Journal of Research Administration
Society of Research Administrators International

IMPLICATIONS OF MODEL CHOICE ON THE ESTIMATION OF EFFICIENT FRONTIERS: A SYSTEMATIC LITERATURE REVIEW AND META-REGRESSION ANALYSIS

AlJawhara M. AlSabah, DrPH, CPH [1*]
[1]New York Medical College School of Health Sciences and Practice & Institute of Public Health, Valhalla, NY 10595
*Corresponding Author: aljawhara_alsabah@alumni.nymc.edu
ORCID 0009-0000-8851-0993

Abstract

The aim of this meta-regression literature review is to examine the effect of modeling choices on technical efficiency scores within the econometric literature based on fundamental concepts rooted in the operations and production management sciences. Building on key modeling frameworks of efficiency analytics and diverse frontier methodologies relevant for use in health services research, the focus of this paper is on the major considerations following the selection of nonparametric data envelopment analysis (DEA) variables/hospital indicators and the choices applied to this estimation technique, accounting for model specification and variables included in the efficiency analysis. The review concludes with an empirical section containing a statistical summary of the literature on hospital efficiency frontier modeling, as well as a meta-regression analysis aimed at identifying the key factors of DEA model specifications or study characteristics that influence efficiency estimates. This step, undeniably, is vital in understanding the limitations and drivers of efficiency score estimations and avoiding certain pitfalls that can reduce the robustness of frontier modeling methods. Based on the meta-regression analysis, it is clear that as the number of variables included in the frontier model increases, the average efficiency predictions drop fairly rapidly when the sample size is fairly small. This significant phenomenon identifies the importance of the sample size effect, indicating that the inclusion of an extra variable into a model with more than 10 (hospital) observations does not alter the average efficiency score very much; as long as the hospital sample size is large and homogenous, the mean technical efficiency shows little change, and the mean efficiency seems to remain constant after a sliding threshold is reached. Therefore, correcting for sample size has a major impact on the assessment of average efficiency estimates and aids in robust empirical analysis that can potentially be deemed scientifically evidence-based to allow for future health policy reforms and allocation redistribution decisions to be made.

Keywords: Efficiency models, health services research, DEA, nonparametric analysis, frontier analysis, efficiency literature review, public health efficiency, hospital productivity.

© 2023The Authors

Introduction

Frontier-based methods for analyzing and estimating efficiency have been applied to many different types of healthcare institutions, including nursing homes, hospitals, health districts/regions, and physician practices. Parametric methods (stochastic frontier analysis; SFA) have gained popularity in recent years, while nonparametric methods (data envelopment analysis; DEA) have long been the dominant tool in this body of literature [1, 2]. A majority of studies utilize efficiency estimates to shed light on policy issues such as ownership and organizational structure [3]. Therefore, this empirical analysis of the literature and meta-regression is necessary to determine how best to model variables and capture measures of hospital efficiency to (i) inform health officials and policy-makers on healthcare funding or care service delivery and capacity utilization and (ii) conduct a post evaluation of healthcare policy effects on hospital behaviors by comparing preintervention baseline measures of hospital efficiency performance and evaluating any improvement/deterioration in efficiency, following policy changes or intervention implementation, to determine their success or failure to meet initiative goals.

The influence of modeling choice on efficiency estimates is widely acknowledged in the efficiency literature. Although most studies do not have a choice in either the sample size or variables used due to data availability, the decision on analytical methods and model specifications, to a larger extent, can be controlled to accommodate the research question(s). Therefore, there are good reasons for examining alternative model specifications and their results to ensure the reliability of the estimation. This is especially important for studies with a policy design focus, as other health economists have pointed out in earlier studies [4-6]. If the efficiency estimates are to inform decision makers on funding or capacity utilization, then incorrectly labeled inefficient hospitals might receive less funding resources or need to trim their production. If post evaluation of a healthcare policy on hospital behaviors is the issue of concern, a biased estimation of efficiency would be misleading to assess the true policy impacts.

Key Modeling Choices and Considerations for Efficiency Estimation Improvement

Variables

The first major decision in modeling production technology relates to output and input choices. Inputs and outputs should be relevant and sufficient to capture the production process. In practice, problems with variable choice come under the form of imperfect measures of inputs and/or outputs, incorrect aggregation, and omitted variables. The inclusion of irrelevant variables is another issue [2, 7]. However, in the hospital efficiency literature, it is far more often that a frontier model fails to capture all aspects of healthcare service production than including an extraneous variable, mainly because of data deficiency. Furthermore, it is suggested that exclusion of relevant variables is likely to be more damaging to frontier models than inclusion of irrelevant variables [8-10]. Although studies far too often do not have choice over quality of input and output data, it is worth emphasizing that findings based on rudimentary measures of inputs/outputs should be interpreted with caution. Omitted variables and aggregation in many situations are mainly attributed to different research questions or data availability, while in other cases, they are due to modeling choice. Its existence usually distorts findings.

Now, let us consider the question of whether it is possible to predict the direction of impact on the average efficiency score by variable inclusion or exclusion in a model. Basically, what considerations should be taken when deciding which hospital variables to include in the efficiency model and what to expect in terms of overestimation of efficiency scores or underestimation of efficiency scores based on variable selection alone? Technically, the inclusion of another variable in the estimated model will increase the dimensions of the frontier. We know this from previous studies that suggest that increasing frontier dimensions may produce higher mean efficiency scores (overestimated efficiency). The magnitude of this effect, however, depends on the omitted variable's correlations with included variables. This means, for example, if the extra variable is an input and it is highly correlated to other input variables, omission of the variable is unlikely to significantly affect the results. However, if the variable is not strongly correlated, then the impact on mean efficiencies can be significant. This is a solid point of consideration when constructing models for hospital efficiency analysis.

One strong example found in the hospital efficiency literature is the study by [11], in which they added case-mix variables to a basic trans log function and found the basic trans log case yielded lower efficiency scores compared to the one with case-mix variables. In fact, the potential impact of dimensionality on efficiency scores was discussed by [12], where the author found that variable set expansion, either through adding new variables or even disaggregating existing variables, may produce an upward trend in mean efficiency scores (important consideration when attempting to transform or decompose data measures originally reported as aggregate hospital variables). Then, again, another study also confirmed that aggregation of many outputs into fewer or one output introduces a downward bias on efficiency estimates, and the more outputs are aggregated, the greater the bias that may be expected [13]. In essence, the modeling process alone is arguably the ultimate determining factor of whether the efficiency scores obtained from our analysis are truly representative of a hospital's efficiency.

Sample Size

In terms of the sample size, number of observations, or decision-making units (DMUs) in an analysis, the opposite effect is generally observed. It seems that the increase in sample size will either push the production frontier up when new observations form part of the new frontier or does not change the frontier at all when new observations lie entirely under the existing frontier [14]. This means that if we decide to include more hospitals to have a larger sample size and these new observations form part of the new frontier, then the units (hospitals) that were once identified as efficient under the old frontier may now be identified as inefficient. Only when a new observation in our sample does not affect the position of the frontier (because it is either on or below the existing frontier) will it not change the status of already identified efficient and inefficient units (hospitals). Therefore, this consideration of the number of hospital units in an efficiency analysis indicates a point of depreciation so that, on average, increasing the sample size is unlikely to result in an increase in mean efficiency scores. This important point was also noted by [15, 16], who both found a negative correlation between the estimated mean efficiency and the number of firms in the industry. The key takeaway here is that when the sample is relatively small, the mean efficiency decreases quickly as the number of observations increases. When sample sizes are large, the mean efficiency shows little change. Above a threshold, a mean efficiency seems to be constant.

