC Health Serv Res* 13:10:65. doi: [10.1186/1472-6963-10-65](https://doi.org/10.1186/1472-6963-10-65).
4. Starfield B (1992) Primary care: concept, evaluation, and policy. New York: Oxford University Press.
5. Amado CA, Dyson RG (2008) On comparing the performance of primary care providers. *Eur J Oper Res* 185: 915–932. doi: [10.1016/j.ejor.2006.02.052](https://doi.org/10.1016/j.ejor.2006.02.052).
6. Hussey PS, de Vries H, Romley J, Wang MC, Chen SS et al. (2009) A systematic review of health care efficiency measures. *Health Serv Res* 44:784–805. doi: [10.1111/j.1475-6773.2008.00942.x](https://doi.org/10.1111/j.1475-6773.2008.00942.x).
7. Jacobs R, Street A, Smith PC (2006) *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. New York: Cambridge University Press.
8. Kringos DS, Boerma WGW, Bourgueil Y, Cartier T, Hasvold T et al. (2010) The European primary care monitor: structure, process and outcome indicators. *BMC Fam Prac* 27:81. doi: [10.1186/1471-2296-11-81](https://doi.org/10.1186/1471-2296-11-81).
9. Schäfer W, Groenewegen PP, Hansen J, Black N (2011) Priorities for health services research in primary care. *Qual Prim Care* 19:77–83.
10. Hollingsworth B, Street A (2006) The market for efficiency analysis of health care organisations. *Health Econ* 15:1055–1059. doi: [10.1002/hec.1169](https://doi.org/10.1002/hec.1169).
11. Hollingsworth B, Dawson PJ, Maniadakis N (1999) Efficiency measurement of health care: a review of non-parametric methods and applications. *Health Care Manag Sci*. 2: 161–172. doi: [10.1023/A:1019087828488](https://doi.org/10.1023/A:1019087828488).
12. Hollingsworth B (2003) Non-parametric and parametric applications measuring efficiency in health care. *Health Care Manag Sci* 6:203–218. doi: [10.1023/A:1026255523228](https://doi.org/10.1023/A:1026255523228).
13. Hollingsworth B (2008) The measurement of efficiency and productivity of health care delivery. *Health Econ* 17:1107–1128. doi: [10.1002/hec.1391](https://doi.org/10.1002/hec.1391).
14. Puig-Junoy J (2000) Efficiency in primary health care: a critical review of frontier measures. *Rev Esp Salud Publica* 74: 483–495.
15. Worthington AC (2004) Frontier efficiency measurement in health care: a review of empirical techniques and selected applications. *Medical Care Res Rev* 61:135–70. doi: [10.1177/1077558704263796](https://doi.org/10.1177/1077558704263796).
16. Smith P (1997) Model misspecification in data envelopment analysis. *Ann Oper Res* 73:233–252. doi: [10.1023/A:1018981212364](https://doi.org/10.1023/A:1018981212364).
17. World Bank. Country and Lending Groups. Available:<http://data.worldbank.org/about/country-classifications/country-and-lending-groups>. Accessed 25 November 2013.
18. WONCA (1991) The Role of the General Practitioner/Family Physician in Health Care Systems; A statement of the World Organisation of National Colleges, Academies and Academic Association of General Practitioners/Family Physicians (Wonca). Victoria: WONCA
19. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 21:e1000097. doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097).
20. Kiadaliri AA, Jafari M, Gerdtham UG (2013) Frontier-based techniques in measuring hospital efficiency in Iran: a systematic review and meta-regression analysis. *BMC Health Serv Res* 13:312. doi: [10.1186/1472-6963-13-312](https://doi.org/10.1186/1472-6963-13-312).
21. Cordero Ferrera JM, Crespo Cebada E, Murillo Zamorano LR (2014) The effect of quality and socio-demographic variables on efficiency measures in primary health care. *Eur J Health Econ* 15:289–302. doi: [10.1007/s10198-013-0476-1](https://doi.org/10.1007/s10198-013-0476-1)
22. Deidda M, Lupiáñez-Villanueva F, Codagnone C, Maghiros I (2014) Using data envelopment analysis to analyse the efficiency of primary care units. *J Med Syst*. 38:122. doi: [10.1007/s10916-014-0122-1](https://doi.org/10.1007/s10916-014-0122-1)
23. Ferreira C, Marques RC, Nicola P (2013) On evaluating health centers groups in Lisbon and Tagus Valley: efficiency, equity and quality. *BMC Health Serv Res* 13:529. doi: [10.1186/1472-6963-13-529](https://doi.org/10.1186/1472-6963-13-529)
24. Pelone F, Kringos DS, Spreeuwenberg P, De Belvis AG, Groenewegen PP (2013) How to achieve optimal organization of primary care service delivery at system level: lessons from Europe. *Int J Qual Health Care* 25:381–93. doi: [10.1093/intqhc/mzt020](https://doi.org/10.1093/intqhc/mzt020).