Orientation

The other important consideration in frontier modeling relates to orientation. The choice of input/output orientation is usually driven by the objective of production units under relevant production and management constraints. In our case of hospital efficiency, hospitals under an expenditure cap scheme tend to maximize output, while hospitals receiving reimbursement based on units of treatment appear to conserve cost. If maximizing output (or outcome) is considered a relevant objective of a hospital, then an output orientation (output-oriented DEA frontier) may be warranted. Alternatively, if the hospital is interested in minimizing inputs or cost, then input-oriented DEA may be selected. In practice, the underlying assumption of input orientation in hospital efficiency studies is that cost (input) minimizes the behavior of hospitals. We assume this and justify it from the viewpoint of hospital managers, who are constantly under the pressure of meeting a budget requirement. However, this assumption has received much criticism in the literature, especially from medical professionals who often argue that their objective is not minimizing cost but improving lives through prevention and treatment of diseases.

Furthermore, orientation has a certain effect on the efficiency score. If the sample in the analysis contains mainly small and few large hospitals, it is expected that most hospitals are operating in the increasing returns to scale region, and therefore an input-oriented DEA approach would produce a higher efficiency level for small hospitals; consequently, higher mean efficiency. The reverse applies to samples with mainly large hospitals. A sample with a balanced mix of hospital sizes is likely to generate similar mean efficiency scores under either output or input orientations. It is noted that this issue only applies for the variable returns to scale (VRS) frontier. In the constant returns to scale (CRS) circumstance, output and input orientations produce identical technical efficiency [17]. This point of consideration indicates that adjustments should be made in terms of the total number of hospital beds (hospital size) to control for higher exaggerated efficiency estimates in a sample of small, medium, and large hospital sizes.

© 2023The Authors

Return to Scale

In economics, returns to scale describe what happens to long-run returns as the scale of production increases, when all input levels including physical capital usage are variable (able to be set by the firm). The concept of returns to scale arises in the context of a firm's production function and relates to whether production units are of the optimal size. This is one of the popular research questions in efficiency analysis. Some production technologies possess the property of constant returns to scale (CRS), and the production size does not matter. Others (and the majority) do not. This raises the question of how returns to scale should be modeled. The CRS assumption is appropriate when all hospitals are operating at the optimal scale (i.e., productivity is scale dependent). However, imperfect competition, government regulations, valid social objectives, and financial and labor constraints may cause the hospital to not operate at the optimal scale [17].

In this case, if we impose CRS in the model, efficiency estimates will be significantly biased. This bias is generally more serious than in the case where VRS is assumed for a CRS technology [18]. Moreover, [19] suggests that imposing CRS will vastly underestimate efficiency, whereas [20] imply that this inappropriate use of returns to scale assumption is particularly damaging when the sample size is small. Table 1 summarizes this section on the expected relationships between efficiency scores and choice of model specifications based on the literature outlined above. The following section statistically evaluates differences in study characteristics, model specifications, and reported efficiency scores in a sample of reviewed studies.

Table 1: *Some expected impacts of modeling choices on mean efficiency estimates*

Factors which push mean efficiency upwards	**Factor with ambiguous impact on mean efficiency**	**Factors which push mean efficiency downwards**
Number of variables	Orientation	Sample size
Pooled panel data		Constant returns to scale (CRS)

Research Methodology

Priority was given for constructing a meta-dataset of reviewed studies and declaring the meta-analysis regression (meta-regression) data in Stata 16.1 statistical package, which is used to perform the meta-regression analysis on study-level summary data. Therefore, most of the time was spent during this phase of the process to ensure that a diverse and globally representative study sample with quality-appropriate findings was finalized. The literature search query applied used the initial “hits” or returns from the many databases searched to review academic and scholarly peer-reviewed studies following specific, predetermined inclusion/exclusion criteria.

© 2023The Authors

Primary studies are then sorted according to keyword usage in the title or abstract with inclusion terms: “efficiency”, “hospital”, “healthcare facility”, “health center”, “data envelopment analysis”, and “DEA”, which helped identify approximately 1,286 publications on healthcare/hospital efficiency [21]. This review followed the recommendations and guidance of the Joanna Briggs Institute (JBI) Global; thus, all methods, including protocol, study selection, critical appraisal, data extraction and synthesis, are supported and approved by the international scientific committee (https://auth.jbisumari.org).

Elimination was based on two main publication factors, studies dated before 2002 (older than twenty 20 years) and non-English language studies; broader exclusion criteria included studies not reporting weighted average technical efficiency scores (or no data given for readers to calculate), no full-text available or accessible, DMUs other than hospitals (i.e., cross-country health system comparisons, health districts/regions, etc.), literature reviews and all other types of study analyses (graduate thesis, dissertations, etc.), studies using cost function to estimate productivity along with efficiency measures, and any hospital efficiency study using parametric stochastic frontier analysis (SFA) and not nonparametric data envelopment analysis (DEA) as the primary technique for frontier estimation. All working papers were additionally eliminated. Finally, studies on allocative efficiency and healthcare services other than hospitals (such as physicians, primary health clinics, diagnostic labs, nursing homes, etc.) were removed from the list.

The screening of identified articles and assessment process for eligibility is based on quality indicators set using the fifteen-point scale by [22] for quality appraisal of the literature. The criteria used to assess study quality include the following: literature review and identification of research gaps; research question and design; validity and reliability; data collection; population and sampling; and analysis and reporting of results. These criteria were rated as 0 (not present or reported), 1 (present but low quality), 2 (present and mid-range quality), or 3 (present and high quality). Studies are rated using the article quality rating sheet [22]. Given the specificity of this review, several papers were continuously eliminated following low-quality ratings, and a few concerns regarding diversity of papers were raised; therefore, to maintain research inclusion while controlling for publication bias, papers rated as present and mid-range quality in their data analysis and interpretation of results were still included.

Articles were then evaluated to determine their relevance and authority in terms of the research question(s) and study objectives, as well as other characteristics, units of analysis, country/region, publication year, methodology, model specifications, and efficiency results and findings – among other exogenous variables (not shown) that were included in the meta-regression based on approaches and model specifications in the primary studies, such as cross-section versus panel data and sample heterogeneity. Most relevant primary studies were identified as journal publications. Studies using any of the different forms of DEA (two-staged Tobit, two-staged Malmquist Productivity Index, bootstrap DEA, dynamic network DEA, etc.)

© 2023The Authors

were enough to warrant their inclusion if the technical efficiency scores were clearly reported. The final meta-dataset consisted of 47 peer-reviewed studies from 27 different countries and published within the past two decades with a median publication year in 2015; the sample size ranged from five (5) decision-making units (DMUs) to 1,259 hospitals. The number of input/output variables were among the model specifications recorded, as well as ownership and hospital type(s) under analysis and orientation choice and returns to scale model used. A summary of the different efficiency frontier characteristics in DEA studies of hospital efficiency is shown in Table 2.

Table 2: *Characteristics of reviewed hospital efficiency studies using DEA*

Article No.	Study	Publication Year	Country	Sample Size	No. of Inputs	No. of Outputs	Orientation	Return to Scale	Hospital Type
1	Stefko et al.	2018	Slovakia	8	3	2	Output-oriented	VRS & CRS	Regional public healthcare facilities
2	Cheng et al.	2015	China	114	3	2	Input-oriented	VRS & CRS	County hospitals
3	Lin et al.	2021	Taiwan	19	5	6	Input-oriented	VRS & CRS	MoH tertiary hospitals
4	Torabipour et al.	2014	Iran	12	3	3	Input-oriented	VRS	University teaching & nonteaching hospitals
5	Jehu-Appiah et al.	2014	Ghana	128	4	4	Output-oriented	VRS	Mixed-ownership district hospitals
6	Ahmed et al.	2019	Bangladesh	62	2	3	Input-oriented	VRS & CRS	Public district hospitals

7	Campanella et al.	2017	Italy	50	3	3	Input-oriented	CRS	Public hospital trusts
8	Kalhor et al.	2016	Iran	54	4	4	Input-oriented	VRS	General hospitals
9	Jat & Sebastian	2013	India	40	3	8	Input-oriented	VRS	District hospitals
10	Yusefzadeh et al.	2013	Iran	23	3	2	Input-oriented	VRS	Public hospitals
11	Masiye	2007	Zambia	30	4	4	Input-oriented	VRS	Public hospitals
12	Dash et al.	2010	India	29	4	5	Input-oriented	VRS	District hospitals
13	Shahhoseini et al.	2011	Iran	12	4	5	Input-oriented	VRS & CRS	Provincial hospitals
14	Farzianpour et al.	2012	Iran	16	3	3	Input- and output-oriented	VRS & CRS	University teaching hospitals
15	Li & Dong	2015	China	14	2	2	Output-oriented	CRS	Public hospitals
16	Medarevic & Vukovic	2021	Serbia	39	3	2	Input-oriented	VRS & CRS	Public general hospitals
17	Xu et al.	2015	China	51	4	3	Input-oriented	CRS	Provincial tertiary hospitals
18	Lobo et al.	2016	Brazil	31	3	1	Output-oriented	VRS	University teaching hospitals
19	Muja	2016	Ugan	18	2	2	Output-	VRS	District

© 2023The Authors

	si et al.		da				oriented		hospitals
20	Flokou et al.	2017	Greece	73	3	3	Input-oriented	VRS & CRS	Public hospitals
21	Jia & Yuan et al.	2017	China	5	2	3	Output-oriented	VRS	Public hospitals
22	Li et al.	2017	China	12	4	3	Input-oriented	VRS & CRS	Public hospitals
23	Giancotti et al.	2018	Italy	41	2	3	Input-oriented	VRS & CRS	Public hospitals
24	Alsabah et al.	2019	Kuwait	15	4	2	Input-oriented	VRS	MoH hospitals
25	Franco Miguel et al.	2019	Spain	25	3	4	Input-oriented	CRS	Mixed-managed public–private hospitals
26	Alatawi et al.	2020	Saudi Arabia	91	4	6	Input-oriented	VRS & CRS	MoH hospitals
27	Hofmarcher et al.	2002	Australia	93	4	2	Input-oriented	VRS	Provincial hospitals
28	Ramanathan	2005	Oman	20	3	3	Input-oriented	VRS & CRS	MoH & public regional hospitals: university teaching hospital, police hospital

© 2023The Authors

29	Mahate & Hamidi	2016	UAE	96	6	3	Output-oriented	VRS & CRS	Private & government hospitals
30	Mogha et al.	2012	India	55	3	1	Output-oriented	VRS & CRS	Private hospitals
31	Sultan & Crispim	2016	Jordan	27	4	3	Input-oriented	VRS & CRS	MoH public hospitals
32	Kontodimopoulos et al.	2006	Greece	17	3	2	Input-oriented	CRS	Rural small-scaled hospitals
33	Wei et al.	2011	Taiwan	21	2	3	Input-oriented	CRS	Public & private medical centers
34	Vitikainen et al.	2009	Finland	40	1	2	Input-oriented	VRS & CRS	Public acute care hospitals
35	Puenpatom & Rosenman	2008	Thailand	92	5	5	Input-oriented	VRS	Provincial public hospitals
36	Prior	2006	Spain	29	4	5	Output-oriented	CRS	Public healthcare network hospitals
37	Butler & Li	2015	United States	57	4	4	Input-oriented	VRS	Rural state hospitals
38	Mitropoulo	2015	Greece	117	4	2	Output-oriented	VRS	Public hospitals

	s et al.								
39	Mitropoulos et al.	2013	Greece	96	4	5	Input-oriented	VRS & CRS	MoH public general hospitals
40	Mehrtak et al.	2014	Iran	18	4	3	Input-oriented	VRS	Provincial general hospitals
41	Linh Pham	2011	Vietnam	101	2	3	Input-oriented	VRS	MoH hospitals
42	Lindlbauer et al.	2016	Germany	749	7	1	Input-oriented	VRS	Mixed ownership acute care hospitals
43	Lee et al.	2008	Korea	106	3	2	Input-oriented	CRS	Mixed-ownership acute care hospitals
44	Khushalani & Ozcan	2017	United States	1259	3	4	Input-oriented	CRS	General medical-surgical hospitals
45	Kawaguchi et al.	2014	Japan	112	10	4	Input-oriented	VRS & CRS	Municipal hospitals
46	Friesner et al.	2008	United States	80	3	4	Input-oriented	VRS & CRS	General, mixed-ownership acute-care hospitals
47	Abo El-Seoud	2013	Saudi Arabia	20	4	4	Input- and output-oriented	VRS & CRS	Privately managed public hospitals

© 2023The Authors

The final meta-dataset contained a total of 4,217 hospitals (pooled sample size) spanning a 19-year period (2002-2021), categorized according to frontier-based study characteristics, DEA model specifications, and estimated mean technical efficiency (MTE). Additionally, studies using panel data that were already pooled and reported MTE as weighted averages from across the study period, their sample size therefore included number of "observations" accounted for over the years and not necessarily individual numbers of hospital units. Reported efficiency scores were assessed with consideration of the different estimates or measures of efficiency, including overall technical efficiency, pure technical efficiency, and scale efficiency, with a primary focus on technical efficiency (TE) scores of the reviewed studies. Since TE is provided by the CRS model while also capturing both pure technical efficiency (PTE) and scale efficiency (SE), whereas the VRS returns to scale model captures PTE devoid of SE effects, studies typically will apply the most relevant model best suited to address research questions and provide the efficiency score of interest; both models can also be pursued in a given study according to research objectives.

Nonetheless, the average TE score being evaluated and recorded for each reviewed study specifically refers to the CCR (CRS) technical efficiency model as described by [23] and based on the extended BCC (VRS) pure technical efficiency model developed by [24] shown in Equation 1. This formula illustrates the fundamentals of CCR and BCC models that follow the assumptions of CRS and VRS technology; whereby the CRS score can be further decomposed into a VRS score and an estimate of scale efficiency, or more often in practice, the scale efficiency (SE) is determined as a quotient (or as a fraction or a ratio in the case of proper division) when dividing technical efficiency (TE) by pure technical efficiency (PTE), or the dividend TE/PTE the divisor.

$$CRS\ score = VRS\ score \times Scale\ efficiency \qquad (1)$$
$$TE = PTE \times SE$$

For the choice of input- or output-orientation, only two studies decided on a mixed orientation approach (i.e., both input- and output-oriented models), in which output orientation was used for sensitivity analysis [25, 26]. A clear majority of studies had selected the input-oriented approach (74.5%) based on the argument that hospitals (especially government or state-funded public hospitals) cannot choose their level of output, which depends on demand for health services. Hospitals then try to conserve inputs, which makes input (or cost) minimization a reasonable assumption for DEA estimation. Some countries have different methods of financing health service providers: instead of payment based on cost history or per diem, reimbursement for hospitals is based on output volume and sector average cost with a cap (global budget). The assumption of maximizing the output level, given the amount of health resources available, has been chosen in those studies to reflect this change. Overall, the distribution of studies using CRS versus VRS assumptions to estimate efficiency scores is unevenly spread out, where the highest proportion of studies favored using a combination of both VRS & CRS models (42.6%),

© 2023The Authors

followed by the VRS model only (38.3%) and last, the CRS model only (19.1%). Technical efficiency (TE) scores, classified by model orientation and returns to scale (under CRS and/or VRS technologies), are presented in Table 3.