25. Cordero-Ferrera JM, Crespo-Cebada E, Murillo-Zamorano LR (2011) Measuring technical efficiency in primary health care: the effect of exogenous variables on results. *J Med Syst* 35:545–54. doi: [10.1007/s10916-009-9390-6](https://doi.org/10.1007/s10916-009-9390-6)
26. Milliken O, Devlin RA, Barham V, Hogg W, Dahrouge S et al. (2011) Comparative efficiency assessment of primary care service delivery models using data envelopment analysis. *Can Public Policy* 37:85–109. doi: [10.3138/cpp.37.1.85](https://doi.org/10.3138/cpp.37.1.85).
27. Ramírez-Valdivia MT, Maturana S, Salvo-Garrido S (2011) A multiple stage approach for performance improvement of primary healthcare practice. *J Med Syst* 35:1015–28. doi: [10.1007/s10916-010-9438-7](https://doi.org/10.1007/s10916-010-9438-7)
28. Varela PS, Martins GA, Fávoro LP (2010) Production efficiency and financing of public health: an analysis of small municipalities in the state of São Paulo – Brazil. *Health Care Manag Sci* 13:112–123. doi: [10.1007/s10729-009-9114-y](https://doi.org/10.1007/s10729-009-9114-y).
29. Kontodimopoulos N, Moschovakis G, Aletras VH, Niakas, D (2007) The effect of environmental factors on technical and scale efficiency of primary health care providers in Greece. *Cost Effectiveness and Resource Allocation* 17:5–14. doi: [10.1186/1478-7547-5-14](https://doi.org/10.1186/1478-7547-5-14).
30. Rosenman R, Friesner D (2004) Scope and scale inefficiencies in physician practices. *Health Econ* 13:1091–1116. doi: [10.1002/hec.882](https://doi.org/10.1002/hec.882). PubMed: 15386682.
31. Andes S, Metzger LM, Kralewski J, Gans D (2002) Measuring Efficiency of Physician Practices Using Data Envelopment Analysis. *Manag Care* 11:48–54.
32. Pinillos M, Antoñanzas F (2002) Primary care: decentralization and efficiency. *Gac Sanit* 16:401–407. doi: [10.1590/S0213-91112002000500006](https://doi.org/10.1590/S0213-91112002000500006).
33. Zavras AI, Tsakos G, Economou C, Kyriopoulos J (2002) Using DEA to Evaluate Efficiency and Formulate Policy Within a Greek National Primary Health Care Network. *J Med Syst* 26:285–292. doi: [10.1023/A:1015860318972](https://doi.org/10.1023/A:1015860318972).
34. Garcia F, Marcuello C, Serrano D, Urbina O (1999) Evaluation of efficiency in primary health care centres: an application of data envelopment analysis. *Fin Account Manage* 15:67–83 doi: [10.1111/1468-0408.00074](https://doi.org/10.1111/1468-0408.00074)
35. Luoma K, Jarvio ML, Suoniemi I, Hjerpe RT (1996) Financial incentives and productive efficiency in Finnish health centres. *Health Econ* 5:435–445. doi: [10.1002/\(SICI\)1099-1050\(199609\)5:5<435::AID-HEC223>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1099-1050(199609)5:5<435::AID-HEC223>3.0.CO;2-Y).
36. Huang YGL, McLaughlin CP (1989) Relative Efficiency in Rural Primary Health Care: An Application of Data Envelopment Analysis. *Health Serv Res* 24:143–58.
37. Linna M, Nordblad A, Koivu, M (2003) Technical and cost efficiency of oral health care provision in Finnish health centres. *Soc Sci Med* 56:343–353. doi: [10.1016/S0277-9536\(02\)00032-1](https://doi.org/10.1016/S0277-9536(02)00032-1).
38. Buck D (2000) The efficiency of the Community Dental Service in England: a data envelopment analysis. *Community Dent Oral Epidemiol* 28:274–280. doi: [10.1034/j.1600-0528.2000.280405.x](https://doi.org/10.1034/j.1600-0528.2000.280405.x).
39. Testi A, Fareed N, Ozcan YA, Tanfani E (2013) Assessment of physician performance for diabetes: a bias-corrected data envelopment analysis model. *Qual Prim Care* 21:345–57.
40. Amado CA, Dyson RG (2009) Exploring the use of DEA for formative evaluation in primary diabetes care: An application to compare English practices. *J Oper Res Soc* 60:1469 –1482.
41. Salinas-Martinez AM, Amaya-Alemán MA, Arteaga-Garcia JC, Nunez-Rocha GM, Garza-Elizondo ME (2009) Technical efficiency in primary care for patients with diabetes. *Salud Publica Mex*. 51:48–58. doi: [10.1590/S0036-36342009000100010](https://doi.org/10.1590/S0036-36342009000100010).