Table 3: *Technical efficiency estimates by choice of returns to scale and model orientation of studies*

	Technical Efficiency (TE) Scores				
DEA Model Specifications	**Mean**	**Std. Dev.**	**Median**	**Min**	**Max**
Input-oriented (n=35)	0.810	0.1046	0.8154	0.52	0.989
Output-oriented (n=10)	0.748	0.1373	0.7716	0.476	0.96
Mixed orientation (n=2)	0.902	0.079	0.902	0.846	0.958
CRS (n=9)	0.8007	0.1256	0.8	0.52	0.96
VRS (n=18)	0.7947	0.1187	0.8045	0.584	0.989
VRS & CRS (n=20)	0.8069	0.1105	0.8027	0.476	0.96

Note. TE scores are shown on a 0-1 scale. Model orientation (input- and/or output-oriented) and returns to scale (VRS and/or CRS) are not restricted to one model choice or a specific approach in frontier analysis. Technical efficiency scores are grouped by choice of returns to scale (RTS) orientation choice due to this very fact; please see meta-regression results and interpretation for more information.

The number of variables contains all inputs and outputs included in the model (dimension). In general, the number of input/output variables included in the model is thought to depend on the sample size; the rule of thumb believed is that 2 or 3 times the sum of input/output variables should be less than the sample size (number of hospital observations). The sample size is generally the number of individual hospitals included in the primary study. In effect, we can assume that a larger sample size and a lower number of input and output variables in a study will be associated with lower efficiency scores since not enough variables are accounted for; however, with proper weighted adjustments to hospital data and suitable model choices, there may not be any such constraints besides the restrictions of the number of variables used in modeling analysis for small sample size studies. Table 4 below shows some descriptive statistics of model specifications recorded from the reviewed studies.

Table 4: *Descriptive statistics of prominent study features*

	Sample size	**No. of inputs**	**No. of outputs**
Mean	89.723	3.55	3.26

© 2023The Authors

Median	40	3	3
Min	5	1	1
Max	1259	10	8

The most common inputs were capital-based (number of hospital beds, etc.) and labor-based variables (counts of human resources and hospital workforce, i.e., number of different specialists, clinicians, allied health professionals, other medical and nonmedical staff). Several output variables were centered around healthcare activities and direct patient services (i.e., number of outpatient visits, discharges, and inpatient services received). The pooled estimate of mean TE was 0.803 (±0.114). This suggests that hospitals could improve their performance by approximately 19.7%.

Although addressed by previous studies in the literature [1, 6, 27, 28], we still wish to repeat and further push a few points here since it is one of the difficulties in developing a productivity model and in preparation of the data. In addition to managerial reasoning for the selection of input and output factors, the computational and data aspects of this selection process are unclear among all 47 reviewed studies. Typically, the choice and the number of inputs and outputs and the (hospital) DMUs determine how good a discrimination exists between efficient and inefficient units. There are two conflicting considerations when evaluating the sample size. One consideration is to include as many DMUs as possible because with a larger population, there is a greater probability of capturing high-performance units that would determine the efficient frontier and improve discriminatory power [29]. However, the other conflicting consideration with a large sample size is that the homogeneity of the dataset may decrease, meaning that some exogenous impacts beyond our control have the potential to affect the final results [30].

An interesting trend emerged when we compared technical efficiency TE scores reported from high-income versus lower to upper middle-income economies, as well as by hospital efficiency studies from developed versus developing countries in Table 5 below. The reported scores in the reviewed studies tell us that on average, hospitals in developing countries are much less efficient than those in the developed world, with approximately 22.2% versus 16.6% inefficiency. The story changes after the mean technical efficiency TE estimates are adjusted by country income levels; introducing high-income economies not necessarily deemed fully developed in terms of development index and usually classified as developing or emerging markets (Gulf Arab states). The difference is now not thus far behind between high-income and lower to upper middle-income countries, with only a 3.2% difference being observed.

The hypothesis is that this large change is primarily a consequence of developing country studies having access to datasets with sample sizes and variables that are smaller relative to developed country studies, suggesting that developing country studies construct DEA efficiency models based on the availability of observations and hospital-level data and not based on reliability or accuracy considerations. The comparison adjusted by income level has shifted the efficiency studies from the Gulf Cooperation Council (GCC) member states to the high-income group with the developed majority European Organization for Economic Cooperation and Development (OECD) countries that were previously grouped separately based on development; this directly points to the weakness in regional study methodology. However, it must be emphasized that comparisons of mean efficiencies across countries (or across any groups) can be misleading unless a single reference frontier is used.

Table 5: *Hospital technical efficiency in high-income and developed countries*

Technical efficiency TE score	Developed countries	Developing countries	High-income economies	Lower to upper middle-income economies
Mean	0.834	0.778	0.817	0.785
Median	0.855	0.79	0.846	0.79
Std. Dev.	0.108	0.114	0.124	0.104
Min	0.52	0.476	0.476	0.584
Max	0.96	0.989	0.96	0.989

Note. Income level determined by GNI per capita (calculated using the atlas method) definition from the World Bank; country development index based on the World Trade Organization (WTO) classification threshold. High-income economies are not necessarily developed countries.

A noteworthy theory cited in some studies is that construction of the DEA efficiency model is less about the number of hospital variables included but rather the broad range of input/output variables accounted for in the efficiency frontier analysis. The logic for this focus on scope of coverage is based on theoretical foundations of nonparametric methods introduced by [31, 32]; the argument is that hospitals are complex organizations of production and should not be treated like other frontier firms in an industry. Thus, attempts are made to ensure that hospital variables used in efficiency models closely mirror the resource intensity of procedures in healthcare delivery units [33]. This raises the issue concerning the aggregation of variables. Constraints on degrees of freedom and zero (0) values in some variables (not missing data) usually lead to aggregation of variables. In most studies, the leading human resources of two primary labor categories of doctors and nurses are produced by aggregating many subcategories of very different skill levels, ranging from junior trainees to specialists or directors of nursing, without much weight adjustment. Aggregation of administrative, allied health professionals and other nonmedical staff or hospital workers is another common practice [34-36].

On the output side, episodes and procedures in healthcare usually differ from one patient to the other, and aggregation is generally needed to reduce the number of outputs. Since the development of case-mix systems that consider the differences in resource consumption for various types of treatments, studies have been using case-mix information to aggregate outputs, often from more than several hundred output categories into one or perhaps two outputs. Many other analyses, most notably studies from developing countries, including high-income countries in the Arabian Gulf, struggle from data deficiency or limited data availability and often use raw counts (or unweighted aggregation) of the total number of inpatient and outpatient service events. This, unsurprisingly, can potentially lead to significantly biased results when certain healthcare units provide complicated case-mix services [37-39].

Univariate and Multivariate Analyses

This section introduces the two main types of methods employed in our empirical evaluation of the literature: univariate and multivariate analyses. The practical application of multivariate statistics to a particular problem may require several types of univariate and multivariate analyses to fully understand the relationships between variables and their relevance to the problem being studied. That said, additional calculations in the analysis, such as estimating weighted averages of pooled TE scores from panel data studies or using simple hypothesis testing such as the independent-samples t test to compare estimated mean TE, among several others, are not detailed in the methodology but mentioned if applied to estimates or displayed in results. In the univariate analysis, mean technical efficiency TE estimates were compared using Wilcoxon's rank-sum test based on different features from the reviewed studies. Also known as the Mann–Whitney two-sample statistic or Mann–Whitney U test, this nonparametric analysis applies to unmatched data and was used to test the hypothesis on whether two independent samples are likely to derive from the same population with the same distribution (i.e., if the two populations have the same shape) [40, 41].

For the multivariate statistical analysis, the dependent variable is the mean efficiency score expressed as a percentage with average TE values now on a scale between 0 and 100. The transformation of efficiency estimates to percentages has no real impact on results and is simply for ease of interpretation. Other considerations in the meta-regression analysis included explanatory variables, such as the dimension regressor, which we expect to have a positive impact on efficiency estimates, while sample size is the opposite. Their effects are likely to be nonlinear and diminish when the dimension and the sample size increase. Among the many functional forms that appear to suit this expectation (quadratic, trans log, etc.), the linear-log model, as indicated by R-squared and adjusted R-squared, proved more effective in capturing the positive and diminishing marginal impact on efficiency estimates we expect as the dimension (number of inputs and outputs) increases and the marginal effect eventually turns negative.

Furthermore, the linear-log mathematical model takes the form of a function whose logarithm equals a linear combination of the parameters of the model, which makes it possible to perform multivariate linear regressions [42-44]; therefore, it was chosen as an ideal candidate for the meta-regression of the reviewed literature. Exogenous variables included in the meta-regression were chosen based on approaches and model specifications in the primary studies, including dimension variables of the frontier model (consisting of inputs, outputs and control variables such as case-mix), sample size, dummy variables to capture the type of data used (cross-section versus pooled panel data) as well as heterogeneity in the sample (lack of homogeneity in terms of hospital type, activity, and ownership), orientation (input versus output), and other explanatory variables such as model specifications (CRS versus VRS technologies) and accessibility factors that may impact the availability of reliable data (developed versus developing countries). Explanatory variables used to explain efficiency (in the one-stage or two-stage estimation approaches) were not included in the count because they do not alter the dimensions of the production space. Again, studies that pool the panel data to construct one frontier instead of estimating a separate frontier for each year will be considered as having a sample size based on the total number of observations that is usually equal to the number of individual hospitals multiplied by the number of years for balanced panels.

Our primary aim is to examine the consistency of efficiency estimates and the effect of model selection on the final technical efficiency score. According to the literature, the number of input and output variables (dimension) in our analysis is expected to have a positive impact on efficiency estimates, whereas sample size is the opposite (please see Table 1); their effects are likely to be nonlinear and diminishing when the sample size and the model dimension increase [7, 45]. Additionally, larger sample sizes and lower numbers of input and output variables included in frontier models are thought to be associated with lower efficiency scores [46]. As such, conducting a rigorous systematic literature review followed by a meta-regression is crucial to statistically identify significant factor(s) of influence in DEA models and average hospital efficiency scores using a diverse dataset of reviewed studies.

Results and Discussion

Variables included in the analysis were chosen based on study approaches and model specifications expected to impact estimated mean efficiency. Table 6 shows Wilcoxon's rank-sum test of average TE scores by variables used in the analysis and by variable subgroups falling above or below the median. This two-sample Wilcoxon rank-sum (Mann–Whitney) test simply tests the null hypothesis (H_o): mean efficiency percent (variable subgroup = = 0) = mean efficiency (variable subgroup = = 1).

Table 6: *Rank-sum mean technical efficiency estimates by variables in the analysis, divided based on median value*

Variable	N	Mean TE (Std. Dev.)	P Value
No. of hospitals			0.120

≤ 40	25	82.82 (±10.51)	
> 40	22	77.45 (±12.05)	
Orientation			0.259
Input	37	81.534 (±10.462)	
Output	10	81.18 (±10.19)	
No. of input & output variables			0.360
≤ 6	23	79.32 (±11.38)	
> 6	24	81.25 (±11.70)	
Data sample			0.116
Panel	20	83.58 (±9.53)	
Cross-section	27	77.88 (±12.32)	
Returns to scale			0.607
CRS	29	80.83 (±11.38)	
VRS	18	79.47 (±11.87)	
Homogeneity			0.264
Yes	23	79.54 (±9.52)	
No	24	81.05 (±13.22)	
Country development status			0.077
Developed economies	19	83.44 (±10.84)	
Developing/emerging economies	28	78.18 (±11.57)	
Sample size/dimension ratio			0.083
< 3	7	87.26 (±6.76)	
≥ 3	40	79.09 (±11.73)	
Overall total	47	80.30 (±11.4)	--

Eight (8) variables are specified to capture the various frontier efficiency model options discussed above. Most studies incorporate approximately three (3) to four (4) input and output variables; notable exceptions include [47] with 10 input variables, [48] using eight (8) output variables, [49] analyzing a sample size of 1,259 observations, and [25] applying a mixed orientation model (both input- and output-oriented approaches). Detailed variable descriptions are presented in Table 7.

Table 7: *Regression variables and definitions*

Variable Name	Label	Variable Definition
TE	Technical efficiency score	Reported average technical

		efficiency scores (0-100 scale)
SIZE	Number of observations	Number of (hospital) observations included in the reviewed studies
DIMENSION	Number of variables	Total number of inputs, outputs, & control variables included in the frontier model (this does not include control variables in second stage analyses)
INPUT_ORT	Orientation dummy	Dummy 0/1 variable takes value of 1 if input-oriented (including if mixed input-output- orientation), and 0 otherwise (output-oriented only)
CRS	Returns to scale	Returns to scale can either be variable or constant returns to scale; dummy variable takes value of 1 if CRS (including both CRS & VRS mix), and 0 otherwise (VRS only)
PANEL	Pooled panel data	This attempts to capture any effects of using pooled panel data instead of cross-sectional data for efficiency frontier construction; dummy variable takes value of 1 if pooled panel data study, and 0 otherwise
HOMOGENEITY	Sample homogeneity in (hospital) ownership status	Efficiency frontier units must be comparable & adjusted for heterogeneity; this dummy variable takes value of 1 if same ownership type in sample, and 0 otherwise
DEVELOPED	Efficiency studies using (hospital) data and/or published from industrialized	Dummy 0/1 variable takes value of 1 if classified “developed” by WTO (more

© 2023The Authors

	countries with economically developed markets NOT BASED ON HIGH-INCOME	advanced postindustrial economies with advanced technological infrastructure & high quality of life), and 0 otherwise *If no data available, IMF reference of $20,000 in 2021 USD nominal GDP per capita used

The dependent variable in the meta-regression is the reported average TE score on a continuous scale of 0-100. Apart from the two variables of sample size and dimension of input/output variables, the rest are dummies that explain different methodological choices. Heterogeneity in sample observations in terms of hospital type, ownership, activities, and level of care, among others, has been associated with higher efficiency scores if no adjustments are applied to homogenize hospital units [50]. Hospital ownership type is included as an additional explanatory variable since failure in accounting for heterogeneity across units of a frontier can affect estimated efficiency scores [51, 52]. Last, many studies estimate frontier models using panel data in a cross-sectional manner (i.e., they pool the panel to construct one frontier instead of estimating a separate frontier for each year). It is expected that a pooled panel sample has less variation than a cross-sectional sample because one hospital will be observed more than once; thus, variation from year to year is expected to be smaller than variation between different hospitals [53]. This can potentially produce higher average efficiency scores (please see Table 1). Therefore, a dummy variable is included in our regression to capture any differences and account for the type of data used (cross-sectional versus pooled panel data).