42. Bates JM, Baines DL, Whyne DK (1998) Assessing efficiency in general practice: an application of data envelopment analysis. *Health Serv Manage Res* 11:103–8. doi:10.1177/095148489801100204.
43. Rabetti ADC, Freitas SFTD (2011) Evaluation of actions concerning systemic arterial hypertension in primary healthcare. *Rev Saúde Pública* 45:258–268. doi: 10.1590/S0034-89102011005000007.
44. Rouse P, Harrison J, Turner N (2010) Cost and Performance: Complements for Improvement. *J. Med. Syst.* 35: 1063–1074. doi: 10.1007/s10916-010-9520-1.
45. Pai CW, Ozcan YA, Jiang HJ (2000) Regional Variation in Physician Practice Pattern: An Examination of Technical and Cost Efficiency for Treating Sinusitis. *J Med Syst* 24:103–117. doi: 10.1023/A:1005568814163.
46. Ozcan YA (1998) Physician benchmarking: measuring variation in practice behavior in treatment of otitis media. *Health Care Manag Sci* 1:5–17. doi: 10.1023/A:1019026114859.
47. Thanassoulis E, Boussofiane A, Dyson RG (1995) Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. *Eur J Oper Res* 80:588–607. doi: 10.1016/0377-2217(94)00139-4.
48. Ozcan YA, Jiang HJ, Pai CW (2000) Do primary care physicians or specialists provide more efficient care? *Health Serv Manage Res* 13: 90–96. doi: 10.1177/095148480001300203.
49. Pelone F, Kringos DS, Valerio L, Romaniello A, Lazzari A et al. (2012) The measurement of relative efficiency of general practice and the implications for policy makers. *Health Policy* 107:258–68. doi: 10.1016/j.healthpol.2012.05.005.
50. Rahman M, Capitman JA (2012) Can More Use of Supporting Primary Care Health Practitioners Increase Efficiency of Health Clinics? Evidence from California's San Joaquin Valley. *J Health Care Finance* 38:78–92.
51. Amado CA, Santos SP (2009) Challenges for performance assessment and improvement in primary health care: The case of the Portuguese health centres. *Health Policy* 91:43–56. doi: 10.1016/j.healthpol.2008.11.008.
52. Giuffrida A, Gravelle H (2001) Measuring performance in primary care: econometric analysis and DEA. *Appl Econ* 33:163–175. doi: 10.1080/00036840122522
53. Giuffrida A (1999) Productivity and efficiency changes in primary care: a Malmquist index approach. *Health Care Manag. Sci.* 2, 11–26. doi: 10.1023/A:1019067223945
54. Bates JM, Baines D, Whyne DK (1996) Measuring the Efficiency of Prescribing by General Practitioners. *Journal of the Operational Research Society*.47:1443–1451. doi: 10.1057/jors.1996.188.
55. Salinas-Jimenez J, Smith P (1996) Data envelopment analysis applied to quality in primary health care. *Ann Oper Res* 67:141–161. doi: 10.1007/BF02187027
56. Szczepura A, Davies C, Fletcher J (1993) Efficiency and Effectiveness in General Practice. *J Manage Med* 7:36–47. doi: 10.1108/eb060565
57. Wagner JM, Shimshak DG, Novak MA (2003) Advances in physician profiling: the use of DEA. *Socio Econ Plan Sci* 37:141–163. doi: 10.1016/S0038-0121(02)00038-1
58. Chilingirian JA, Sherman HD (1997) DEA and primary care physician report cards: Deriving preferred practice cones from managed care service concepts and operating strategies. *Ann Oper Res* 73:35–66. doi: 10.1023/A:1018993515090
59. Chilingirian JA, Sherman HD (1996) Benchmarking physician practice patterns with DEA: A multi-stage approach for cost containment. *Ann Oper Res* 67:83–116. doi: 10.1007/BF02187025
60. Allen R, Athanassopoulos A, Dyson RG, Thanassoulis E (1997) Weights Restrictions and Value Judgments in DEA: Evolution, Development and Future Directions. *Ann Oper Res* 73:13–34. doi: 10.1023/A:1018968909638
61. Cooper WW, Seiford LM, Tone K (2000) *Data envelopment analysis: a comprehensive text with models, applications, references, and DEA-Solver software.* Boston, Kluwer Academic Publishers.
62. Murillo-Zamorano LR (2004) Economic Efficiency and Frontier Techniques. *J Econ Survey* 18:33–77. doi: 10.1111/j.1467-6419.2004.00215.x.
63. Pedraja-Chaparro F, Salinas-Jimenez J, Smith P (1999) On the quality of the data envelopment analysis model. *The Journal of the Operational Research Society* 50 (6), 636 – 644. doi: 10.1057/palgrave.jors.2600741.
64. Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. *Manage Sci* 39:1261–4. doi: 10.1287/mnsc.39.10.1261.