Table 8 contains some descriptive statistics of the dependent and explanatory variables. The pooled mean TE was 80.30 (±11.4), with the highest being 98.9 and the lowest 47.6. Interestingly, this considerable variation in efficiency scores comes from studies in the Middle East with similar variable measures being used to estimate the frontier, as well as hospital activity data included in the analysis; however, the model specifications clarify the distinction: (i) heterogeneity in type of hospital, ownership status, and hospital activities (sample included both teaching and nonteaching hospitals; teaching and research activities of university hospitals were not accounted for and other general differences in ownership and hospital management were poorly handled), pooled panel data, input-oriented, VRS, 3 inputs (dimension), 12 hospitals (size) = 98.9 average TE score; and (ii) heterogeneity in hospital ownership status (sample included both private and government hospitals; differences were unadjusted but instead hospitals were grouped by ownership type and analyzed separately then merged unweighted efficiency frontiers for comparison), cross-sectional, output-oriented, VRS & CRS, 6 inputs (dimension), 96 hospitals (size) = 47.6 average TE score [37, 54]. This is a striking example of how the choice of models, input/output variables, and quality control adjustments can significantly alter efficiency estimates and study robustness, which leads one to question the degree to which policy can be influenced by this type of performance indicator, and if indeed basic measures are done correctly, what can be drawn from hospital efficiency studies?

Table 8: *Descriptive statistics of modeling choices in estimating the production possibility frontier*

Variable	Mean	Median	Std. Dev.	Min	Max
TE	80.30	80	11.41	47.60	98.90
SIZE	89.723	40	205.046	5	1,259
DIMENSION	6.809	6	2.223	3	14
INPUT_ORT	0.787	1	0.414	0	1
CRS	0.617	1	0.491	0	1
PANEL	0.426	0	0.410	0	1
HOMOGENEITY	0.489	0	0.505	0	1
DEVELOPED	0.404	0	0.496	0	1

Note. Efficiency scores (TE) are shown on a 0-100 percentage scale, similar to how mean TE estimates were applied in the meta-regression and analyzed as dependent variables.

The choice of functional form is driven by the possible impacts of the two continuous variables: dimension and sample size. Dimension is expected to have a positive impact on efficiency estimates, while sample size is the opposite. Their effects are likely to be nonlinear and diminish when the dimension and the sample size increase. The functional form that appears to suit this expectation is the linear-log model. Described by [55] and [17] and recommended by [45], the linear-log function is used in the following estimation with the specifications as follows:

$$\text{MTE} = \beta 0 + \beta 1 \ln(\text{Size}) + \beta 2 \ln(Dimension) + \beta 3(input - oriented) + \beta 4(CRS) + \beta 5(Panel) + \beta 6(Homogeneity) + \beta 7(Development) \tag{2}$$

where MTE is the mean technical efficiency TE. The marginal effect of dimension on efficiency estimates is expressed by the partial derivative or differential below in Equation 3:

$$\frac{\partial \text{MTE}}{\partial \text{Dimension}} = \beta 1 \frac{1}{Dimension} \tag{3}$$

The marginal effect of sample size on efficiency is also expressed by Equation 4:

© 2023The Authors

$$\frac{\partial \text{MTE}}{\partial \text{Siz}} = \beta 2 \frac{1}{Size} \tag{4}$$

Consequently, when the dimension increases, a positive β1 will ensure that the marginal effect approaches zero but does not turn negative. The opposite happens to size; a negative value of β2 allows the marginal effect of size on efficiency to approach zero from below as sizes increase [45].

Ordinary least squares (OLS), a type of linear least squares method for estimating the unknown parameters in a linear regression model, was used for this model since ordinary least squares regression is considered a consistent estimator [56]. That said, it is not necessary for us to use Tobit or limited dependent variable procedures, which are usually used when the dependent variable is bounded; there is no mean efficiency of 0 or 1 (or 100 in the case of percentage scale) in the meta-data; therefore, Tobit estimates are identical to their OLS counterparts. Meta-regression was analyzed in Stata/SE 16.1 (StataCorp LLC, College Station, TX), and a correlation matrix identified the absence of multicollinearity between the independent variables. The regression results are displayed in Table 9.

Table 9: *Results of meta-regression analysis*

Variables	**Coef. (Std. Err.)**
Ln (SIZE)	-4.072562 *** (1.573681)
Ln (DIMENSION)	6.798539 ** (1.186624)
INPUT_ORT	5.278795 (3.780669)
CRS	-0.9297522 * (3.164268)
PANEL	4.573849 * (1.127381)
HOMOGENEITY	-1.058234 ** (1.203779)
DEVELOPMENT	7.999337 ** (3.421742)
Constant	89.03749 *** (6.434255)

© 2023The Authors

F-statistics	3.5657
R-squared	0.3902
Adjusted R-squared	0.3615

Note. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The estimated coefficient for SIZE, capturing the effect of sample size on mean efficiency, is negative, while that for DIMENSION, the regressor that represents the influence of the number of input and output variables on efficiency, is positive. Both sample size and dimension are significant at the 1% and 5% levels, respectively, and in line with expectations according to the reviewed literature. The negative sign of the coefficient for SIZE indicates that, holding everything else constant or all other variables equal, increasing the number of hospital observations will yield a lower mean technical efficiency score. For example, when evaluated at the median sample size of 40 hospital observations, the marginal effect of SIZE is -0.102; however, the marginal effect is larger at smaller sample sizes. Upon evaluating a sample size of 30 hospital observations, the marginal effect now becomes -0.136; looking at a sample size of 20, for instance, the marginal effect is larger at -0.204, suggesting that the addition of observations could lead to a reduction in mean efficiency.

The effect of DIMENSION on the average efficiency score is more substantial. The marginal effect is 1.133 when evaluated at the sample median of 6 variables. However, as the number of variables decreases, the marginal effect is larger. For example, a value of 3 variables results in a marginal effect of 2.266, suggesting that the addition of extra variables could lead to an increase in mean efficiency. These larger effects at low SIZE and DIMENSION values seem to statistically show relevant DEA assumptions cited in the literature. Based on the meta-regression analysis, it is clearer that as the number of variables included in the frontier model increases, the average efficiency predictions drop fairly rapidly when the sample size is small. The author(s) of reference [16] also arrived at a similar conclusion on the sample size effect, further showing how the inclusion of an extra variable into a model with more than 10 (hospital) observations does not alter the average efficiency score very much. They observed that when the sample size was large, the mean technical efficiency showed little change, and the mean efficiency seemed to remain constant after a threshold. Therefore, correcting for sample size has a major impact on the assessment of average efficiency estimates [15].

As expected, although significant only at the 10% level, the coefficient for the PANEL variable produces a positive sign, suggesting that the use of a pooled panel tends to produce higher average efficiency scores of 4.57 percentage points. A possible explanation for this is that a single hospital is observed more than once in a pooled panel, and therefore, any variation from year to year is expected to be smaller than variations between different hospitals when cross-sectional data analysis is used. This can potentially produce higher average efficiency scores. The lesson learned from this is to ensure that separate production frontiers are created for each year in a cross-sectional manner instead of pooling the entire panel data into a single efficiency

© 2023The Authors

frontier and attempting to analyze hospitals at yearly cross-sections. Another variable found barely significant at the 10% level is the estimated coefficient for CRS, which displays a negative and $p<0.1$ significant effect on the mean efficiency score. The magnitude of the CRS coefficient implies that choosing a CRS technology over a VRS returns to scale will reduce the mean efficiency estimate by approximately one (1) percentage point.

Conclusions

While heterogeneity is assumed to be associated with higher efficiency scores or exaggerated efficiency estimates, only heterogeneity in hospital type and ownership status was statistically significant [57]. As the regression suggests, maintaining homogeneity (uniform hospital sample) is expected to reduce overestimated efficiency scores seen in heterogeneous data analysis by reducing the mean efficiency by approximately one (1) percentage point compared to nonhomogeneous hospital samples. Last, data analysis studies published from developed countries are statistically significant and reported an average of approximately 8 percentage points higher efficiency scores compared to data analysis studies published from developing/emerging countries. Among several reasons, the main findings suggest that developing countries suffer from weak studies due to aggregation of input categories, no adjustment for differences in case-mix and quality of care between hospitals, small sample size, little adjustment for heterogeneity in hospital sample, and no attempt to evaluate the misspecification in applied models. This raises issues of validity, usefulness, and generalizability of studies from the developing world. The lack of data is above all the main explanation for such limitations; therefore, improving data collection and processing in hospital databases of developing nations is critical in promoting quality in efficiency studies aimed at informing policy. Future efficiency analysis studies should therefore apply the proper adjustments and model considerations gleaned from the findings of this meta-regression analysis.

Declarations

Study Limitations

Not applicable/none.

Acknowledgments

Much appreciation to the graduate faculty at the School of Health Sciences and Practice at New York Medical College, along with research fellows at the Institute of Public Health, for their positive support.

Funding source

No funding source exists/none.

Competing Interests

No declaration of potential competing interests exists in this publication.

© 2023The Authors

References

[1] J. Aparicio, C. A. K. Lovell, and J. T. Pastor, *Advances in Efficiency and Productivity*, 1st ed. Cham: Springer International Publishing : Imprint: Springer,, 2016, pp. 1 online resource (VI, 415 pages 65 illustrations, 36 illustrations in color.

[2] A. Assaf and K. M. Matawie, "Improving the accuracy of DEA efficiency analysis: a bootstrap application to the health care foodservice industry," *Applied Economics,* vol. 42, no. 27, pp. 3547-3558, 2010/11/01 2010, doi: 10.1080/00036840802112497.

[3] S. Daidone and F. D'Amico, "Technical efficiency, specialization and ownership form: evidence from a pooling of Italian hospitals," *Journal of productivity Analysis,* vol. 32, no. 3, pp. 203-216, 2009.

[4] C. Arfa, H. Leleu, M. Goaïed, and C. van Mosseveld, "Measuring the capacity utilization of public district hospitals in Tunisia: Using dual data envelopment analysis approach," *International Journal of Health Policy and Management,* vol. 6, 2016//2016.

[5] E. Z. Asbu, M. D. Masri, and A. Kaissi, "Health status and health systems financing in the MENA region: roadmap to universal health coverage," *Global Health Research and Policy,* vol. 2, no. 1, p. 25, 2017/09/04 2017, doi: 10.1186/s41256-017-0044-9.

[6] G. Bevan, J. K. Helderman, and D. Wilsford, "Changing choices in health care: Implications for equity, efficiency and cost," (in eng), *Health Econ Policy Law,* vol. 5, no. 3, pp. 251-67, Jul 2010, doi: 10.1017/s1744133110000022.

[7] R. Jacobs, P. C. Smith, and A. Street, *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. Cambridge: Cambridge University Press, 2006.

[8] N. Kontodimopoulos, P. Nanos, and D. Niakas, "Balancing efficiency of health services and equity of access in remote areas in Greece," *Health Policy,* vol. 76, no. 1, pp. 49-57, 2006/03/01/2006, doi: https://doi.org/10.1016/j.healthpol.2005.04.006.

[9] P. Mitropoulos, I. Mitropoulos, and A. Sissouras, "Managing for efficiency in health care: the case of Greek public hospitals," *The European Journal of Health Economics,* vol. 14, no. 6, pp. 929-938, 2013/12/01 2013, doi: 10.1007/s10198-012-0437-0.

[10] K. Nguyen and T. Coelli, "Quantifying the effects of modeling choices on hospital efficiency measures: A meta-regression analysis," University of Queensland, School of Economics, 2009. [Online]. Available: https://EconPapers.repec.org/RePEc:qld:uqcepa:74

[11] T. Fixler, J. C. Paradi, and X. Yang, "A data envelopment analysis approach for measuring the efficiency of Canadian acute care hospitals," *Health Services Management Research,* vol. 27, no. 3-4, pp. 57-69, 2014, doi: 10.1177/0951484815601876.

[12] R. Goudarzi, N. RjabiGilan, S. R. Ghasemi, S. Reshadat, R. Askari, and M. Ahmadian, "Efficiency measurement using econometric stochastic frontier analysis (SFA) method, case study: hospitals of Kermanshah University of Medical Sciences," *J Kermanshah Univ Med Sci,* vol. 17, 2014//2014.

[13] M. D. Rosko, "Impact of internal and external environmental pressures on hospital inefficiency," (in eng), *Health Care Manag Sci,* vol. 2, no. 2, pp. 63-74, May 1999, doi: 10.1023/a:1019031610741.

© 2023The Authors

[14] H. S. Brown, 3rd, "Managed care and technical efficiency," (in eng), *Health Econ,* vol. 12, no. 2, pp. 149-58, Feb 2003, doi: 10.1002/hec.712.

[15] R. Fare, S. Grosskopf, M. Norris, and Z. Zhang, "Productivity growth, technical progress, and efficiency change in industrialized countries," *American Economic Review,* vol. 84, no. 1, pp. 66-83, 1994.

[16] B. Sun, L. Zhang, W. Yang, J. Zhang, D. Luo, and C. Han, "Data envelopment analysis on evaluating the efficiency of public hospitals in Tianjin, China," *Transactions of Tianjin University,* vol. 22, no. 2, pp. 182-188, 2016.

[17] T. J. Coelli, D. S. P. Rao, C. J. O'Donnell, and G. E. Battese, *An introduction to efficiency and productivity analysis*. springer science & business media, 2005.

[18] J. F. Burgess and P. W. Wilson, "Hospital Ownership and Technical Inefficiency," *Manage. Sci.,* vol. 42, no. 1, pp. 110–123, 1996.

[19] R. Webster and S. Johnson, "Comparing Techniques for Measuring the Efficiency and Productivity of Australian Private Hospitals," 1998.

[20] V. J. M. Cantor and K. L. Poh, "Integrated Analysis of Healthcare Efficiency: A Systematic Review," *Journal of Medical Systems,* vol. 42, no. 1, p. 8, 2017/11/22 2017, doi: 10.1007/s10916-017-0848-7.

[21] R. Liberati et al., "1Guidelines and Guidance The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration," 2009.

[22] C. Mitton, C. Adair, E. McKenzie, S. B. Patten, and B. W. Perry, "Knowledge transfer and exchange: Review and synthesis of the literature," *Milbank Quarterly,* vol. 85, 2007//2007.

[23] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research,* vol. 2, 1978//1978.

[24] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management science,* vol. 30, no. 9, pp. 1078-1092, 1984.

[25] F. Farzianpour, S. Hosseini, T. Amali, S. Hosseini, and S. S. Hosseini, "The evaluation of relative efficiency of teaching hospitals," *Am J Appl Sci,* vol. 9, 2012//2012.

[26] M. S. Abou El-Seoud, "Measuring efficiency of reformed public hospitals in Saudi Arabia: An application of data envelopment analysis," *International Journal of Economics and Management Sciences,* vol. 2, no. 9, pp. 44-53, 2013.

[27] D. Adham, B. Issac, G. Sadeghi, P. Mohammad, A. Hossein, and E. Salarkhah, "Contemporary use of hospital efficiency indicators to evaluate hospital performance using the Pabon Lasso model," *Eur J Bus Soc Sci,* vol. 3, 2014//2014.

[28] M. Bahadori, "Combining multiple indicators to assess hospital performance in Iran using the Pabon Lasso Model," *Australasian Medical Journal,* vol. 4, 2011//2011.

© 2023The Authors

[29] I. Contreras, "A review of the literature on DEA models under common set of weights," *Journal of Modeling in Management,* vol. 15, no. 4, pp. 1277-1300, 2020, doi: 10.1108/JM2-02-2019-0043.

[30] M. Ghahremanloo, A. Hasani, M. Amiri, M. Hashemi-Tabatabaei, M. Keshavarz-Ghorabaee, and L. Ustinovičius, "A novel DEA model for hospital performance evaluation based on the measurement of efficiency, effectiveness, and productivity," 04/28 2020, doi: 10.2478/emj-2020-0001.

[31] A. Worthington, "Frontier Efficiency Measurement in Health Care: A Review of Empirical Techniques and Selected Applications," *Medical care research and review : MCRR,* vol. 61, pp. 135-70, 07/01 2004, doi: 10.1177/1077558704263796.

[32] A. C. Worthington, "Frontier Efficiency Measurement in Healthcare : A Review of Empirical Techniques and Selected Applications," 2009.

[33] A. Worthington, "An empirical survey of frontier efficiency measurement techniques in education," *Education Economics,* vol. 9, 10/20 1999.

[34] A. D. Alatawi, L. W. Niessen, and J. A. M. Khan, "Determinants of Technical Efficiency in Public Hospitals: The Case of Saudi Arabia," (in eng), *Health Econ Rev,* vol. 10, no. 1, p. 25, Aug 1 2020, doi: 10.1186/s13561-020-00282-z.

[35] H. Li and S. Dong, "Measuring and Benchmarking Technical Efficiency of Public Hospitals in Tianjin, China

A Bootstrap–Data Envelopment Analysis Approach," *Inquiry,* vol. 52, pp. 1-5, 2015.

[36] A. M. Alsabah, H. Haghparast-Bidgoli, and J. Skordis, "Measuring the efficiency of public hospitals in Kuwait: A two-stage data envelopment analysis and a qualitative survey study," *Econometrics: Econometric & Statistical Methods - Special Topics eJournal,* 2019.

[37] "Mahate A, Hamidi S. Frontier efficiency of hospitals in United Arab Emirates: an application of data envelopment analysis. J Hosp Adm. 2015. https://doi.org/10.5430/jha.v5n1p7."

[38] R. Ramanathan, "Operations assessment of hospitals in the Sultanate of Oman," *International Journal of Operations & Production Management,* vol. 25, 2005//2005.

[39] R. Kalhor, S. Amini, M. Sokhanvar, F. Lotfi, M. Sharifi, and E. Kakemam, "Factors affecting the technical efficiency of general hospitals in Iran," *Journal of the Egyptian Public Health Association,* vol. 91, 2016 2016.

[40] F. Wilcoxon, *Biometrics Bulletin,* vol. 1, no. 6, p. 80, 1945.

[41] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics,* vol. 18, no. 1, p. 50, 1947, doi: 10.1214/aoms/1177730491.

[42] G. Jandaghi, H. Z. Matin, M. Doremami, and M. Aghaziyarati, "Efficiency evaluation of Qom public and private hospitals using data envelopment analysis," *Eur J Econ Finance Adm Sci,* vol. 22, 2010//2010.

© 2023The Authors

[43] A. C. Worthington, "Frontier Efficiency Measurement in Health Care: A Review of Empirical Techniques and Selected Applications," *Medical Care Research and Review,* vol. 61, no. 2, pp. 135-170, 2004, doi: 10.1177/1077558704263796.

[44] S. Kohl, J. Schoenfelder, A. Fügener, and J. O. Brunner, "The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals," *Health care management science,* vol. 22, no. 2, pp. 245-286, 2019.

[45] A. A. Kiadaliri, M. Jafari, and U. G. Gerdtham, "Frontier-based techniques in measuring hospital efficiency in Iran: a systematic review and meta-regression analysis," (in eng), *BMC Health Serv Res,* vol. 13, p. 312, Aug 15 2013, doi: 10.1186/1472-6963-13-312.

[46] B. Hollingsworth, "Non-Parametric and Parametric Applications Measuring Efficiency in Health Care," *Health Care Management Science,* vol. 6, no. 4, pp. 203-218, 2003/11/01 2003, doi: 10.1023/A:1026255523228.

[47] H. Kawaguchi, K. Tone, and M. Tsutsui, "Estimation of the efficiency of Japanese hospitals using a dynamic and network data envelopment analysis model," *Health Care Management Science,* vol. 17, no. 2, pp. 101-112, 2014/06/01 2014, doi: 10.1007/s10729-013-9248-9.

[48] T. Jat and M. Sebastian, "Technical efficiency of public district hospitals in Madhya Pradesh, India: a data envelopment analysis," (in eng), *Glob Health Action,* vol. 6, p. 21742, Sep 24 2013, doi: 10.3402/gha.v6i0.21742.

[49] J. Khushalani and Y. A. Ozcan, "Are hospitals producing quality care efficiently? An analysis using Dynamic Network Data Envelopment Analysis (DEA)," *Socio-Economic Planning Sciences,* vol. 60, pp. 15-23, 2017/12/01/2017, doi: https://doi.org/10.1016/j.seps.2017.01.009.

[50] B. Hollingsworth, & Peacock, S. J., *Efficiency Measurement in Health and Health Care.*, 1st ed. ed. New York: Routledge, 2008.

[51] C. Jehu-Appiah et al., "Ownership and technical efficiency of hospitals: evidence from Ghana using data envelopment analysis," *Cost Effectiveness and Resource Allocation,* vol. 12, no. 1, p. 9, 2014/04/08 2014, doi: 10.1186/1478-7547-12-9.

[52] J. M. Kirigia, E. Z. Asbu, W. Greene, and A. Emrouznejad, "Technical efficiency, efficiency change, technical progress and productivity growth in the national health systems of continental African countries," *Eastern Africa Social Science Research Review,* vol. 23, no. 2, pp. 19-40, 2007.

[53] A. Hoff, "Second stage DEA: Comparison of approaches for modeling the DEA score," *European journal of operational research,* vol. 181, no. 1, pp. 425-435, 2007.

[54] A. Torabipour, M. Najarzadeh, M. Arab, F. Farzianpour, and R. Ghasemzadeh, "Hospitals productivity measurement using data envelopment analysis technique," (in eng), *Iran J Public Health,* vol. 43, no. 11, pp. 1576-81, Nov 2014.

[55] K.-H. Nguyen and T. Coelli, "Quantifying the effects of modeling choices on hospital efficiency measures: a meta-regression analysis," School of Economics, University of Queensland, Australia, 2009.

© 2023The Authors

[56] R. Kieschnick and B. D. McCullough, "Regression analysis of variates observed on (0, 1): percentages, proportions and fractions," *Statistical modeling,* vol. 3, no. 3, pp. 193-213, 2003.

[57] Y. A. Ozcan, *Health care benchmarking and performance evaluation*. Springer, 2008.

© 2023The Authors